Journal of Healthcare Engineering

Multiple Criteria Decision-Making Approaches for Healthcare Management Applications

Lead Guest Editor: Hao-Chun Lu
Guest Editors: Hsiao-Ting Tseng, Chi-Hua Chen, and Chia-Yu Lin

Multiple Criteria Decision-Making Approaches for Healthcare Management Applications

Journal of Healthcare Engineering

Multiple Criteria Decision-Making Approaches for Healthcare Management Applications

Lead Guest Editor: Hao-Chun Lu
Guest Editors: Hsiao-Ting Tseng, Chi-Hua Chen, and Chia-Yu Lin

Contents

Contents

Hindawi
Journal of Healthcare Engineering
Volume 2023, Article ID 5934523, 22 pages
https://doi.org/10.1155/2023/5934523

Research Article

Integrating Fuzzy Multiobjective Programming and System Dynamics to Develop an Approach for Talent Retention Policy Selection: Case on Health-Care Industry

Chia-Lun Lo[1] **and Ya-Han Hu**[2]

[1]*Department of Health-Business Administration, Fooyin University, Kaohsiung 831301, Taiwan*
[2]*Department of Information Management, National Central University, Taoyuan 320317, Taiwan*

Correspondence should be addressed to Chia-Lun Lo; allenlo.tw@gmail.com

Received 20 May 2021; Revised 2 May 2022; Accepted 25 November 2022; Published 18 February 2023

Academic Editor: Chia-Yu Lin

The demand for medical services has been increasing yearly in aging countries. Medical institutions must hire a large number of staff members to provide efficient and effective health-care services. Because of high workload and pressure, high turnover rates exist among health-care staff members, especially those in nonurban areas, which are characterized by limited resources and a predominance of elderly people. Turnover in health-care institutions is influenced by complex factors, and high turnover rates result in considerable direct and indirect costs for such institutions (Lo and Tseng 2019). Therefore, health-care institutions must adopt appropriate strategies for talent retention. Because institutions cannot determine the most effective talent retention strategy, many of them simply passively adopt a single human resource (HR) policy and make minor adjustments to the selected policy. In the present study, system dynamics modeling was combined with fuzzy multiobjective programming to develop a method for simulating HR planning systems and evaluating the suitability of different HR policies in an institution. We also considered the external insurance policy to be the parameter for the developed multiobjective decision-making model. The simulation results indicated that reducing the turnover rate of new employees in their trial period is the most effective policy for talent retention. The developed procedure is more efficient, effective, and cheaper than the traditional trial-and-error approaches for HR policy selection.

1. Introduction

Population aging engenders changes in population structure. Therefore, in aging societies, increasing attention has been paid to issues such as medical care, medical economics, psychology, and social welfare policies. Because of the low muscle strength and flexibility of elderly people, considerable resources must be spent on their health care [1]. However, the challenges involved in caring for elderly people impose high pressure on health-care staff, which results in a high staff turnover rate. Accordingly, staff turnover is a major problem facing health-care institutions in aging countries.

Traditionally, medical service providers maintain their competitiveness by hiring professional personnel to minimize care risks as well as human errors in instrument or device operation and medication provision [2]; thus, labor constitutes more than 45% of the total cost incurred by medical institutions. Some hospitals use human resource planning (HRP) strategies involving a hiring freeze to achieve cost reduction and profits. Other hospitals even occasionally adopt the lowest standard of labor as a business strategy. However, hospitals cannot predict the number of patient visits. In addition, a sudden increase in the number of staff members leads to an increase in the burden on the original staff; this is because the original staff must mentor the new staff. Some new staff members might also make mistakes because of insufficient experience, which can negatively affect the reputation of the hospital they work for. The medical industry is a labor-intensive industry, and cost

reduction attempts involving personnel cuts can lead to reduced medical service quality and endanger patients' lives. In addition, a high number of adverse events can occur in a high-burden environment, which increases the rate of human resource (HR) loss. Therefore, cost reduction strategies in the medical industry are associated with moral dilemmas.

Medical institutions must adopt flexible strategies to achieve effective team building and avoid brain drain. In particular, medical institutions can secure their competitiveness by hiring sufficient personnel. Furthermore, when designing HRP strategies, HR planners must consider internal and external factors to ensure that develop comprehensive HR plans and system policy evaluation models from an organizational system perspective.

In the health-care field, root cause analysis (RCA) is one of the most common methods for systematic policy evaluation. RCA is based on secondary data and analysis data obtained through discussions among individuals with different subjective judgments and professional backgrounds [3]. Therefore, RCA can only produce a consensus decision instead of an objectively correct answer [4]. Moreover, the results of the root cause explanations might differ over time. Thus, it is not suitable for HRP issues. HRP is a complicated process and can be defined as a systematic analysis of HR needs to ensure that adequate numbers of employees with the required skills are available at a given time [5]. Several studies have used statistical techniques and mathematical models to forecast HR demand and supply and to execute HRP. However, because of their fundamental limitations, these techniques and models cannot be used for dynamic structural analysis and the identification of delayed feedback effects. Actions undertaken on the basis of inaccurate demand forecasts can occasionally produce results that are contrary to the intended ones. Thus, a systematic tool that can consider complex and dynamic factors must be developed for evaluating HRP policies [6].

In the system dynamics (SD) approach, features such as feedback delays and nonlinear relationships are considered for accurately determining the HR demand and supply. Accordingly, SD modeling is suitable for studying the behavior of a dynamic HRP system [6]. The system dynamics (SD) approach can deal with forecasting inaccuracies and potential mismatches and towards understanding and policy design, which use dynamic features such as feedback delays and nonlinear relationships are considered by very few modelers to deal with HR demand and supply uncertainties. In order to study the behavior of a dynamic HRP system, SD modeling is an appropriate tool.

Several uncertain factors influence the medical actions or administrative decisions taken in a health-care institution. Furthermore, such an institution can have multiple goals related to medical personnel, patients, insurance, and the institution. An SD model considers internal and external factors that affect a system and decreases the percentage of wrong decisions; such a model can thus be beneficial for such institutions. Nevertheless, administrative decisions are complex. Therefore, a fuzzy model can facilitate the task of identifying appropriate administrative decisions. In the medical field, no closed-form solution is generally available for the problem of determining appropriate administrative decisions. The most reasonable approach for administrative decision-making in the medical field involves creating a fuzzy set of multiple decision parameters to determine the optimal decision for achieving a goal. Accordingly, the present study combined SD modeling and fuzzy multiobjective programming (MOP) to develop an HRP system for evaluating the appropriateness of different HR policies in a health-care institution. The SD approach was adopted to create complex and nonlinear models to determine causal feedback relationships [6].

Several studies have adopted SD approaches that entail the consideration of an entire system for analysis. Nonetheless, according to our search of the PubMed database, no study has adopted SD modeling for analyzing nursing talent retention policies in medical institutions. Hence, the present study applied SD modeling to examine the suitability of different labor policies in medical institutions; the study also adopted a hybrid modeling approach to simulate the number of patients that can be accepted in the event of changes in the number of nursing personnel.

2. Literature Review

The medical industry is a labor-intensive service sector; thus, HRP is essential in this industry. In general, approximately 30% of the personnel in an acute hospital are nursing staff members. Nurses thus constitute the main personnel in hospital operations. In modern medical treatment procedures, nurses must provide highly professional care to patients. A lack of sufficient nursing staff for posttreatment care might lead to high risks of patient relapse. Thus, nursing staff members play crucial roles at every level in medical institutions.

In general, a higher ratio of nursing staff to patients is associated with superior patient recovery [7, 8]. However, nursing jobs are highly exhausting, and nursing staff must spend considerable time interacting with and satisfying the demands of patients or their families; these requirements impose a high burden on nursing staff members. Consequently, medical institutions experience a high turnover of nursing staff [9]. According to Waldman et al. [10], Most of the hospital staffs are nursing personnel; therefore, the turnover cost for nursing personnel accounts for approximately thirty percent of the total turnover cost. A study indicated that more than 20% of the nursing staff in hospitals had the intention of quitting their job each year in general [11]. A shortage of nursing staff can considerably increase patient mortality rates [12, 13] and in-hospital infection rates [14]. Moreover, a lack of adequate nursing staff might lead to insufficient preservice training and evaluation skills, which result in communication disruptions and thus medical adverse events, including falls [15], medication or transfusion errors [16], treatment delays, complications during or after surgery [11, 17], and prolonged hospital stays [11].

According to Chiu et al. [18], the cost of training a new registered nurse is US$15,825 in the U.S. In addition, insufficient nursing experience results in productivity reduction, the estimated cost of which is US$5,245–US$16,102; it

also engenders medical adverse events, which are also associated with a high implicit cost. Therefore, planning strategies for hiring nursing personnel constitute a crucial component of the overall labor planning process of hospitals. The insufficiency of nursing staff is a problem faced by hospitals worldwide [18].

Hospital managers must develop suitable incentives to increase nursing staff retention. However, if managers adopt only salary hikes as a strategy for nursing staff retention, their hospitals would experience high operational pressure because their expenses would increase while their income would remain limited by health insurance systems. Failure to address the problem of nursing staff turnover might lead to a deterioration in nursing quality, which would lead to patients providing poor reviews regarding the hospital. Poor reviews can result in a sharp decrease in hospital income. In Taiwan's insurance payment system, nursing expenses are included in hospital expenses. The salary covered by this system for each nursing personnel is less than $20,000. Thus, hospitals can reduce their nursing expenses only by reducing the number of nursing staff employed [19], hiring contracted staff instead of employees, adopting a working-hour-based salary system, and not recruiting new staff even when vacancies exist. Decreases in the ratio of nursing staff to patients result in staff demoralization and thus high turnover rates. Therefore, hospitals should employ a suitable number of nursing staff such that they minimize expenses while maintaining high health-care quality.

HR quality is more important than HR quantity in most of the industries [20], including the medical industry. Health-care staff members typically differ in terms of literacy, health-care capability, and service quality. Therefore, "clinical advancement systems" have been developed in the nursing industry. Continual on-the-job training can help nursing staff to improve their nursing skills, knowledge, and self-growth, thus increasing their service quality. The current clinical advancement system in Taiwan is based on Benner's theoretical structure, which comprises five stages of skill acquisition in nursing knowledge: novice, advanced beginner, competent, proficient, and expert. The purposes of this system are to help nursing staff transform from novices to experts through step-by-step learning [21, 22] and to reduce staff turnover by higher job achievement.

Managing turnover is crucial for firms because retaining the best talents is essential to remain competitive in the 21st century. Many HR managers have termed the current era as the era of the war for talent [23]. To address the problem of turnover, numerous studies have attempted to determine the key factors that drive turnover [24–27]. However, health-care staff turnover is a dynamic problem because the interrelationships between the factors that influence it to vary. Furthermore, the assumptions underpinning the models constructed in previous studies for assessing health-care staff turnover might become invalid over time. This can thus engender uncertainties regarding the impact of strategies suggested by such models. Consequently, organizations face high risk in strategy implementation, especially when the implemented strategies are ineffective. Considering these limitations, static decision-making models might be unsuitable for modeling health-care staff turnover because they cannot integrate all the variables of a real situation. The basic assumptions of decision-making models are that criteria and alternatives are fixed a priori and that a decision occurs only once; that is, a decision does not involve spatial or temporal considerations. These assumptions limit the validity of the results of the aforementioned models, especially when parameter values change over time and the decision matrix is not fixed or static. In addition, multicriteria decision-making models for talent retention policy selection focus on the cause-effect relationships between individual factors; hence, these models are not comprehensive. Multicriteria decision-making models also generally cannot provide a complete understanding of the complexity of the problem of talent retention policy selection with respect to human-related factors.

Forrester identified a deficiency in the decision-making process of complex dynamic systems [28]. To overcome this deficiency, he proposed SD modeling tools that allow the refining and simulation of mental models for different decision-making policies. Studies have applied SD to address several management problems and provide decision support for managers in many fields [29–32]. A suitable decision-making model must be able to tolerate vagueness or ambiguity because fuzziness and vagueness are common characteristics of most decision-making problems [33]. Accordingly, fuzzy logic (FL) is an essential way for an effective decision-making model [34–37]. The fuzzy theory is the most commonly used concept for solving problems related to imprecise data and ambiguous human judgments in the selection of a talent retention policy. Fuzzy set theory was created to adapt mathematical tools of logic to different types of uncertainty, such as vagueness and approximation, which are characteristics of natural language and human mental models. FL enables the representation of human knowledge through linguistic IF-THEN expressions, which are typical of approximate reasoning [38]. Moreover, FL can be used to obtain solutions to many real-world problems that involve some degree of imprecision and ambiguity, such as talent retention policy selection. Data shortage, measurement errors, or the subjectivity of human judgment can result in uncertain information exhibiting a fuzzy or stochastic nature. Fuzzy models are commonly used because they address various types of uncertainties.

Hybrid models combine different forecasting methods to estimate policy performance. The main advantage of these models is that they can combine different methods and thus leverage their advantages; nevertheless, verifying the rationality of the approach used to combine such methods is difficult. Moreover, hybrid models are characterized by a long computational time.

The first researchers to integrate FL and SD were Pankaj et al. [39], who proposed a method for the qualitative analysis of causal loops by using fuzzy linguistic uncertainties to incorporate the perceptions and beliefs of the modeler. Their motivation was based on the understanding that natural language is the optimal means of expressing the relationships between variables in human mental models. This perception is also the reason behind the development of

most hybrid models. Several studies have integrated FL into SD models so that they could consider fuzzy parameters (i.e., their relations, arithmetic, or soft descriptions) when data are unavailable or are not sufficiently credible or when certain parameters exhibit fuzziness [40–42]; these studies did not use FL to define policies to control the developed SD models but instead used it to handle the uncertainty in some model parameters. These models were controlled through a classic approach entailing the use of crisp parameters. Song et al. [43] and Orji and Wei [44] have used SD models to simulate alternative scenarios and then ranked them through MOP. Similarly, Chang and Ko [45], Xu et al. [46], and Wu and Xu [47] have used MOP to control each step of an SD model. This approach is similar to the methodology proposed in this study in that the controller is completely dissociated from the model to be controlled in both approaches. Sabounchi et al. [48] used FL to model decision rules in an SD model of users' transportation preferences.

On the basis of the preceding literature on the combination of FL with SD, we can determine that no study has combined FL with SD modeling for the selection of talent retention policies. Accordingly, the main aim of this study was to design a multiobjective FL-based SD method to handle constraints and uncertain parameters in the evaluation of talent retention policies.

3. Methodology

The SD approach is based on systems thinking and is aimed at facilitating learning tasks in complex, feedback, multiloop, multistate, and nonlinear systems in which humans live [49]. It involves using systematic thinking tools to modify mental models repeatedly for implementing reflective learning. Sterman [50] suggested that the SD approach is the best method for resolving complex and dynamic problems. An SD model that considers the interrelationships between internal and external factors can provide specific information related to a complex and dynamic problem and then provide suggestions for the perfect strategic decisions for solving the problem, which would be helpful for managers. In this study, we used the general modeling method proposed by Forrest [51]. SD models are designed to determine the structure of a complex system through the comprehension of concepts such as feedback, stocks, flows, time delays, and nonlinearity [52]. The major components of an SD model include stocks, flows, rates, and auxiliar. Time delay is the most crucial factor in an SD model.

3.1. Case Study and Problem Specifications. We conducted a case study on a private and not-for-profit metropolitan hospital in the North of Taiwan. Although this hospital is not a medical center, it is the largest institution with the most beds in the area it is located. The hospital adopts a performance-oriented administrative culture. Various performance management systems have been implemented in this hospital for various posts. Consequently, this hospital is far ahead of the other hospitals in the area in terms of financial performance in the health insurance application. However, the hospital is located in a remote area; thus, it encounters difficulties in retaining personnel. Problems related to HR outflow exist for all the positions in the hospital. Although each hospital employee signs a contract when accepting their job offer, they usually choose to leave the job when their contract expires. Despite improvements in the transportation infrastructure that have removed geographic constraints between the rural area in which the hospital is located and the nearby urban area, employees who remain with the hospital after their contract period are usually locals. In addition, the hospital is a leader in the medical industry; thus, to compete with hospitals in urban areas in attracting talents, it has raised the salaries of its employees, which has led to an increase in its operational costs. Consequently, the net profits of the hospital have decreased within a very short period. However, the hospital still has been unable to solve the problem of HR outflow. Other hospitals that are located in the same area and are of the same level as the case hospital are mostly public institutions or associated with a religious organization, and they have a strong focus on services (e.g., integrated delivery system). Employees of these other hospitals work there because they share altruistic ideals. Thus, these other hospitals need not increase the salaries of their personnel to remain competitive with hospitals located in urban areas. The national health insurance (NHI) system in Taiwan has been adjusted to the managed care-related related policy in recent 20 years. Hospital income cannot be increased without limitation; however, HR costs are still very high.

In this study, the SD modeling process involved five essential steps [53]: (a) problem articulation (boundary selection), (b) dynamic hypothesis formulation, (c) simulation model creation, (d) testing, and (e) policy design and evaluation. The information used for the study was collected from archival records and interviews with health-care personnel, namely, five health-care managers and two directors in the nursing department of the case hospital. Moreover, five strategies were considered as alternative talent retention policies for assessment.

3.2. Causal Loop Diagram. The first step in developing an SD model is to define a causal loop diagram. Causal loop diagrams are useful for identifying the feedback loops involved in a process and for representing the feedback structures of systems. The developed SD model captures the relationships between all nodes, including patients, medical staff, hospitals, hospitals' HR policies, and total hospital expenses, to define the system boundary. In this study, the causal loop diagram for the problem of talent retention policy selection focused on nursing staff duty, work pressure, and external information disclosure criteria.

The causal loop diagram described the five aforementioned talent retention policies with respect to the three criteria (nursing staff duty, work pressure, and information disclosure criteria). The developed SD model assumes that when the hospital recruits a new staff member, the workload in the team increases, resulting in this or other staff members resigning and leaving the system; this represents a sequential

process, with the new staff member being the first point (stock) of the sequence. The rate of stock inflow was determined. This study set the system boundary on the basis of relevant internal factors, including the number of patients and care staff members; external factors, including population served by the hospital, number of other competing hospitals, insurance budget, and hospital administration policy. The problem specification and the causality of the overall system are described as follows.

Most medical utilization models consider the mutations of epidemic diseases and special health check-ups, but these items are not covered by insurance systems. Medical staff members experience work-related stress when their hospitals adopt new technologies that are challenging to operate. Thus, medical institutions must frequently retrain their staff members to improve their health-care provision skills, which can reduce their work-related stress. Moreover, new employees increase the burden and stress exerted on existing employees because they are unfamiliar with their job processes; the high proportion of new employees (with 30% of front-line staffs of the hospitals being new employees) can be attributed to the high turnover rate of health-care personnel. Because of the burden and stress induced by new employees, staff morale is relatively low, and the health-care quality decreases; this leads to the resignation of existing personnel, thereby creating a vicious circle. A decrease in health-care personnel leads to a decrease in health-care supply and an increase in the duty ratio (health-care demand/health-care supply), which imposes extra stress on front-line personnel and leads to a further increase in their turnover rate. The expansion limits set for hospitals by the health-care insurance organization when preparing its budget are based on regional totals and the calculation method of point values. Hospitals frequently adopt incentive measures to increase their market share, and these measures often increase their medical resource consumption and expenses. Moreover, the performance system causes increases in the number of patients and the personnel workload. Medical institutions require a prolonged time to recruit new employees in response to increases in workload. Failure to promptly recruit new employees to take up the excess workload leads to stress among existing personnel, and this ultimately increases their turnover rates; the high turnover increases the institution's expenses. In summary, factors such as population aging lead to an increase in the demand for medical services; nevertheless, the supply of such services is limited and is governed by the total number of people paying insurance premiums. Thus, relevant authorities have developed many policies to restrain the growth of hospitals' expenses and reduce their expenses if possible. However, as displayed in the causal loop diagram, a reduction in hospital expenses might increase health-care risks. When insufficient health-care personnel is available to meet patients' health-care demands, some balancing policies must be adopted to ensure that patients' rights and treatments are not negatively affected.

To estimate the performance of different strategies for workload reduction, factors related to revenue sustainability, such as the total insurance budget/cost associated with a particular nursing staff member, should be considered first. The performance of different strategies was estimated by experts by using fuzzy questionnaires and real data collected from the case hospital. The following five HR balancing strategies were evaluated in this study [6].

3.2.1. Policy 1: Changing the Care Model. Policy 1 involves replacing the original primary nursing model with a functional care model. However, although the adoption of the functional care model can temporarily increase efficiency, the root problem of insufficient medical supply still exists. Thus, personnel are still under stress and may eventually resign. Policy 1 might help balance the causal loop over the short term; however, it can also lead to a long-term time delay effect. Therefore, policy 1 has unsuitable effects over the long term.

3.2.2. Policy 2: Increasing Working Hours. Policy 2 involves making health-care personnel work overtime. The number of patients and the workload of long-term care institutions are somewhat fixed, and employees often work overtime in these institutions. However, health care is a persistent job. The tasks involved in each shift must be handed over to the person working the next shift, and many varied tasks are conducted in hospitals. In the short term, increasing working hours can increase the medical supply; however, over the long term, overtime might lead to high turnover rates and thus a time delay effect.

3.2.3. Policy 3: Increasing the Number of Patients Attended to by Each Health-Care Professional. Patients must be cared for even when the supply of manpower is insufficient. Patients cannot be discharged because of an inadequate workforce. Therefore, the number of patients assigned to a single health-care personnel is increased to meet care needs. When the number of health-care employees in a hospital decreases, the duty rates for the remaining employees are increased because the number of patients is unchanged. According to the literature, when the number of patients to be cared for by each health-care professional decreases by 1, the corresponding health-care quality, staff morale, and job satisfaction decrease [12, 14], which results in personnel resignations over the long term. Thus, policy 3 results in a time delay effect over the long term.

3.2.4. Policy 4: Increasing the Number of Care Staff Members. Although an upper bound exists for the number of patients assigned to each health-care personnel (i.e., a maximum of 2.5 beds assigned to a single health-care personnel), hospitals usually consider their entire staff, including those in administrative units, outpatient units, and other special units, when evaluating the number of patients to assign to their personnel. Thus, the actual number of available personnel is different from the reported number of personnel. Consequently, when the number of care personnel decreases, the workload of the remaining personnel increases considerably. Because the HR departments of hospitals consider only the total number of personnel as a basis for recruitment, the

number of care personnel hired might be insufficient, which negatively affects the provided health care and increases the workload of health-care personnel. This also engenders a time delay effect.

3.2.5. Policy 5: Reducing the Number of Available Beds. Under policy 5, the numbers of beds and health-care personnel are regulated according to the ratio of beds to total personnel. When the number of personnel is insufficient, some wards can be closed or the bed availability can be reduced. The removal of beds results in a decrease in long-term medical demand. However, this measure does not help reduce the immediate medical demand or the number of person-days of hospitalization. Consequently, the workload might increase in the short term. The main problem associated with removing beds is that hospitals cannot turn away patients, and reducing the number of beds would result in complaints about long waiting periods for hospitalization. Thus, the stress on employees might increase over the long term, which is a time delay effect.

On the basis of the preceding analysis, this study established a diagram to demonstrate the possible reason for the high turnover of health-care staff (especially nursing staff) and to outline policies that can be adopted by health-care institutions to address the problems caused by the high turnover, as presented in Figure 1. The next section presents the stock and flow diagrams as well as their equations, all of which were used to analyze the relationship of the policies and the stabilize of the HR structure in the case hospital. We attempted to identify a suitable strategy for handling care staff turnover.

3.3. Stock Flow Diagram. The proposed SD model combines HR demand and HR supply to evaluate HR policies. On the basis of the aforementioned causal loop diagram, we created a stock-flow diagram. Subsequently, the dynamic equations for each element in the stock flow diagram were added to the developed SD model.

The ultimate purpose of an SD model is to help a manager simulate the influences of management decisions on system growth and stability in a management system. The manager can then develop measures or policies to improve system performance. Therefore, only the manager can decide whether the SD model helps improve the actual management performance.

In the proposed methodology, the system to be governed (modeled using SD stock and flow language) is separated from the human decision-making system (policies). A stock node (symbolized by a rectangle) represents a point where content can accumulate and deplete. A flow node (symbolized by a valve) is a rate of change in a stock node, and it represents an activity, which fills in or drains the stock node. An auxiliary node or a constant node can store an equation or a constant. Finally, the connectors, represented by simple arrows, are the information links representing the cause and effects within the model structure; the double arrows represent physical flows. Double lines across the arrows indicate time-delayed information.

We created our stock flow diagram for the health-care subsector by using the number of care staff members and the workload information. Providing health care to patients is a skill. The experience of health-care personnel influences the quality of the care they provide and their risks of making errors during care provision. According to the HRP model proposed by Nkomo, experience differences between personnel should be considered to reduce errors in evaluations. Therefore, the present study referenced the skill acquisition model developed by Hubert and Dreyfus to design an HR classification structure [54]. However, interviews with administrators from the case hospital indicated that the developed five-level structure was excessively complex. Therefore, we categorized nurses in the second year of their job as advanced nurses, those in the third year of their job as senior nurses, and those holding a position in the administration or nursing department as experts. The stock-flow diagram for HRs in nursing is displayed in Figure 2.

The trial period for each employee in the case hospital is 3 months, and those who pass their trial period become official employees. New recruits are junior nurses, who become senior nurses after gaining 1 year of experience. Every year, the case hospital conducts a survey of employees' willingness to remain at the hospital. Those who wish to resign can be classified into the category of "expected to leave office." According to the manager of the case hospital, personnel changes after the expiry of contracts are usually related to transfers from a first-line post to another post or resignations due to family needs. As revealed by annual survey data, the turnover rates of the case hospital are 65% and 30% for advanced and senior personnel, respectively.

3.3.1. Calculating Number of Care Staff Members and Workload. To determine HR supply and HR demand, the current number of staff members must be calculated. Moreover, weights should be assigned for various factors for calculation. According to the interviews conducted in this study, the case hospital adjusts its number of job vacancies on the basis of its current HRs. The main recruitment stages are usually implemented after two national exams every year; however, the number of people the hospital has recruited and the number of people it reports to have recruited are two factors that should be considered for calculation. In addition to the workload demand, the number of people of different seniority levels leaving the hospital should be considered during the calculation of the quantity of resources required to fill these vacancies. The time required to find suitable candidates usually differs for different seniority levels. An interviewee stated that finding suitable candidates immediately was impossible. Therefore, in the case of departures, the additional workload was spread among the remaining personnel. Consequently, the weight pertaining to workload could be set to 1.

Because the seniority of first-line health-care personnel might influence their service quality, a suitable weight should be adopted for personnel at each seniority level. Therefore, we set a weight of 1 for senior staff and experts. The interviews also indicated that during the trial period,

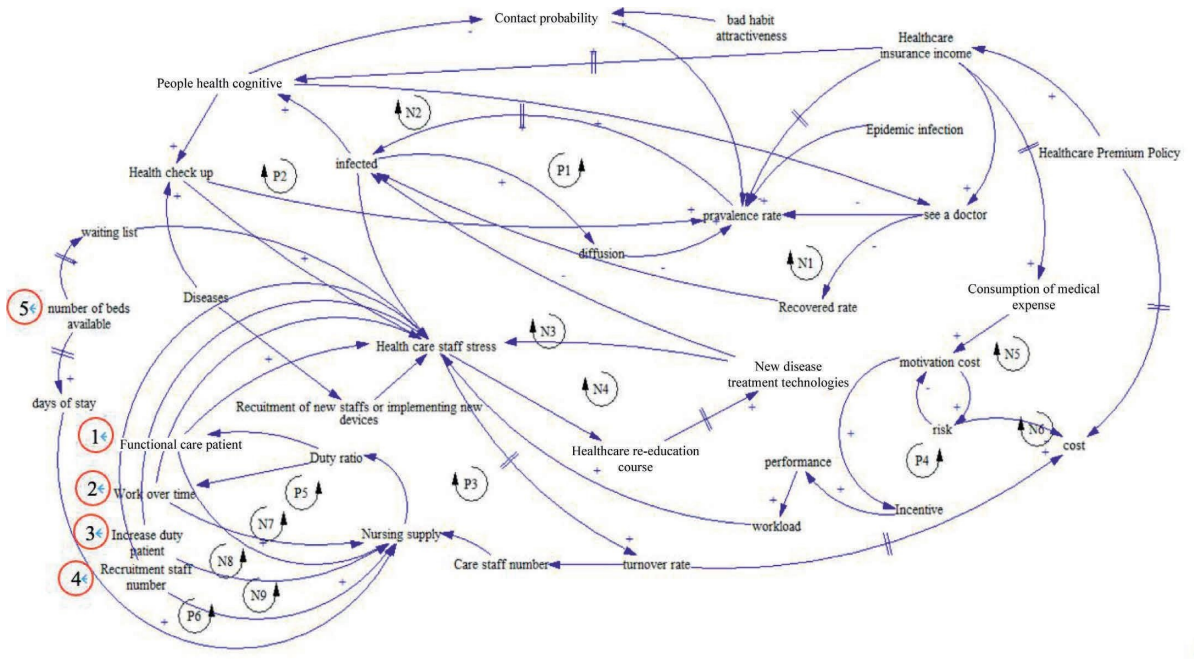


Figure 1: Causal feedback loop diagram of the react policies about loss of care staffs. Note. P1~P5: positive loop; N1–N9: negative loop; circle1~circle5: strategy.

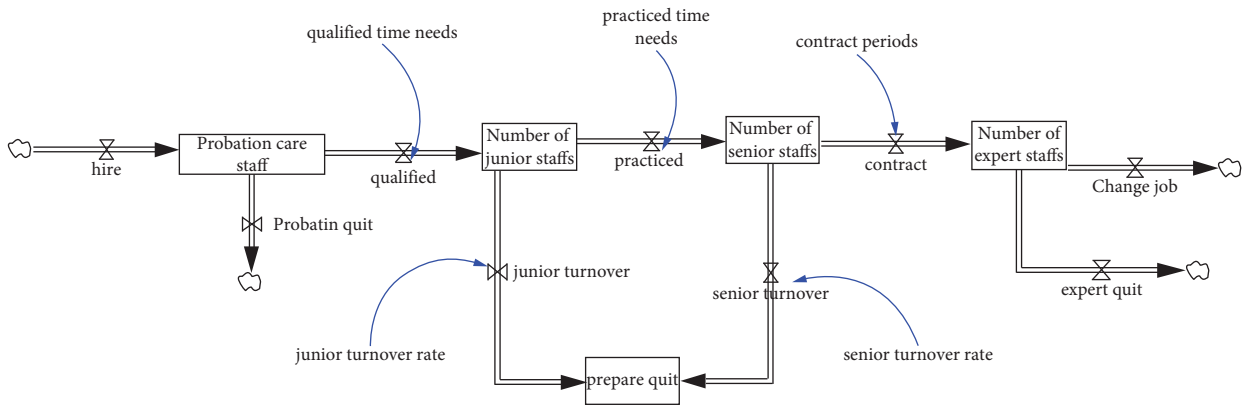


Figure 2: The stock-flow diagram for human resources in nursing.

personnel usually spend their time learning. Moreover, they provide care to patients with simple conditions. After the trial period, personnel are fully integrated into the health-care provision system. Senior staff must spend time teaching new staff in the trial period. Accordingly, we set a weight of −1 for senior staff providing assistance during the trial period. Moreover, directors cannot engage in health care when they are interviewing junior staff. Therefore, we set a weight of 0.8 for directors. Finally, we set a weight of 0.9 for registered nurses who plan to leave the job after the expiration of their contracts. Because the weight of the resignation possibility includes negative numbers, we adopted a setting to avoid negative values.

A higher person-days of hospitalization might lead to higher workloads for health-care staff. However, for calculation, an assumption of a three-shift system (day, evening, and overnight shifts) would be more convenient than a 24 hours way; thus, person-days were converted into hours. The maximum number of patients assigned to each care personnel in the three-shift system was set to 8, 10, and 12 for the day, evening, and overnight shifts, respectively. Moreover, the time required for patient care varies with disease severity. Hence, we also considered disease severity in terms of the Carlson comorbidity index and set the risk ratio to 1.2 [55]. We also set the daily nursing hours to 8 h per day. Because working overtime is highly common in hospitals, a corresponding parameter was added to the system dynamic model.

Health-care workloads might increase because of changes in health-care tasks, and increased workload is a

major reason for HR outflow. This time delay factor should be considered before applying any response measures. Thus, the nursing workload ratio was set to 1. A nursing workload ratio higher than 1 indicates excessive workload. The time required to recruit new staff to fill vacancies was set to 6 months. Because the provided health care affects people's health considerably, medical service quality must be considered when assigning jobs. In addition, upper bounds should be set for the workload. Accordingly, the upper bound was set to 12, 15, and 18 patients for the day, evening, and overnight shifts, respectively. The literature reveals that a positive relationship exists between work pressure and HR outflow. This study thus developed an HR outflow function (Figure 3) and a stock-flow diagram (Figure 4) for the analysis of staff workload.

In general, if a manager wishes to make a suitable policy decision, they must consider strategic approaches toward multiple policy objectives. Moreover, they should avoid linear thinking to avoid making wrong decisions. However, determining the best multiobjective policy through humans' limited thinking loops is difficult. Therefore, the MOP model can be used to design different objective functions under a set of constraints for decision-making in systems involving two or more goals.

As mentioned, every system has unique objectives and requirements pertaining to workforce planning. To achieve all objectives and find the best solution, all the adopted systems must be balanced in every loop. Patients generally wish to receive the best (most expensive) service; however, the insurance expense system of a hospital must minimize the hospital's claimed expenses. Therefore, a hospital management system is typically designed to increase the hospital's revenue by maximizing its income and minimizing its costs (the labor cost accounts for the highest percentage of all costs in a hospital). Employees prefer their labor payment to be increased to the highest amount possible. If the labor payment is insufficient, employees' willingness to offer high-quality services might be influenced, which can considerably influence patients' perceptions of the hospital. Thus, optimizing the aforementioned multiobjective parameters through SD modeling can facilitate a more comprehensive policy consideration process. We conducted simulations for the multiobjective optimization model developed in this study by adding insurance and workforce constraints to the model. We also considered the effects of turnover on the balance of possible institutional changes or structural workforce changes. These effects were determined by decision nodes, which were represented by the growth rate of each industry, in the simulations. Some of the constraints used in the baseline scenario are described as follows.

(1) Population Subsystem. A strong relationship exists between medical demands and population. We defined the demand subsector and population subsector. Population data obtained from the Department of Statistics, Ministry of the Interior, Taiwan, revealed that by the end of 2013, a total of 458,456 people resided in the county in

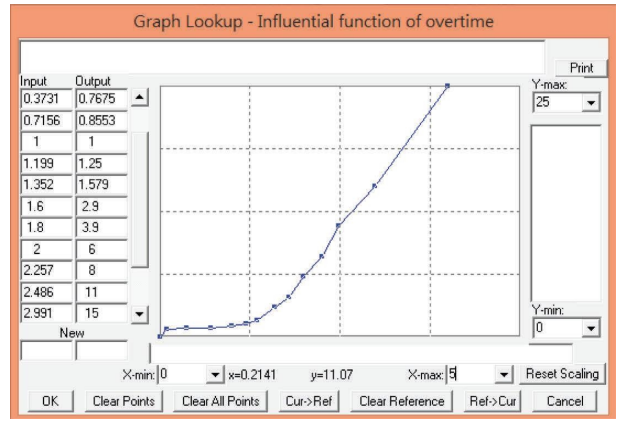


Figure 3: The function of human resource outflow due to increase in workload.

which the case hospital is located. Among them, 88,112 people (20% of the county population) were aged older than 65 years. Overall, elderly people constitute 11.53% of the population of Taiwan in 2013. Therefore, the aforementioned county has a high population of elderly people. Table 1 presents the population trend over the past 10 years in this county. Population aging engenders increased medical demand and increased workload for health-care staff.

(2) Revenue Subsystem. The Taiwan NHI system is different from those of other countries with a family doctor system. In Taiwan, no appointment system exists for outpatient services; however, moral constraints prevent doctors from declining any patient. Most Taiwanese hospitals have a culture of performance incentives based on physician fees. Therefore, the numbers of patients in Taiwanese hospitals are very high, which leads to high workloads and work pressure for health-care staff. Accordingly, we collected data on the monthly numbers of patients in the case hospital to define our simulation parameters. With regard to medical revenue, the data collected by this study from the case hospital included medical income from inpatient and outpatient services covered by the NHI program, medical income from medical services paid by patients, and nonmedical income. The main revenue source of the case hospital was the income obtained from services covered by the NHI program; information regarding all other sources of revenue was confidential. Therefore, our simulations were related to the services covered by the NHI in the case hospital. Partial revenue data for the case hospital are presented in Table 2.

(3) Care Provider Subsystem. The purpose of this study was to analyze the influences of the HRP policies of a medical institution on the outflow of its health-care staff. Thus, the HR supply data used in this study comprised the number of nursing personnel in the case hospital and the number of personnel resigning from the case hospital. We collected monthly data regarding the experience levels of the nursing staff (N–N4) and the number of personnel of each

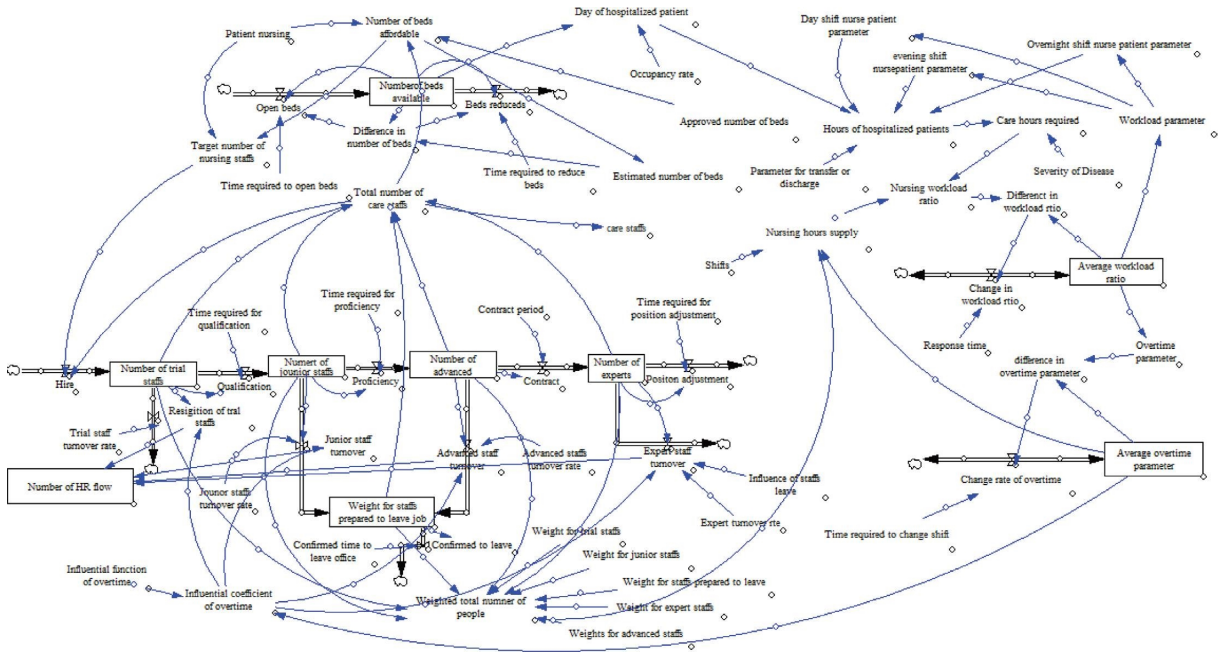


Figure 4: The stock-flow diagram of the microcare staff's workload model.

Table 1: Population of the case administrative district.

Year	Birth		Death		Social increase rate		Population	Elderly rate
2000	6388	0.014	3162	0.007	−2624	−0.006	465186	0.145
2001	5487	0.012	3221	0.007	−3969	−0.009	465799	0.148
2002	5092	0.011	2935	0.006	−4319	−0.009	464107	0.152
2003	4701	0.010	3305	0.007	−2876	−0.006	463284	0.155
2004	4428	0.010	3441	0.007	−3852	−0.008	462286	0.158
2005	4098	0.009	3161	0.007	−3412	−0.007	461586	0.159
2006	3952	0.009	3221	0.007	−3025	−0.007	460426	0.16
2007	3796	0.008	2935	0.006	−1715	−0.004	460398	0.162
2008	3683	0.008	3305	0.007	−3501	−0.008	460902	0.164
2009	3683	0.008	3441	0.007	−2412	−0.005	461625	0.168
2010	3637	0.008	3544	0.008	−2266	−0.005	460486	0.173
2011	3448	0.008	3490	0.008	−1577	−0.003	459061	0.18
2012	3930	0.009	3476	0.008	−1895	−0.004	458595	0.186
2013	3531	0.008	3495	0.008	−519	−0.001	458456	0.192

Source: budget, accounting and statistics department of local county; P.S.: birth rate: means the total number of births per 1000 of population each year; young age means population by age of 0–14; working age means population by age of 15–64; elderly means population by age of 65 and over; social increase rate means the immigrant rate subtracting the emigrant rate; total population = (young age + working age + elderly) ∗ (1 + social increase rate).

level who resigned from the hospital (including those who resigned and those who could not perform first-line tasks because of different reasons) over a 5-year period. In general, the peak period for nursing staff outflow in the case hospital was determined to be from July to September and from February to April. Table 3 presents a summary of the number of nursing staff members who resigned from the case hospital over the 5-year period, which increased exponentially.

3.4. Fuzzy MOP Model. An MOP model is typically used to maximize or minimize different objective functions under a set of constraints and is essential for decision-making in systems involving two or more goals. In a system for selecting a suitable strategy for health-care talent retention, each subsystem has unique goals, and each item must be optimized to achieve a suitable trade-off among the subsystems. In general, in such a system, the health-care institution's revenue and insurance revenue must be maximized, whereas the HR cost must be minimized. To achieve these goals, system parameters that strongly influence the system output must be identified. Accordingly, we developed an MOP model and applied it to a dynamic system to optimize its parameters. The notations used in this section are listed in Table 4.

Table 2: Medical visits and points of case hospital for revenue subsector.

Year/month		Hospital outpatient		Hospital inpatient	
		Visits	Medical points	Visits	Medical points
101	1	56722	164,447,252	4,436	165,710,659
101	2	44812	108,866,638	3,882	118,505,577
101	3	47992	128,145,124	4,271	143,449,344
101	4	55516	161,820,331	4,313	173,307,089
101	5	54391	159,923,394	4,363	163,446,223
101	6	46246	125,912,544	4,174	148,963,647
101	7	58280	173,265,634	4,458	173,488,691
101	8	53153	152,568,364	4,372	155,011,434
101	9	46291	128,705,569	4,117	142,019,141
101	10	58660	177,711,212	4,304	166,300,818
101	11	51426	144,346,940	4,192	157,847,699
101	12	54385	159,835,276	4,313	170,175,597
102	1	52916	147,071,264	4,995	209,700,289
102	2	53858	135,073,360	4,276	156,556,231
102	3	55692	152,597,484	4,762	204,766,443
102	4	61262	169,234,139	4,792	206,670,431
102	5	62737	166,681,236	5,238	213,385,929
102	6	57035	147,975,918	4,301	165,745,638
102	7	63862	173,854,512	4,640	176,675,969
102	8	59569	153,677,310	4,261	162,280,972
102	9	60040	157,683,212	4,201	162,262,866

Table 3: A comparison of each year of nursing staff outflow.

Year	Number of staffs	Average staff turnover rate
2009	324	0.027
2010	325	0.045
2011	323	0.047
2012	322	0.05
2013	325	0.052
2014	314	0.081

Table 4: List of parameters.

Notation	Explanation
f_1	Objective related to a health-care institution's revenue
f_3	Objective related to human resource planning
G_2	Patient care revenues
$E1$	Revenue of hospital
$E3$	Revenue of district hospital
x_1	Growth rate of hospital
x_3	Growth rate of hospital
x_5	Proportion of investment in the remaining staff
C_1	Global budget for western medicine consumption
C_3	Global budget for dental medicine consumption
C_5	Reduction in hospital expenses after the implementation of the global budget
s_i	Consumption coefficient of fee i to human resource loss
α_1	Restrictive impact factor coefficient for the global budget system
f_2	Objective related to the insurance system
G_1	Patient care revenue under health insurance
G_3	Nonpatient revenue
$E2$	Revenue of a metropolitan hospital
$E4$	Revenue of a clinic
x_2	Growth rate of a metropolitan hospital
x_4	Administration fee in the global budget
C_2	Global budget of western primary medicine consumption
C_4	Global budget of Chinese medicine consumption
a_i	The exhausted of care staff equivalent transformation coefficient of the insurance of fee i
e_{ij}	Consumption coefficient of fee i per unit of output for the jth hospital level
α_2	Implement impact factor on coefficient of human loss policy

3.4.1. Objective Function

(1) Objective Related to a Health-Care Institution's Revenue. Health-care institutions are nonprofit organizations; however, they must generate sufficient revenue to operate sustainably. Some health-care institutions have also started commercializing their services. They must earn revenue that at least equals the operating cost and the capitalized cost related to the new investments. In addition, under the influences of an unlimited insurance policy and private patient fees, achieving maximum revenue has become a major objective of most health-care institutions. This objective can be considered the first objective in this study and can be expressed as follows:

$$\begin{aligned} f_1 &= E_1 + E_2 + E_3 + E_4, \\ E_1 &= G_1(1+x_1) + G_2(1+x_1) + G_3(1+x_1), \\ E_2 &= G_1(1+x_2) + G_2(1+x_2) + G_3(1+x_2), \\ E_3 &= G_1(1+x_3) + G_2(1+x_3) + G_3(1+x_3), \\ E_4 &= G_1(1+x_4) + G_2(1+x_4) + G_3(1+x_4). \end{aligned} \tag{1}$$

(2) Objective Related to the Insurance System. Unlimited growth in the global insurance budget is impossible, and this budget influences the economy and civil atmosphere of a nation. Rapid aging in a society engenders increased health-care demand, and a gap between health-care demand and supply is a developmental bottleneck for health-care institutions. Therefore, reducing insurance fees is a major objective of the National Health Insurance Administration (NHIA) of Taiwan, which can be considered the second objective in this study. This objective can be expressed as follows:

$$\begin{aligned} f_2 &= C_1 + C_2 + C_3 + C_4 - C_5, \\ C_1 &= G_1(1+x_1)a_{11} + G_2(1+x_2)a_{12} + G_3(1+x_3)a_{13}, \\ C_2 &= G_1(1+x_1)a_{21} + G_2(1+x_2)a_{22} + G_3(1+x_3)a_{23}, \\ C_3 &= G_1(1+x_1)a_{31} + G_2(1+x_2)a_{32} + G_3(1+x_3)a_{33}, \\ C_4 &= G_1(1+x_1)a_{41} + G_2(1+x_2)a_{42} + G_3(1+x_3)a_{43}, \\ C_5 &= x_4\alpha_1, \end{aligned} \tag{2}$$

where C_1, C_2, C_3, and C_4 represent the global budgets of western medicine consumption, western primary medicine consumption, Chinese medicine consumption, and dental medicine consumption, respectively, at the three hospital levels (medical centers, metropolitan hospitals, and district hospital). Moreover, C_5 denotes the insurance amount saved through any budget control policy.

(3) Objective Related to the HR Cost. The expansion of a hospital is limited by the total population in its region and the global budget of insurance fee point calculation systems, which can result in a crowding-out effect. Thus, hospital expenses should be decreased to raise hospital profits; the HR cost accounts for the highest proportion of the total expenses of a health-care institution. However, according to the causal loop diagram, expense reductions might increase the risks associated with patient care. Therefore, the HR cost must be minimized such that patient safety is not compromised. This can be considered the third objective in this study and can be expressed as follows:

$$f_3 = s_1c_1 + s_2c_2 + s_3c_3 - x_5\alpha_2. \tag{3}$$

The final item in equation (3) is the NHI claim reduction resulting from the health-care institution's internal adjustment policy.

3.4.2. Constraints. Decision-makers select their preferences for the relationships between factors or variables in an HRP system, and these preferences are then considered as constraints for MOP.

(1) Constraint Related to the Growth of a Health-Care Institution. Health-care demands and expenses increase with the growth of the elderly population in an area. However, increases in the use of insurance resources and institution costs must be limited while meeting the health-care demand. Therefore, an upper bound for the growth rate of a health-care institution is a constraint in an HRP system. Moreover, the growth rate of NHI fees must be higher than the development rate of health-care institutions in the relevant country. Accordingly, the constraint for the growth of a health-care institution can be expressed as follows:

$$f_1 = \frac{(E_1 + E_2 + E_3)}{\mathrm{E}} \le d_1, \tag{4}$$

where d_1 represents the minimum growth rate of insurance fees.

(2) Investment Constraints. Most health-care institutions in Taiwan are nonprofit organizations covered by insurance, which indicates that they have limited income. They cannot invest in other businesses and must use their revenue for their own needs. In general, the manager of a health-care institution must strive to achieve a cost allocation of 33% for each of the following aspects of their institution: HRs, medicines, and utilities. Therefore, the proportion of investment into HRP is limited. Let d_2 and d_3 represent the upper limits of the global budget and HRP budget, respectively; hence, the following equations can be derived:

$$\begin{aligned} x_4 &\le d_2, \\ x_5 &\le d_3. \end{aligned} \tag{5}$$

(3) Insurance Source Constraints. Insurance consumption mainly varies with the age of, health-care literacy of, and medical advice obtained by an individual. Moreover, the density of hospitals in an area and the proportion of the population seeking medical advice in this area govern the consumption of insurance resources, especially after the implementation of the capitation and fee-for-service reimbursement system. To control the growth rate of insurance consumption, the NHIA can set fixed thresholds or

constraints to regulate the frequency at which people can seek medical advice and the frequency of hospitalization (e.g., global budget or diagnosis-related group system). Accordingly, the following proportional constraints can generally be applied:

$$\begin{aligned} b_{11} &\le \frac{a_1c_1}{a_1c_1 + a_2c_2 + a_3c_3 + a_4c_4} \le b_{12}, \\ b_{21} &\le \frac{a_2c_2}{a_1c_1 + a_2c_2 + a_3c_3 + a_4c_4} \le b_{22}, \\ b_{31} &\le \frac{a_3c_3}{a_1c_1 + a_2c_2 + a_3c_3 + a_4c_4} \le b_{32}, \\ b_{41} &\le \frac{a_4c_4}{a_1c_1 + a_2c_2 + a_3c_3 + a_4c_4} \le b_{42}. \end{aligned} \tag{6}$$

Moreover, the insurance budget for the current year should be less than that for the previous year. Hence, the following inequality should be satisfied:

$$\frac{f_2}{f_1} \le \mathrm{IF}(1 - d_4), \tag{7}$$

where IF denotes the insurance fee for the previous year and d_4 is the average rate of increase in this fee.

(4) Constraint Related to the HR Cost. The HR cost generally accounts for the highest proportion of the total cost incurred by a health-care institution. Therefore, from an administrator's viewpoint, an increase in HR costs should be minimized. The constraint for the increase in HR cost is similar to that for the increase in insurance resource consumption, and the following inequality should be satisfied:

$$\frac{f_3}{f_1} \le \mathrm{HC}(1 - d_5), \tag{8}$$

where HC denotes the HR cost in the previous year and d_5 denotes the average rate of increase in this cost.

On the basis of the aforementioned analysis, the established multiobjective model can be expressed as follows:

$$\begin{aligned} & f_1 = (E_1 + E_2 + E_3 + E_4), \\ & f_2 = s_1c_1 + s_2c_2 + s_3c_3 + s_4c_4 - c_5, \\ & f_3 = s_1c_1 + s_2c_2 + s_3c_3 - Gx_5\alpha_2, \\ & f_1 = \frac{(E_1 + E_2 + E_3)}{E} \le d_1, \\ & x_4 \le d_2, \\ & x_5 \le d_3, \\ s.t & \begin{cases} b_{11} \le \dfrac{a_1c_1}{a_1c_1 + a_2c_2 + a_3c_3 + a_4c_4} \le b_{12}, \\ b_{21} \le \dfrac{a_2c_2}{a_1c_1 + a_2c_2 + a_3c_3 + a_4c_4} \le b_{22}, \dfrac{f_2}{f_1} \le IF(1 - d_4), \\ b_{31} \le \dfrac{a_3c_3}{a_1c_1 + a_2c_2 + a_3c_3 + a_4c_4} \le b_{32}, \dfrac{f_3}{f_1} \le HC(1 - d_5). \\ b_{41} \le \dfrac{a_4c_4}{ac_1 + a_2c_2 + a_3c_3 + a_4c_4} \le b_{42}, \end{cases} \end{aligned} \tag{9}$$

3.4.3. Fuzzy Extensions. In a multiobjective model, many coefficients and parameters must be determined. Some of these parameters can be obtained from historical records such as local statistical records. For parameters with inadequate, incomplete, imprecise, or inconsistent data, domain experts can subjectively determine the values for these parameters. Because HRP involves uncertainties and great influence, each considered HR policy in this study could not be implemented realistically in the case institution. As mentioned, fuzzy programming is useful for solving programming problems involving uncertainties. Accordingly, this study incorporated fuzzy programming

into the aforementioned multiobjective model to handle potential uncertainties in the selection of health-care-related HRP.

In general, the triangular fuzzy parameters, which can be expressed in terms of a triplet of crisp numbers [i.e., $(r_1, r_2, r_3$, where $r_1 < r_2 < r_3]$, are used to describe fuzzy coefficients. The membership function of a fuzzy parameter can be expressed as follows:

$$\mu(x) = \frac{x - r_1}{r_2 - r_1}, \text{if } r_1 \le x \le r_2;\ \frac{x - r_3}{r_2 - r_3}, if\ r_2 \le x \le r_3; 0, \text{otherwise.} \tag{10}$$

Fuzzy parameters are assumed to be triangular; therefore, the fuzzy equivalent of a can be denoted as $\tilde{a}$. Let be x a decision vector; ξ be a fuzzy vector; $f_i(x, \xi)$, $i = 1, 2, \ldots, n$, be an objective function; and $g_i(x, \xi)$ be a constraint function. Therefore, the following equation can be obtained for a fuzzy multiobjective decision model:

$$\text{Max}\{f_1(x, \xi), \ldots, f_i(x, \xi), \ldots, f_n(x, \xi) s.t.\ g_i(x, \xi) \le 0, j = 1, 2 \ldots\ p\}. \tag{11}$$

This model can be used to solve fuzzy programming problems. Because of this model's fuzzy nature, its precision cannot be compared with that of a mathematical model. Domain experts must thus determine the expected values of the model parameters. For the triangular fuzzy parameter $\tilde{a} = (r_1, r_2, r_3)$, the expected value is $\tilde{a} = 1/3 \times (r_1 + r_2 + r_3)$. By calculating the expected values for fuzzy parameters, experts can transform a fuzzy multiobjective model into a classic crisp MOP model, whose solution can be derived using the method described in the following section.

3.4.4. Method for Solving Multiobjective Problems. In general, four types of methods are used to solve multiobjective problems: methods with no articulation of preference information, methods with a priori articulation of preference information, methods with a progressive articulation of preference information, and methods with a posteriori articulation of preference information. Solving a multiobjective problem basically involves transforming it into a single-objective problem. The most common approach for this transformation is the weighted-sum method, which involves a posteriori articulation of preferences. To illustrate the weighted-sum method, the following multiobjective problem can be used as an example:

$$\{f_1(x), \ldots., f_i(x), \ldots, f_n(x)\} \text{s.t. } x \in X. \tag{12}$$

The first step in the weighted-sum method involves solving the following functions: $\max f_i(x)$s.t., $x \in X$ and $\min f_i(x)$s.t., $x \in X$, where $i = 1, 2, \ldots\ldots\ n$. These functions are solved to obtain the optimal value for each single-objective problem.

The second step entails applying the optimal values obtained in the first step to define new functions as follows:

$$F_i(x) = \frac{f_i^{\max} - f_i(x)}{f_i^{\max} - f_i^{\min}}, i = 1, 2, \ldots., n. \tag{13}$$

The third step involves transforming the original problem into a single-objective problem as follows:

$$\min F(x) = w_1 F_1(x) + \ldots + w_i F_i(x) + \ldots. + w_n F_n(x), s.t.\ x \in X, \tag{14}$$

where w_i is the relative weight of the ith objective such that $w_i \ge 0, \sum_{i=1}^{n} w_i = 1$.

The solution to the problem expressed in the preceding equation is a Pareto-optimal solution, which is satisfactory for the original problem. By using the weighted-sum method, a decision-maker can easily adjust the importance of each objective. Consider, for example, the three objectives outlined in Section 3.6.1. The parameters w_1, w_2, w_3, w_4, and w_5 denote the relative weights of the objectives pertaining to health-care institution development, the patient health system, the insurance spending system, the staff assignment system, and care risk management, respectively. If the importance of the health-care institution's development must be emphasized, w_1 can be set to 0.6 and w_2–w_5 can be set to 0.1. However, if all five objectives are considered to be of equal importance, w_1–w_5 are set as 1/5.

3.5. Model Test. According to Coyle [56], the validity of an SD model does not depend on its absolute accuracy; instead, it depends on its suitability for solving the complex and nonlinear problems. Forrester [28] also stated that an SD model developed for a complex high-order nonlinear system with numerous nodes cannot be validated using general statistical methods. These statements are valid because SD models reflect real-world situations. If some nonsignificant

node relationships in an SD model are removed, the purpose of developing the model might not be achieved [49]. Considering these statements, we adopted Sterman's [49] suggestion and performed a model behavior test and model structure tests. The following model structure tests were performed using the Vensim 6.3 DSS simulation software tool: structure verification, dimensional consistency, and extreme condition tests.

In the structure verification test, we performed two interviews: one before the causal loop diagram was created and one after the flow diagram was created. By exploring the case hospital's management problems, we determined the main simulation structure required for conducting policy simulations in this study. After identifying and confirming the relationships, we performed syntax checking, single-equation checking, lookup usage checking, model checking, and multiple-equation checking by using functions of the Vensim software.

In the dimensional consistency test, we mainly evaluated whether clear and reasonable constants and node values could be calculated by unit synonyms and unit-checking tools to avoid inconsistent unit problems. The simulations conducted in this study were based on real data obtained from the case hospital; therefore, month, which is often used as the unit in models developed for selecting general human affairs strategies, was adopted as the basic unit of the developed SD model. The unit of all the nodes was set to month for consistency. Before starting the simulation procedure, we used the Vensim unit checking tool to ensure that all parameter units were consistent. The formulae and parameters are listed in Table 5. In the extreme condition test, we used two conditions to confirm whether the adopted simulation structure was correct. In the first method, the insurance income was set to 0 when the population was 0. In the second method, the care demand was set to 0 when the occupancy rate was 0.

In the model behavior test, different sets of data were used for different purposes. This test was used to determine whether the simulation results matched the real data from the case hospital. We attempted to determine whether the workload ratio and HR outflow would increase with the occupancy rate. All the simulation results were as expected and exhibited a trend of exponential growth.

4. Simulation and Results

Simulations for the designed system were conducted using the developed SD model. The model parameters were determined from hospital data provided by the Ministry of Health and Welfare in Taiwan, open data provided by the Taiwanese government, and information provided by the case hospital. Baseline scenarios were designed on the basis of the real revenue and nurse-patient ratios of the case hospital. We performed microcosm and macrocosm simulations in this study.

4.1. Policy Simulation Results Obtained with a Microcosm Model for the Case Hospital. The microcosm effects of different HR policies on the case hospital were analyzed using five practical conditions. Regarding HR outflow, primary-level health-care personnel often suggest that the standard patient-nurse ratio should be reduced. Specifically, they suggest that the number of patients requiring care should be decreased to reduce their heavy workload, which can mitigate the problem of HR outflow. Accordingly, we first simulated the effects of reducing the aforementioned ratio in the case hospital. We considered the three nursing models that are currently adopted in hospitals: primary nursing, team nursing, and functional nursing. Primary nursing is currently the most commonly used nursing model. In this model, one nurse must take care of several patients. Nurses only make plans for their patients and complete their tasks during their working hours (in this study, we assumed that three shifts exist per day and reduced the patient-nurse ratio for evaluation). Team nursing involves a group of nurses taking care of a group of patients. Different groups of patients are taken care of by nurses with different levels of experience. Job arrangements for nurses in the same group are based on their experience. We considered nurses of different levels, divided them into four groups, and performed a simulation using the same number of patients for the groups. Functional nursing is a task-centered nursing model that involves assigning specific tasks to personnel. Because some studies have indicated that the efficiency of the functional nursing model is high, we selected seven main tasks: material management, medication (injection), nursing planning, temperature, pulse, and respiration (TPR) measurement, making beds, training, and caring, as the basis for our calculations. We performed the simulation by increasing the quantity of HRs by 1.2.

The simulation results obtained for the nursing model and patient-nurse ratio are displayed in Figure 5. The care workload ratio did not decrease considerably when the patient-nurse ratio was reduced or when the functional nursing model was used. However, when the team nursing model was used, the care workload ratio was effectively reduced. These results indicate that reducing the standard patient-nurse ratio did not effectively reduce the care workload ratio. Instead, it led to uncertainty in the number of available beds, which increased administrative costs. When nurses with different experience levels work in groups, the pressure imposed on them decreases, which consequently reduces the personnel turnover rate.

In the second simulation, we increased the level of overtime, and the results are shown in Figure 6. An increase in overtime initially caused a decrease in workload. Nevertheless, the workload subsequently increased to the original level with time. Thus, the aforementioned policy was not helpful in reducing HR outflow.

In the third simulation, we adjusted the levels of experience of the nurses to increase the health-care efficiency and number of patients cared for. We simulated the effects of applying different policies, such as incentives, retention bonuses, and high salaries, for nurses with different levels of experience. We found that when the turnover rate of senior personnel was minimized, the workload did not decrease.

Table 5: The formula using the software 'Vensim' to simulate the talent retention policy model.

Level/rate	Formula or default observations
Income of institution	INTEG (insurance income + noninsurrance income + nonmedical insurance) – cost – global budget cost
Scheduling parameter	0.2
Total number of staffs	Number of trial staffs + number of junior staffs + number of advanced staffs + number of staffs planning to resign + number of expert staffs
Number of trial staffs	INTEG (hire – qualification – number of trial staffs resigned)
Number of junior staffs	INTEG (Qualification – proficiency)
Number of advanced staffs	INTEG (Proficiency – contract – number of junior staffs resigned)
Number of staffs planning to resign	INTEG (number of junior staffs resigned + number of advance staffs resigned + number of senior staffs resigned- number of staffs confirmed to resign)
Number of expert staffs	INTEG (contract- position adjustment – number of expert staffs resigned)
Qualification	(Number of trial staffs – number of trial staffs resigned)
Time required for qualification	3
Proficiency	Number of junior staffs/time required for proficiency
Time required for proficiency	6
Contract	(Number of advance staffs – number of junior staffs resigned)/contract period
Contract period	24 (2 year)
Position adjustment	Number of senior staffs/time required for position adjustment
Time required for position adjustment	24
Number of trial staffs resigned	Number of trial staffs ∗ trial staff turnover rate
Number of junior staffs resigned	Number of junior staffs ∗ junior staff turnover rate
Number of staffs confirmed to resign	Number of staffs planning to resign/time required for staffs confirmed to resign
Number of senior staffs resigned	Number of senior staffs ∗ senior staff turnover rate
Trial staff turnover rate	0.3
Junior staff turnover rate	0.5
Senior staff turnover rate	0.3
Confirmed time to leave office	1
Total number of care staffs	Number of trial staffs + numbert of junior staffs + numbert of senior staffs + number of staffs planning to resign + numbert of expert staffs
Time required for reassignment	1
Time required for hiring	3
Weighted total number of people	MAX (number of trial staffs ∗ weight for trial staffs + number of junior staffs ∗ weight for junior staffs + number of advance staffs ∗ weight for senior staffs + number of prepared to leave job ∗ weight for staffs prepared to leave job + number of expert staffs ∗ weight for advanced and senior staffs, 0)
Weight for trial staffs	−0.1
Weight for staffs prepared to leave job	0.9
Weight for advanced and senior staffs	1.2
Weight for junior staffs	0.8
Influence of staffs leaving office	1⟶4⟶7⟶8 > 9 (trial staffs⟶senior staffs⟶advanced staffs⟶senior staffs⟶expert staffs)
Care hours required	Person-times of hospitalization ∗ seriousness of disease
Person-times of hospitalization	Number of hospitalized patients ∗ 24/3 ∗ (1/Day shift nurse-patient parameter + 1/Evening shift nurse-patient parameter + 1/Overnight shift nurse-patient parameter) ∗ parameter for transfer or discharge
Parameter for transfer or discharge	0.9
Seriousness of disease	1.5
Supplied hours	Weighted total number of people ∗ 8 ∗ shifts ∗ average overtime parameter
Shifts	22
Workload ratio	XIDZ (care hours required, supplied hours, 50)
Average workload ratio	INTEG (different in workload ratio, 1)
Different in workload ratio	Workload ratio ∗ average workload ratio
Change in workload ratio	Different in workload ratio/responding time
Responding time	1
Day shift nurse-patient parameter	8 ∗ workload parameter

Table 5: Continued.

Level/rate	Formula or default observations
Evening shift nurse-patient parameter	10 ∗ workload parameter
Overnight shift nurse-patient parameter	12 ∗ workload parameter
Workload parameter	MIN (average workload ratio, 1.5)
Overtime parameter	MIN (average workload ratio, 2)
Average overtime parameter	INTEG (change rate of overtime, 1)
Difference in overtime parameter	Overtime parameter-average overtime parameter
Change rate of overtime	Difference in overtime parameter/time required to change shift
Time required to change shift	6
Influential coefficient of overtime	Influential function of overtime (ovetime parameter)
Resignation of trial staffs	Number of trial staffs ∗ MIN (trial staff turnover rate ∗ influential function for trial staffs, 1)
Resignation of junior staffs	Number of advanced staffs ∗ MIN (junior staffs turnover rate ∗ influential function for junior staffs, 1)
Resignation of senior staffs	Number of senior staffs ∗ MIN (senior staffs turnover rate ∗ influential function for senior staffs ∗ influential function for senior staffs, 1)
Number of beds available	INTEG (number of open beds − number of beds reduced)
Number of open beds	IF THEN ELSE (difference in number of beds ≥ 0, difference in number of beds/time required to open beds, 0)
Number of beds reduced	IF THEN ELSE (difference in number of beds < 0, MIN (−difference in number of beds/time required to reduce beds, number of beds available, 0)
Time required to open beds	3
Time required to reduce beds	9
Difference in number of beds	Approved number of beds-number of beds available
Estimated number of beds	SMOOTH (number of beds affordable, 3)
Number of beds affordable	MIN (total number of care staffs ∗ patient-nurser standard, approved number of beds)
Approved number of beds	750
Patient-nurser standard	2.5
Target number of nursing staffs	Number of beds affordable/Patient-nurser standard
Number of hospitalized patients	Delay (number of beds available ∗ 30 ∗ occupancy rate, 1)
Occupancy rate	70 (including reduce beds)
Effect of implementation of technology	Time required for proficiency ∗ weight for each staffs
Coefficient for administrative tasks	MIN (average workload ratio, 1.2)

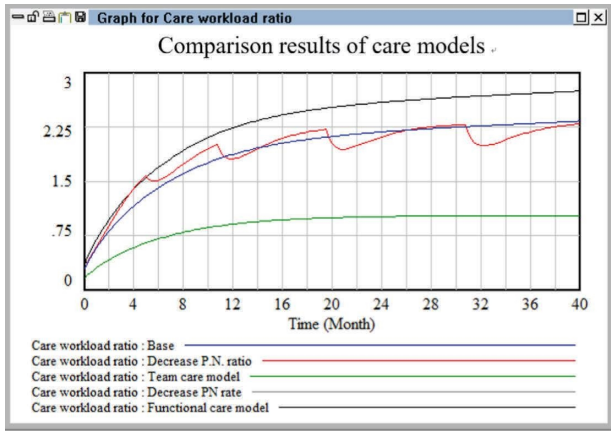


Figure 5: Comparison evaluation results between baseline and different care models.

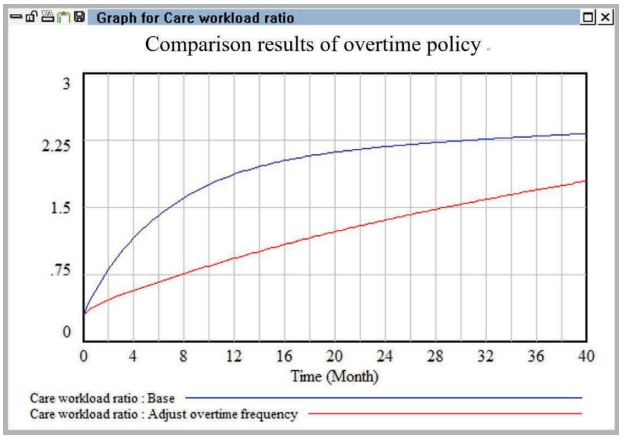


Figure 6: Comparison evaluation results between baseline and overtime policy.

However, when the turnover rate of personnel working in their trial period was reduced to 0, the overall HR turnover and workload decreased. This result can be used as a reference for designing salary policies in the medical field, and it contradicts the line of thinking adopted by current hospital managers. Specifically, hospital managers attempt

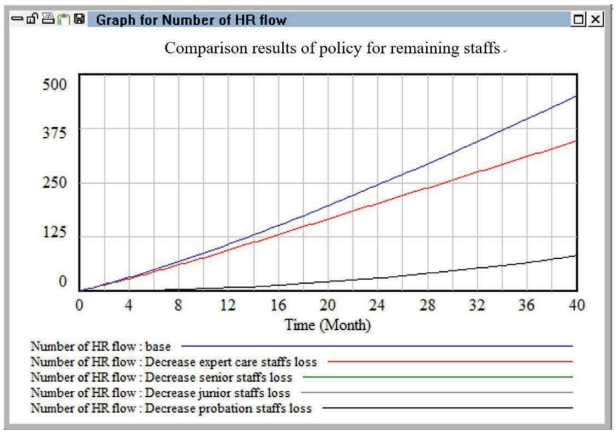


Figure 7: Comparison evaluation results among policy for remaining differential care staffs.

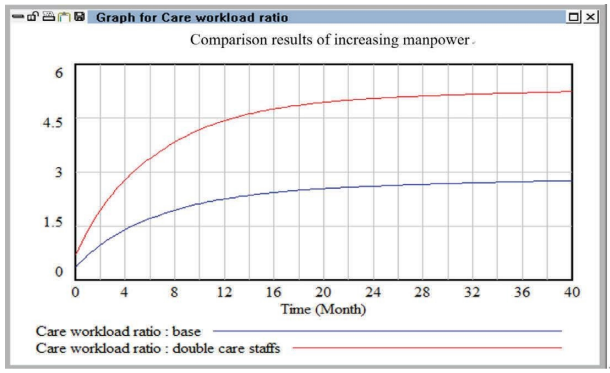


Figure 8: Comparison evaluation results between baseline and increase manpower.

to control HR outflow by promoting senior and advanced personnel. Moreover, they attempt to reduce the contribution of the salaries of primary-level personnel to the overall hospital cost; nevertheless, these personnel constitute a critical group of HRs providing adequate first-line health-care in a hospital. Therefore, managers should consider other policy strategies. Figure 7 indicates that reducing the turnover rate of new employees in their trial period is the most effective policy for HR outflow reduction.

We did not consider management costs in the fourth simulation. In this simulation, we doubled the number of health-care personnel in the case hospital and then examined the corresponding effect on the workload. This examination revealed that the workload increased considerably after a short period. Therefore, the strategy of doubling the number of personnel is unsuitable for reducing HR outflow. Furthermore, as displayed in Figure 8, directly hiring an excessive number of new employees increases communication costs for senior health-care personnel, which in turn increases health-care risks.

The fifth simulation involved reducing the number of beds, and the simulation results are presented in Figure 9. A reduction in the number of beds had no effect on the nursing

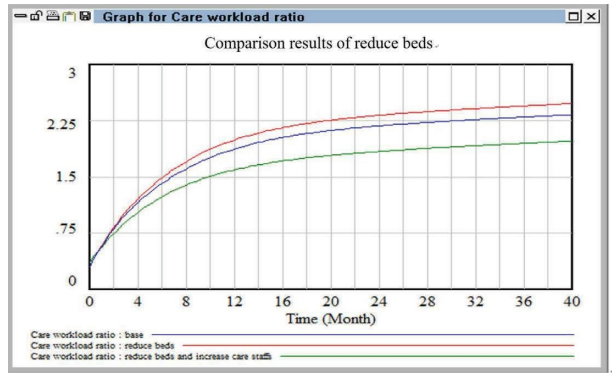


Figure 9: Comparison evaluation results between baseline and reduce beds.

workload under the general health-care model. However, when we combined this policy with the policy of increasing the number of health-care personnel (i.e., when we reduced the number of beds while increasing the number of health-care personnel), the HR outflow decreased. Nevertheless, over the long term, a reduction in the number of beds resulted in increases in the occupancy rate and workload.

Various factors might influence HRP. Traditionally, predictions regarding the total HRs required in an institution are made on the basis of a single aspect. The processes adopted by managers in policy-related decision-making are simple, and their thinking can easily become linear. Examining problems involving systematic thinking is difficult. Moreover, SD modeling is highly suitable for HR policy simulations. However, as indicated in this study, the HR outflow of health-care personnel can be influenced by not only internal factors but also external factors, which include increased medical demand due to population aging in the relevant area, competition between institutions, and NHIA policies (e.g., in Taiwan, nursing fees are part of ward fees and thus receive less attention from managers). Nevertheless, SD modeling can provide some crucial insights. Accordingly, our simulation results are valuable for hospital managers. The following section describes the simulation results that were obtained by using a simple representation of the developed SD model to emphasize the relationships between the nodes of the stock flow diagram and external related systems.

4.2. Policy Simulation Results Obtained with a Macrocosm Model for the Case Hospital. In the area where the case hospital is located, the demand for medical services has been increasing because of the growth of the elderly population. Moreover, the development of the transportation infrastructure in the area has increased the convenience of traveling, meaning that more people can easily travel to the hospital for medical attention; this has increased personnel workload and thus aggravated the problem of insufficient medical HRs. We conducted a macrocosm simulation of Taiwan's medical resource growth by using a simplified representation of the designed SD model (Figure 10), in which we combined three objectives: hospital, insurance, and health-care HR systems to the model.

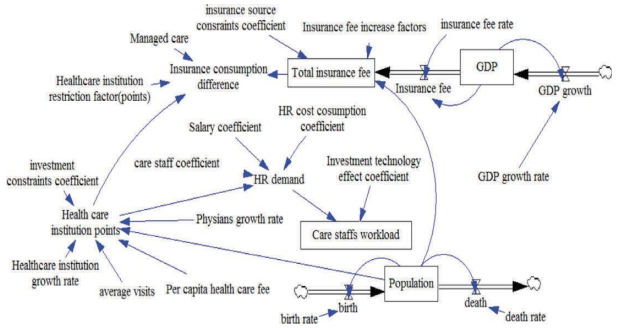


Figure 10: The stock-flow diagram of the macrocare staff workload model.

Table 6: Data of the coefficients.

Notation	Value	Unit
G_1	64028.72	10^5 NT$/month
G_3	6259.06	10^5 NT$/month
a_2	13	%
s_1	42.75	%
s_3	0.05	%
e_{21}	16.46	%
e_{12}	88.47	%
e_{32}	0.42	%
e_{23}	8.4	%
G_2	9087.46	10^5 NT$/month
a_1	76	%
a_3	10	%
s_2	52.25	%
e_{11}	83.14	%
e_{31}	0.4	%
e_{22}	11.11	%
e_{13}	91.6	%
e_{33}	0.1	%

Three baseline scenarios with different rates of population growth and health-care revenues (10%, 20%, and 30% of population increase rate) were designed. The parameters of the simplified SD model were determined using statistical information obtained from Taiwan's Directorate General of Budget, accounting and statistics, and open data obtained from the Taiwan Ministry of Health and Welfare. The parameters of this model are presented in Table 6.

In multiobjective decision-making, the parameters considered (e.g., decisions related to a health-care institution's overall expenses, its insurance costs, and the workload of its care personnel) depend on the decision-maker's preferences. In this study, we assumed that the decision-maker uses fuzzy language to express uncertain quantities, and we transformed these uncertain quantities into fuzzy numbers. For example, the growth rate for each level of a health-care institution or clinic is between approximately 0.06 and 0.15 (i.e., $0.06 < X_i < 0.15$, $i = 1, 2, 3$). Therefore, in this study, all the fuzzy coefficients were assumed to be triangular fuzzy parameters, and a multiobjective model was developed on the basis of the model described in Section 3.4.

Table 7: Trade-offs between the objectives.

Case number	fw_1	fw_2	fw_3
Case 1	0.9	0.05	0.05
Case 2	0.8	0.1	0.1
Case 3	0.5	0.3	0.2
Case 4	0.33	0.33	0.33
Case 5	0.1	0.8	0.1
Case 6	0.1	0.1	0.8

Because of constraints related to time and data collection, the scope of the macrocosm simulation was limited to western medicine; thus, data for Chinese medicine and dental medicine were excluded. Different solutions were obtained for different objective weights (Table 7). Because the first and second objectives were more important than the third objective, two weight selection methods were used. In the first method, the importance of the first and second objectives was emphasized (e.g., cases 1, 2, and 3 adjust the ratio step by step). Moreover, in the second method, the importance of the three objectives was balanced (e.g., case 3 and 4). Last, some extreme cases were also arranged (e.g., cases 5 and 6).

The macrocosm simulation results are illustrated in Figures 11–14. The proportion of health-care revenue in cases 1 and 2 changed with time (Figure 11), similar to the evaluation results obtained in the next case. Moreover, regarding the second objective, an exponential growth was observed in case 5 (Figure 12). The proportion of revenue in case 2 was slightly reduced relative to that in case 1; however, the proportion in case 2 increased considerably with time, which indicates that the insurance fee was a sensitive parameter and it alter very fast. If the current industrial structure in Taiwan remains unchanged, the insurance consumption in Taiwan would increase rapidly. In addition, the simulation results obtained when a high weight was assigned to the third objective (Figure 13) were similar to those obtained when a high weight was assigned to the first objective. However, the effect of different cases strategy was temporary, and the growth curve eventually became smooth. In case 4, the same weight was assigned to the three objectives; hence, we expected the growth rates to be consistent. However, we noted negligible differences between the simulation results obtained in Cases 4–6 (Figure 14). Accordingly, the simulation results indicate that adopting a balanced strategy for realizing multiple objectives is advantageous for health institutions.

Finally, we substituted real data into our SD overall model. When we set the same value for the GDP growth rate and population growth rate of Taiwan and considered insurance premiums to be positive, we noted that the required growth rate of the health-care in Taiwan was maintained at approximately 3.813 and that 700 care personnel were added to the workforce (approximately equal to the number of care staff members in a medium-sized medical center). This thus led to the achievement of a balance between medical demand and medical supply. These results indicate that health-care staff members in Taiwan are being overloaded at work.

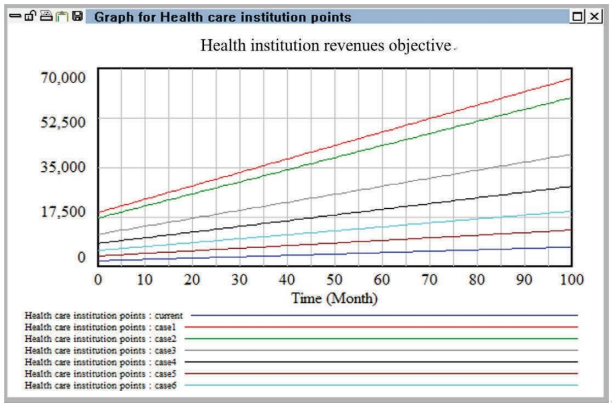


Figure 11: Comparison results among six cases in the objective one.

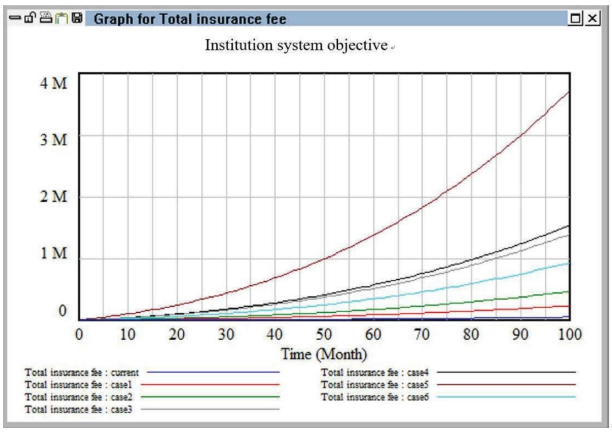


Figure 12: Comparison results among six cases in the objective two.

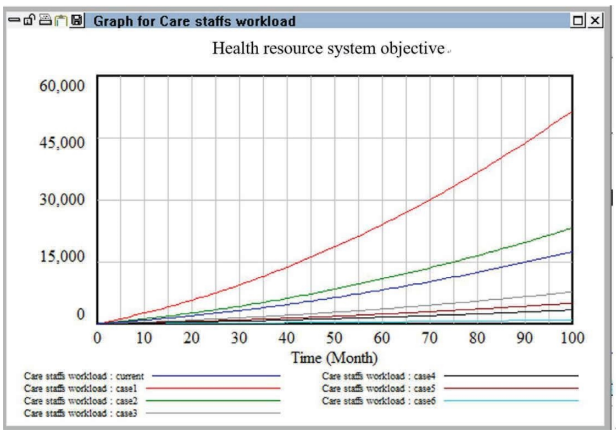


Figure 13: Comparison results among six cases in the objective three.

Taiwan's NHI program is one of the world's best health insurance systems. A financial crisis will still happen if medical expenses get more and more increasing. Therefore, to make the best use of NHI fees, cost-saving policies regarding medical institutions' scales must be implemented. Increases in medical demand and limited income can increase the workload of medical staff members, which could predispose them to emotional risks. Accordingly, we provide the following suggestions for reducing employee workload in hospitals in the same area as the case hospital:

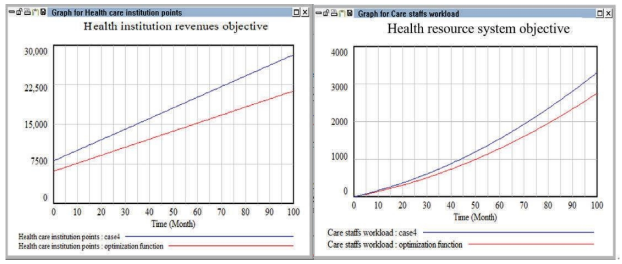


Figure 14: Health institution growth rate after multiple objectives adjusting function.

4.2.1. Optimal Distribution of the Level of Care in the Hospital. Currently, three metropolitan hospitals and five district hospitals, all of which are acute hospitals, are located in the same area as the case hospital. This area has a large population of elderly people, is positioned by the Taiwanese government as a tourism-friendly area, and is suitable for retirement; therefore, the development of diversified long-term care systems should be promoted in this area on the basis of local residents' characteristics. The development of a subacute medical care system, such as that in the United States, or an intermediate care system, such as that in the United Kingdom, can reduce the workload of health-care personnel caused by high demands in acute hospitals.

4.2.2. Development of Subsidiary Medical Institutions at Local Universities to Attract an Increasing Number of Outstanding Talents for Permanent Stay. The study interviews indicated that many doctors and other health-care personnel of the case hospital use their leave days to take degree courses or to teach in order to obtain a teaching certificate. In addition to wasting time due to commuting to their learning or teaching venues, these employees leave the case hospital after they obtain their degree. Therefore, to resolve this problem, subsidiary medical institutions can be established in local universities.

4.2.3. Promotion of Health and Improvement of Self-Care Capabilities among Elderly People. The interview results indicated that in the area where the case hospital is located, changes in family structures have forced many of the elderly people to live alone, which can negatively influence their healing. The hospitals in the area should cooperate with social workers to develop diversified health promotion programs for improving residents' health awareness, health literacy, and self-care. Such programs can reduce the medical demand.

4.2.4. Development of Integrated Information Platforms and Automatic Delivery Systems. Improving employees' work efficiency is the most effective strategy for reducing the pressure related to HR costs caused by increased medical demands. The central government can encourage hospitals in the considered area to increase their overall work efficiency through measures such as financial support, supporting policies, and incentive mechanisms. For example, the computerization of medical record systems can improve work efficiency. Data and interfaces from different systems should be continually integrated. Moreover, the hospitals in this area can develop a personalized health platform for residents, which can enable them to identify more people in need of medical services. They can apply the capitation payment system to reduce the overall medical demand. The automatic delivery system used in the case hospital can be adopted in other hospitals in the area to reduce employees' workload.

5. Conclusions and Managerial Implication

Because of advances in various forms of media, the information asymmetry between doctors and patients has decreased. However, disputes often occur between health-care personnel and patients, which can create mental pressure for health-care personnel and result in HR outflow from medical institutions. Nearly all hospitals in Taiwan are covered by the NHI program; thus, their income is limited. However, medical demands are increasing because Taiwan's elderly population is growing. In addition to medical centers and university hospitals, small and medium-sized hospitals are facing the problem of HR outflow. This phenomenon is most commonly observed among nursing staff members, for whom the workload is high but the salary is low.

Managers usually modify the HR policies of their medical institutions through trial-and-error approaches to overcome the problem caused by nursing personnel turnover. Such trial-and-error approaches can lead to the failure of HR policies, resulting in the waste of money and other resources. To overcome this problem, this study developed an SD model to simulate the effects of various HR strategies in a case hospital in different scenarios. Moreover, we constructed an SD-integrated MOP model to simulate the long-term effects of different HR strategies in the case hospital under various scenarios. This model can assist HR policymakers in medical institutions to achieve trade-offs between different objectives when attempting to reduce HR outflow.

This study has some limitations. First, data were collected from only one hospital. The scale of the developed model was reduced to cover only the field of HRs. Future studies can investigate hospitals of different levels and examine the different strategies resulting in findings between long-term care institutions and general hospitals. Moreover, the workload-related system developed for health-care staff in this study is a preliminary framework for controlling their workload. However, under the NHI system, the NHI system is more favorable to institutions that are already at a stronger position. Future studies can consider factors related to external competition from other hospitals to determine whether such factors lead to an unnatural distribution of HRs; the findings of such studies can serve as a reference for the establishment of policies in relevant institutions. Third, only fuzzy uncertainty was considered in the developed MOP model. Future studies can consider other types of uncertainty, such as stochastic uncertainty, to examine additional risk factors for HR outflow.

Data Availability

The data used to support the finding of this study are included within the article.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research was supported in part by the Ministry of Science and Technology of the Taiwan under Grant nos. MOST-107-2410-H-242-001, MOST 109-2637-H-242-001, and MOST 110-2410-H-242-001-MY2.

References

[1] J. G. Hoogerduijn, M. J. Schuurmans, M. S. Duijnstee et al., "A systematic review of predictors and screening instruments to identify older hospitalized patients at risk for functional decline," *Journal of Clinical Nursing*, vol. 16, no. 1, pp. 46–57, 2007.

[2] A. L. Tucker and A. C. Edmondson, "Why hospitals don't learn from failures: organizational and psychological dynamics that inhibit system change," *California Management Review*, vol. 45, no. 2, pp. 55–72, 2003.

[3] T. P. Hofer and R. A. Hayward, "Are bad outcomes from questionable clinical decisions preventable medical errors? A case of cascade iatrogenesis," *Annals of Internal Medicine*, vol. 137, no. 5_Part_1, pp. 327–333, 2002.

[4] N. Leveson, "A new accident model for engineering safer systems," *Safety Science*, vol. 42, no. 4, pp. 237–270, 2004.

[5] I. Aburawi, K. Hafeez, and A. Abdulsadig, "Modelling recruitment, training in workforce planning using system dynamics," *Academy of Contemporary Research Journal*, vol. 2, no. 3, pp. 86–97, 2013.

[6] C. L. Lo and H.-T. Tseng, "Applying fuzzy arithmetic operations in system dynamics for simulation and evaluation of talent retention policy model," in *Proceedings of the 7th International Conference on Business and Social Science*, Kyoto, Japan, March 2019.

[7] J. Needleman, P. Buerhaus, S. Mattke, M. Stewart, and K. Zelevinsky, "Nurse-staffing levels and the quality of care in hospitals," *New England Journal of Medicine*, vol. 346, no. 22, pp. 1715–1722, 2002.

[8] L. H. Aiken, D. M. Sloane, L. Bruyneel et al., "Nurse staffing and education and hospital mortality in nine European countries: a retrospective observational study," *The Lancet*, vol. 383, no. 9931, pp. 1824–1830, 2014.

[9] J. C. T. Yin and Ke-P. A. Yang, "Nursing turnover in Taiwan: a meta-analysis of related factors," *International Journal of Nursing Studies*, vol. 39, no. 6, pp. 573–581, 2002.

[10] J. D. Waldman, F. Kelly, S. Aurora, and H. L. Smith, "The shocking cost of turnover in health care," *Health Care Management Review*, vol. 29, no. 1, pp. 2–7, 2004.

[11] A. M. Rafferty, S. P. Clarke, J. Coles et al., "Outcomes of variation in hospital nurse staffing in English hospitals: cross-sectional analysis of survey data and discharge records," *International Journal of Nursing Studies*, vol. 44, no. 2, pp. 175–182, 2007.

[12] L. H. Aiken, S. P. Clarke, D. M. Sloane, J. Sochalski, and J. H. Silber, "Hospital nurse staffing and patient mortality, nurse burnout, and job dissatisfaction," *JAMA*, vol. 288, no. 16, pp. 1987–1993, 2002.

[13] P. Griffiths, C. Saville, J. E. Ball, J. Jones, and T. Monks, "Beyond ratios - flexible and resilient nurse staffing options to deliver cost-effective hospital care and address staff shortages: a simulation and economic modelling study," *International Journal of Nursing Studies*, vol. 117, Article ID 103901, 2021.

[14] K. P. Yang, "Relationships between nurse staffing and patient outcomes," *Journal of Nursing Research*, vol. 11, no. 3, pp. 149–158, 2003.

[15] N. Dunton, B. Gajewski, R. L. Taunton, J. Moore, and J. Moore, "Nurse staffing and patient falls on acute care hospital units," *Nursing Outlook*, vol. 52, no. 1, pp. 53–59, 2004.

[16] A. E. Rogers, W. T. Hwang, L. D. Scott, L. H. Aiken, and D. F. Dinges, "The working hours of hospital staff nurses and patient safety," *Health Affairs*, vol. 23, no. 4, pp. 202–212, 2004.

[17] R. J. Croteau, "Sentinel events, root cause analysis and the trustee," *Trustee: The Journal for Hospital Governing Boards*, vol. 56, no. 3, pp. 33-34, 2003.

[18] Y. L. Chiu, R. G. Chung, C. S. Wu, and C. H. Ho, "The effects of job demands, control, and social support on hospital clinical nurses' intention to turnover," *Applied Nursing Research*, vol. 22, no. 4, pp. 258–263, 2009.

[19] M.-J. Wang, Nurses are famously known as an angel in white, but low-paid doormats, 2012, http://www.gvm.com.tw/Boardcontent_20302.html.

[20] S. M. Nkomo, "Strategic planning for human resources Let's Get Started," *Long Range Planning*, vol. 21, no. 1, pp. 66–72, 1988.

[21] B. Haag-Heitman and A. Kramer, "Creating a clinical practice: development model," *American Journal of Nursing*, vol. 98, no. 8, pp. 39–43, 1998.

[22] M. M. Shapiro, "A Career ladder based on Benner's model an analysis of expected outcomes," *The Journal of Nursing Administration*, vol. 28, no. 3, pp. 13–19, 1998.

[23] M. Duru-Bellat, "The global auction. the broken promises of education, jobs, and incomes," *Revue Française de Sociologie*, vol. 53, no. 1, pp. 145–148, 2012.

[24] L. Rhoades, R. Eisenberger, and S. Armeli, "Affective commitment to the organization: the contribution of perceived organizational support," *Journal of Applied Psychology*, vol. 86, no. 5, pp. 825–836, 2001.

[25] E. M. Whitener, "Do "high commitment" human resource practices affect employee commitment?: a cross-level analysis using hierarchical linear modeling," *Journal of Management*, vol. 27, no. 5, pp. 515–535, 2001.

[26] Y. –C. Wei, "Do employees high in general human capital tend to have higher turnover intention? The moderating role of high-performance HR practices and P-O fit," *Personnel Review*, vol. 44, no. 5, pp. 739–756, 2015.

[27] W. F. Cascio, "The economic impact of employee behaviors on organizational performance," *California Management Review*, vol. 48, no. 4, pp. 41–59, 2006.

[28] J. W. Forrester, *System Dynamics Modelling: A Practical Appraoch*, CRC Press, Boca Raton, FA, USA, 1961.

[29] S. Dabirian, "Estimating the effects of human resource motivation on construction projects performance using system dynamics," *Journal of Structural and Construction Engineering*, vol. 8, no. 11, 2022.

[30] R. Pakdel and S. H. Hosseini, "Designing a national human resource development model using system-dynamic approach," *Quarterly Journal of Training and Development of Human Resources*, vol. 16, no. 16, p. 0, 2018.

[31] R. Yuana, E. A. Prasetio, R. Syarief, Y. Arkeman, and A. I. Suroso, "System dynamic and simulation of business model innovation in digital companies: an open innovation approach," *Journal of Open Innovation: Technology, Market, and Complexity*, vol. 7, no. 4, p. 219, 2021.

[32] H. Liu, Y. Yu, Y. Sun, and X. Yan, "A system dynamic approach for simulation of a knowledge transfer model of heterogeneous senders in mega project innovation," *Engineering Construction and Architectural Management*, vol. 28, no. 3, pp. 681–705, 2020.

[33] A. H. I. Lee, W.-C. Chen, and C. J. Chang, "A fuzzy AHP and BSC approach for evaluating performance of IT department in the manufacturing industry in Taiwan," *Expert Systems with Applications*, vol. 34, no. 1, pp. 96–107, 2008.

[34] M. Y. Bayrak, N. Celebi, and H. Taskin, "A fuzzy approach method for supplier selection," *Production Planning & Control*, vol. 18, no. 1, pp. 54–63, 2007.

[35] M. Bevilacqua and A. Petroni, "From traditional purchasing to supplier management: a fuzzy logic-based approach to supplier selection," *International Journal of Logistics Research and Applications*, vol. 5, no. 3, pp. 235–255, 2002.

[36] C. Kahraman, U. Cebeci, and Z. Ulukan, "Multi-criteria supplier selection using fuzzy AHP," *Logistics Information Management*, vol. 16, no. 6, pp. 382–394, 2003.

[37] S. M. Ordoobadi, "Development of a supplier selection model using fuzzy logic," *Supply Chain Management: International Journal*, vol. 14, no. 4, pp. 314–327, 2009.

[38] D. C. de Salles, A. C. Gonçalves Neto, and L. G. Marujo, "Using fuzzy logic to implement decision policies in system dynamics models," *Expert Systems with Applications*, vol. 55, no. 15, pp. 172–183, 2016.

[39] Pankaj, K. Seth, and Sushil, "A fuzzy set theoretic approach to qualitative analysis of causal loops in system dynamics," *European Journal of Operational Research*, vol. 78, no. 3, pp. 380–393, 1994.

[40] P. Kunsch and J. Springael, "Simulation with system dynamics and fuzzy reasoning of a tax policy to reduce CO2 emissions in the residential sector," *European Journal of Operational Research*, vol. 185, no. 3, pp. 1285–1299, 2008.

[41] F. Campuzano, J. Mula, and D. Peidro, "Fuzzy estimations and system dynamics for improving supply chains," *Fuzzy Sets and Systems*, vol. 161, no. 11, pp. 1530–1542, 2010.

[42] V. Karavezyris, K. P. Timpe, and R. Marzi, "Application of system dynamics and fuzzy logic to forecasting of municipal solid waste," *Mathematics and Computers in Simulation*, vol. 60, no. 3-5, pp. 149–158, 2002.

[43] J. Song, D. Song, and D. Zhang, "Modeling the concession period and subsidy for BOT waste-to-energy incineration

projects," *Journal of Construction Engineering and Management*, vol. 141, no. 10, Article ID 04015033, 2015.

[44] I. J. Orji and S. Wei, "An innovative integration of fuzzy-logic and systems dynamics in sustainable supplier selection: a case on manufacturing industry," *Computers & Industrial Engineering*, vol. 88, pp. 1–12, 2015.

[45] Y. C. Chang and T. T. Ko, "An interactive dynamic multi-objective programming model to support better land use planning," *Land Use Policy*, vol. 36, pp. 13–22, 2014.

[46] J. Xu, Y. Deng, and L. Yao, "Sustainable development-oriented industrial restructuring modeling and analysis: a case study in Leshan," *Clean Technologies and Environmental Policy*, vol. 16, no. 2, pp. 267–279, 2014.

[47] Z. Wu and J. Xu, "Predicting and optimization of energy consumption using system dynamics-fuzzy multiple objective programming in world heritage areas," *Energy*, vol. 49, pp. 19–31, 2013.

[48] N. S. Sabounchi, K. P. Triantis, S. Sarangi, and S. Liu, "Dynamic simulation modeling and policy analysis of an area-based congestion pricing scheme for a transportation socioeconomic system," *Transportation Research Part A: Policy and Practice*, vol. 59, pp. 357–383, 2014.

[49] J. D. Sterman, "Learning in and about complex systems," *System Dynamics Review*, vol. 10, no. 2-3, pp. 291–330, 1994.

[50] J. D. Sterman, *Business dynamics: systems thinking and modeling for a complex world*, Vol. 19, Irwin/McGraw-Hill , Boston, MA, USA, 2000.

[51] J. Forrester, "Jay Forrester Introduction," 2014, http://mitsloan.mit.edu/faculty/detail.php?in_spseqno=41467.

[52] J. K. Doyle and D. N. Ford, "Mental models concepts for system dynamics research," *System Dynamics Review*, vol. 14, no. 1, pp. 3–29, 1998.

[53] B. Hu and H. R. Vetter, "System dynamical analysis for interdisciplinary research on human resource development," in *Proceedings of the 2008 International Conference of the System Dynamics Society*, The System Dynamics Society, Athens, Greece, July 2008.

[54] H. L. Dreyfus and S. E. Dreyfus, "The ethical implications of the five-stage skill-acquisition model," *Bulletin of Science, Technology & Society*, vol. 24, no. 3, pp. 251–264, 2004.

[55] M. E. Charlson, P. Pompei, K. L. Ales, and C. R. MacKenzie, "A new method of classifying prognostic comorbidity in longitudinal studies: development and validation," *Journal of Chronic Diseases*, vol. 40, no. 5, pp. 373–383, 1987.

[56] R. G. Coyle, *System Dynamics Modeling. A Practical Approach*, Chapman & Hall, London, UK, 1996.

Hindawi
Journal of Healthcare Engineering
Volume 2022, Article ID 9969604, 12 pages
https://doi.org/10.1155/2022/9969604

Research Article

Evaluating Organ Donation Decision in ICU Patients' Families by Analytic Network Process Approach

Chia-Lun Lo,[1] **Hsiao-Yun Chang**,[2] **and Guang-Mao Lee**[3]

[1]*Department of Health-Business Administration, Fooyin University, Kaohsiung City 83102, Taiwan*
[2]*Department of Nursing, Chang Gung University of Science and Technology, Taoyuan City 33303, Taiwan*
[3]*Department of Nursing, Kaohsiung Municipal United Hospital, Kaohsiung City 80457, Taiwan*

Correspondence should be addressed to Chia-Lun Lo; allenlo.tw@gmail.com

Received 21 May 2021; Revised 28 January 2022; Accepted 14 March 2022; Published 15 April 2022

Academic Editor: Chia-Yu Lin

The imbalance between supply and demand for organs has been a global crisis, despite the efforts of transplant coordinators from healthcare institutions to promote donor registration. Because the patient's family has legal rights over the patient's remains, they can easily undermine any efforts spent on organ procurement by simply refusing the patient's consent before death in practice. Most related studies seldom mention the decision-making on organ donation from patients' families. The objectives of this study are to find what are the priorities of those factors acting as the pillars of organ donation by patients' families. This study applied the analytic network process (ANP) to the prioritization factors contributing toward the willingness of families to donate organs of intensive care unit patients. The purposive sampling method used structured questionnaires and ANP questionnaires to enroll 180 patients' families from five intensive care units who met the criteria in the regional teaching hospital of southern Taiwan. Through the ANP analysis, it was found that when family members made organ donation decisions, the weights of the four domains are as follows: psychology—47.6%, externality—20.3%, spirituality—19.7%, and physiology—12.3%. The main decision-making factors that influenced the weighting factors were "attitude" (31.5%), "physician's experience" (0.88%), "religion" (19.3%), and "organ selection" (31.9%). These results could assist organ donation teams to take the best strategies for persuading people to agree with organ donation and formulating an individual organ donation plan.

1. Introduction

According to the 2017 data of the Health Resources and Services Administration, an agency of the US Department of Health and Human Services, only 33,612 (28.16%) of the 119,362 individuals awaiting organ transplant have successfully received the transplant. The widening gap between the number of organ donors and recipients is a considerable challenge for governments worldwide. In Taiwan, the Human Organ Transplant Act, which was promulgated in 1987, clearly stipulates that a transplant operation must be performed only after the organ donor has been certified dead by his/her attending physician. Typically, brain death is used as a legal definition of death, and the guidelines for the determination of brain death specify that brain death must be determined when the individual in question is in an intensive care unit (ICU) and is receiving constant perfusion through life support equipment capable of monitoring the individual's hemodynamic parameters. The organs of an individual who is confirmed as brain dead can be donated with the individual's prior consent or with his/her family's consent. Different systems and laws govern organ donation throughout the world, including opt-out (presumed consent) and opt-in presumptive approaches. In fact, in most countries, such approaches are often not rigorously applied because health-care staff members have the family's well-being at heart. Caring for a bereaved family who is going through the unique and distressing experience of a loved one being pronounced brain dead is a responsibility for health-care staff.

The patient's wishes determine the organ donation process. If the patient's wishes are unknown, health-care

staff in the given institution deems that decisions should be made by the patient's family. Difficulties can also emerge when the wishes of the patient are inconsistent with those of the patient's family because health-care staff members are obliged to follow the wishes of the patient, potentially creating legal or ethical tensions with the patient's family [1]. In practice, because the patient's family members have legal rights over the patient's remains, they can easily undermine any efforts expended on organ procurement by refusing consent on behalf of the patient before death [2, 3]. Therefore, identifying the factors that affect family consent to organ donation is of utmost importance.

However, the situation is complex, with the family obliged to make decisions about organ donation within a short time despite needing to discuss the matter with other relatives and friends. Extensive communication and explanation tend to characterize this stage of the organ donation process. Identifying which opinion is the most legitimate one among those of relatives is difficult [4, 5]. Family members' subjective opinions on organ donation can also be influenced by their distinct social, cultural, and religious backgrounds. Some studies have noted that families' education level, funerary customs, and openness of communication with medical personnel are factors affecting their decisions about a patient's organ donation [6, 7]. Family members thus play a key role in the decision on whether the patient's organs will be donated because they overrule the patient's wishes in some cases.

Knowledge of the patient's wishes is generally the strongest and most consistent predictor of donation. A family's awareness of its relative's donation wishes is strongly associated with honoring those wishes [8–16]. In addition, studies on organ shortage have generally focused on ICU nursing personnel's attitude toward and knowledge of organ donation, indicating that individuals with extensive organ knowledge and a positive attitude are more inclined to accept organ donation than others [7, 17–19]. Siminoff et al. [18] reported that attitude toward organ donation varied by professional background; a positive attitude toward and consenting to organ donation were positively correlated among individuals from a medical background. Notably, up to 78.3% of the respondents to a survey deemed their immediate family members to be the only ones (apart from themselves) with the right to decide whether their organs might be donated [18]. Nevertheless, a patient's family refusing to comply with the patient's wishes for organ donation despite that patient signing an organ donor card is not an uncommon situation in clinical settings, especially given that the law protects the rights of patients' families over the patients' remains. Therefore, identifying the factors affecting families' decisions regarding organ donation at the critical moment is crucial to the success of organ procurement and could help organ procurement coordinators and medical staff minimize family distress, fulfil the patient's wishes, and increase the rate of organ donation.

The literature does not provide sufficient evidence to determine the opinions of patients' families in deciding on donating organs of their family member. Making a quick decision on whether to donate the organs of a loved one is a difficult, complicated, and multifaceted process. In this study, we assigned weights to the main factors affecting organ donation decisions made by families of ICU patients. Moreover, the relevant literature on psychological factors and cause-and-effect relationships (the qualitative and quantitative perspectives, respectively) was reviewed. Following this literature review, we combined various factors to propose a practical methodology. We employed a multiple-criteria decision-making (MCDM) model to prioritize the factors affecting organ donation. Many MCDM calculation methods can be used to derive values for prioritization, thereby solving decision-making problems. The analytic hierarchy process (AHP), proposed by Stoeckle [20], was designed for supplier selection and involves compiling comparative evaluation criteria of supplier performance; the AHP has been used widely to solve complex real-world problems. The AHP typically considers the unidirectional relationships among the factors and structures in the problem as a hierarchy. By contrast, if the problem context is a network where goal, criteria (and, where applicable, subcriteria), and alternatives are considered, the analytic network process (ANP) considers such factors to be nodes of the network, with bidirectional relationships.

Therefore, this study employed the ANP as the basic framework and used scientific modeling to explore the factors affecting the decisions of families regarding donating organs of family members in an ICU. In the ANP framework, the relationships between and the weights of diverse factors were compared and ranked on the basis of importance. The findings can offer assistance to organ procurement programs, enabling greater success in organ procurement to be achieved. The purpose of this study was to identify the prominent psychological drivers of patients' families providing consent on behalf of ICU patients regarding organ donation; to this end, the following research questions were addressed:

(1) What are the prevailing attitudes of the families of ICU patients in relation to organ donation?

(2) What are the main factors affecting family acceptance of ICU patients' organ donation?

(3) What is the priority level of each identified factor influencing organ donation consent?

2. Literature Review

Organ transplantation is a treatment method for end-stage organ failure whereby a patient's life is extended and quality of life is improved. Organ donation refers to the donation of body organs or tissues, which are surgically removed from a donor and transplanted to one or multiple recipients. In real-world settings, three types of donation systems are commonly used in the procurement of organs: (1) the opt-out (presumed consent), (2) opt-in, and (3) presumptive approaches. The findings of different countries after implementation of these systems imply that their effectiveness largely depends on how the systems address the factors affecting the willingness to donate, which may include educational, cultural, and social backgrounds [10].

Organ donation is a sensitive topic, particularly from a traditional Chinese cultural perspective and in relation to patients who are at the end stage of their lives. Despite studies not having reached a consensus on the factors influencing the decision to donate organs, several candidate factors have been proposed and can be categorized into demographic, physiological, psychological, spiritual, and external factors as follows.

2.1. Demographic Factors. Araujo and Siqueira [11] and Goz et al. [12] asserted that age does not affect attitude toward organ donation, but Cohen et al. [13] reported that age and attitude were significantly positively correlated. Furthermore, in one study, women were more positive about organ donation than men [14], whereas Siminoff, Gordon, Hewlett, and Arnold argued that men are more likely to donate organs because they bear greater responsibility than women from the perspective of traditional ideals [8]; Jung reported no differences in organ donation based on sex [15].

Numerous studies have suggested that educational attainment influences intention regarding organ donation [11, 16]. Pouraghaei et al. [16] also reported that individuals in full-time employment were more likely to have organ donation intention. Akgun et al. [17] observed that respondents with a medical background were more likely to intend to donate than those with a background in another profession. Finally, organ donation intention tends to be less prevalent in close-knit families because family members cannot bear to damage the remains of loved ones or refuse donation to avoid intrafamily conflict [11].

2.2. Physiological Factors. Siminoff et al. [18] reported that family-decided organ donation was more prevalent (65.1%) when the patient died of external trauma. Thus, a correlation may exist between cause of death and organ donation intention. In Asia, preserving the remains of the deceased, particularly their sensory organs, is a long-standing custom. In clinical settings, the family commonly agrees to donate any part of the deceased except the corneas in accordance with the rationale that the deceased will be unable to journey to the afterlife effectively if deprived of eyesight. However, Siminoff et al. [18] reported that the consent of the family was unaffected by the body part desired for donation.

2.3. Psychological Factors. If the deceased has not left a will or an advance directive, the family of the deceased is granted legal ownership of the deceased's remains. Therefore, the family's attitude is a major factor affecting the decision regarding organ donation. Numerous studies have confirmed that individuals with a positive attitude toward organ donation are more likely to have organ donation intention [7, 19, 20]. Nursing personnel in ICUs generally has a positive attitude toward organ donation, and patients' families who consult nurses on such matters may be influenced by their positive attitude.

2.4. Spiritual Factors. Hsieh et al. [21] listed a number of causes for patients' families refusing organ donation, including traditional customs, family concerns, value systems, cognitive differences, and organs being deemed medically unsuitable for transplantation; they identified custom as the primary cause of most people's traditional belief that the remains of the deceased must be left intact or fear of bringing greater suffering (e.g., mutilation) to their loved one after death. Davison and Jhangri [22] concurred that cultural beliefs have a considerable effect on organ donation intention. Holman et al. [23] further delineated the effects according to religious belief; Catholics (77%) and Orthodox Christians (73%) were more likely to have organ donation intention than Protestants (43%). Conversely, Goz et al. [12] reported that religious belief does not affect people's organ donation decision. In the Netherlands, Witkamp et al. [24] interviewed the families of critically ill patients and reported that those who had discussed death with the patient were more open to the idea of organ donation than those who had not. Simpkin et al. [25] and Wu [26] both noted that discussing death with family members gives individuals more insight into death, reduces their anxiety about death, and affects their intention regarding organ donation.

2.5. External Factors. Organ procurement typically involves medical professionals explaining the relevant details to a patient's family in a quiet place. Gortmaker et al. [27] observed that conducting such discussions in a nursing station offered a greater likelihood of success than in the corridor or the ward (56%, 52%, and 30%, respectively). Simpkin et al. [28] contended that, if possible, organ procurement dialogue with families should be conducted in a secluded place to improve the likelihood of success. Niles and Mattice [29] conducted a retrospective study and noted that although attempting organ procurement discussion before versus after the death of a patient did not exert a significant effect on the granting of consent, doing so immediately upon the death of a patient reduced family consent to 32%–37%. Instead, the ideal time for the organ procurement discussion was the point at which brain death was declared (65.4%).

3. Analytic Network Process

This study employed the ANP as its main framework to analyze the factors that influence the organ donation decisions of patients' families. To resolve problems involving uncertainty and multiple criteria, Saaty [30] developed the AHP to prioritize such criteria in decision-making. The AHP uses stratified and independent approaches to solve complex decision-making problems by identifying the weights of decision alternatives by using quantified calculation methods. However, decision-making problems are highly diverse and complex, and one limitation of the AHP is that the factors that can influence a decision must be independent. Therefore, in some cases, the AHP is not applicable. To address this deficiency, Saaty extended the AHP to develop a new decision-making framework: the ANP. The ANP has been described as a special case of the AHP [31]. The AHP

considers the unidirectional relationships among the factors and structures in the problem as a hierarchy, whereas the ANP structures the problem context as a network in which the goals, criteria (and, where applicable, subcriteria), and alternatives are considered nodes of the network. Thus, unlike the AHP, the ANP allows for loops and feedback between nodes to represent interdependency among the factors. The ANP is based on the pairwise comparisons of the AHP whereby criteria are compared with each of the decision alternatives, necessitating an additional set of comparisons where alternatives are pairwise-compared against each criterion. The ANP allows interactions and feedback among decision criteria to be included through the introduction of evaluation scales into the decision model. The ANP can systematically solve problems involving multiple decision-making criteria by employing a nonlinear network structure based on the problem type [31]. Moreover, the ANP is commonly used to solve multiple-criteria decision problems that cannot be defined using a hierarchical structure, such as product planning, strategic decision-making, and other multidimensional analyses. In these cases, relations are found not only among criteria in the same hierarchy but also among criteria in different hierarchies. Therefore, the ANP was the more appropriate framework than the AHP, given our research goal, and it was adopted in this study accordingly.

The main purposes of developing a network structure were to confirm the problem, clearly describe and identify the decision criteria, define the factors involved with each criterion and subcriterion, identify the interrelations among the criteria, and graph the ANP model. The flow chart of the proposed approach is given in Figure 1. There are four major steps in the ANP, namely, (1) development of a network structure and creation of pairwise comparison matrices, (2) calculation of the relative weights of those matrices, (3) construction of a supermatrix, and (4) selection of the best solution [31]. In accordance with the recommendation of Saaty [31], the following steps should be followed to perform the ANP.

3.1. Developing Pairwise Comparison Matrices. Saaty [31] asserts that the relative importance of n factors (i.e., subcriteria) with respect to a specific element in the immediate upper level should be assessed through a pairwise comparison that utilizes a 9-point scaling system. Patients' families were the targeted respondents for the pairwise matrix comparisons. The pairwise comparison matrices were constructed, the relative importance and weights of the criteria were calculated, and pairwise comparisons of the criteria were made on the basis of a scale from 1 to 9, as detailed in Table 1. The comparisons included comparisons of criteria with factors and comparisons among factors, which included comparisons among factors of the same criterion and among factors of different criteria. The participants were invited to compare two criteria, or two factors—say A and B—in a certain situation. If they believed A was more important than B, they recorded a score of 5–9 on the scale. If they disagreed that A was more important, they recorded a score of 1–4.

3.2. Normalizing the Matrices and Obtaining Priority Vectors. The results of each pairwise comparison were arranged in a pairwise comparison matrix (**W**). Matrix **W** was normalized by dividing each element of the matrix by its column sum. To obtain a priority vector, the rows of **W** were averaged using the arithmetic mean.

3.3. Evaluation of Consistency. Similarly, after the participants had given answers for all items, the obtained data were used to create $n \times n$ comparison matrices. To prevent errors in the research results, the comparison matrices were strictly consistent before and after. Thus, evaluating the consistency of the scale results by using the consistency index (CI) and consistency ratio (CR) was essential. This was how the consistency of the potential decisions made could be assured. The decision-makers' decisions were consistent when the CR was under 0.1.

3.4. Developing the Weighted Supermatrix. After confirming the consistency by using the CI and CR, the eigenvector of each matrix was obtained and used as the weight of the given matrix. An unweighted supermatrix could be constructed using the weight of each matrix according to the interrelations of the factors. Then, the supermatrix was normalized to create the weighted supermatrix. This supermatrix could efficiently solve the problem of criterion dependence. It was composed of multiple submatrices. Each submatrix was composed of the interrelations among the relevant factors. By comparing it with the other matrices, the corresponding eigenvectors could be obtained as the weight of the submatrix. The convergent values of the limit supermatrix were the weights of the corresponding criteria, and these weights represented the relative importance of the criteria (constructs) and factors used to evaluate the alternatives for the research problem. The alternatives could then be prioritized on the basis of these weights.

4. Methods

In this study, a literature review was performed to identify factors affecting the organ donation attitude of families with a member in an ICU. In the first step, we adopted a cross-sectional questionnaire to examine the attitude of the respondents; then the findings were integrated with our clinical experience to produce study variables of the ANP questionnaire to investigate the patient's families with a positive attitude who are more likely to donate organs by the ANP framework questionnaire in the second step. Because the research framework was based on the ANP in this study, the Super Decision 2.6.0 (http://www.superdecisions.com/) and WEKA 3.8.0 open-source data mining (http://www.cs.waikato.ac.nz/ml/weka/) software packages were used to calculate the dependence and feedback between the constructs and factors, upon which the weights of factors were determined to establish a decision tree.

4.1. Samples. Participants were recruited from five types of ICU—the subacute respiratory, medical, neurosurgery, surgical, and cardiac ICUs—of a regional teaching hospital

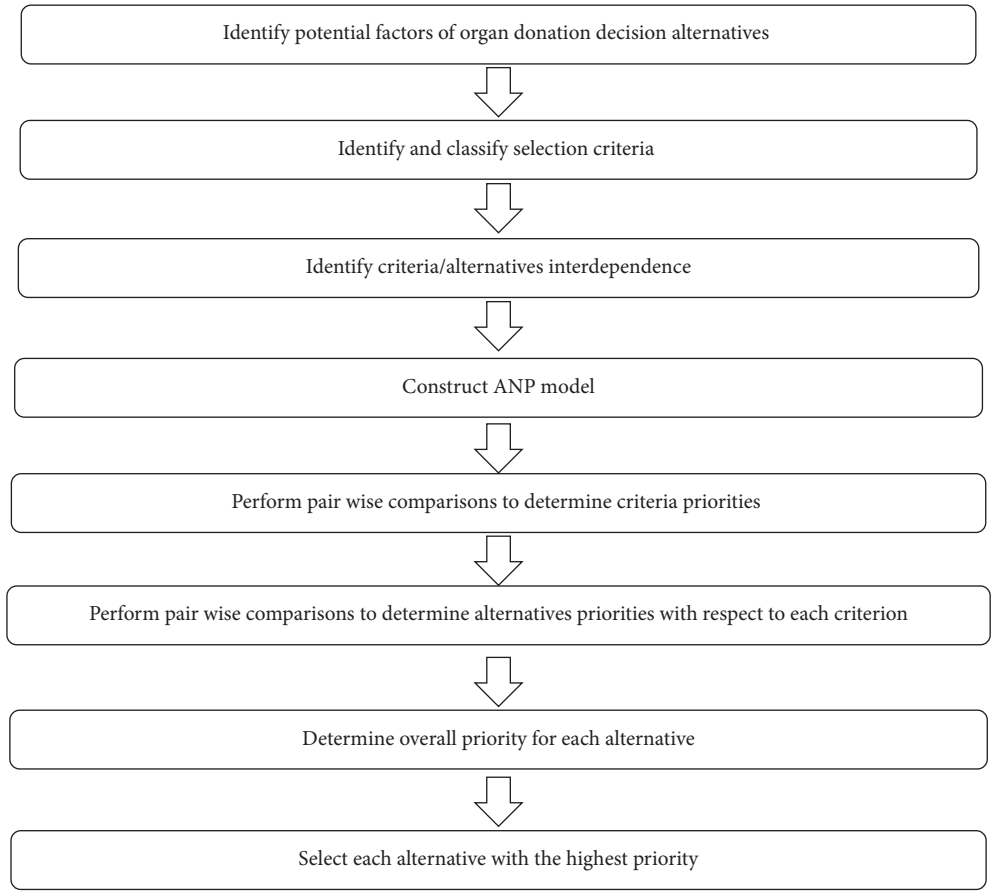


Figure 1: Process flowchart for organ donation decision-making with ANP methodology.

Table 1: Pairwise comparison for AHP or ANP preferences.

Importance	Definition	Explanation
1 2	Equal importance Weak or slight	Two activities contribute equally to the objective.
3 4	Moderate importance Moderate plus	Experience and judgment slightly favor one activity over the other.
5 6	Strong importance Strong plus	Experience and judgment strongly favor one activity over the other.
7 8	Very strong or demonstrated importance Very, very strong	One activity is very strongly favoured over the other; its dominance is demonstrated in practice.
9	Extreme importance	The evidence favoring one activity over the other is of the highest possible order of affirmation.

in Taiwan. Each participant, who was the main decision-maker in the respective patient's family, was required to indicate his/her attitude toward organ donation by using a Likert Scale. Then, quartile clustering was performed to select a most positive group to answer the ANP questionnaire with 4 constructs and 18 factors pertaining to organ donation. When estimated using *G**Power 3.1 with a significance level of 0.05, power of 0.80, and effect size of 0.15, the minimum sample size required was determined to be 150 participants [32]. In consideration of some questionnaires potentially not being returned, this study's sample size was determined as 180 participants.

The inclusion criteria were (1) being the main decision-maker in the family of a patient in an ICU and (2) being an adult aged ≥20 years. The exclusion criteria were (1) having received a diagnosis of a mental disorder and (2) being deaf or mute or having comprehension impairment. Discussing the prospect of participating in the study with ICU patients' family members was difficult when they were at the bedside of the patient, in a bad mood, or feeling rushed. Hence, observing the sentiment of patients' families first and then selecting an appropriate and comfortable time to ask for their participation in our study were essential. We spent nearly 1 year collecting all the study questionnaires.

Table 2: Descriptive statistics of attitude about organ donation.

Positively/negatively	Items	$\mu \pm \sigma$
Positively worded items	"*Organ donation is a good thing and the right thing.*	4.30 ± .776
	Donating the organs of a loved one makes me feel they are still alive; it is a continuation of their life.	4.14 ± .891
	Organ donation helps others by saving their lives.	4.39 ± .728
	Organ donation is an expression of altruistic love.	4.21 ± .784
Negatively worded items	*I do not want to donate the organs of my loved one because I am worried about the opinions of family and friends.*	3.47 ± 1.043
	I do not want to donate the organs of my loved one because I'm worried that the hospital would not do their best in saving my loved one.	3.66 ± 1.099
	Donating organs means the remains of my loved one will be incomplete; that is why I reject it.	3.74 ± 1.015
	Our bodies—to every hair and bit of skin—are received by us from our parents, and we must not presume to injure or wound them; that is why organ donation is unacceptable.	3.74 ± 1.015

Fortunately, only eight individuals declined our request to participate before the desired 180 completed questionnaires were obtained.

4.2. Research Process. The research process of the present study comprised three phases. First of all, the topic was determined, and a thorough understanding of the relationship of the variables was gained on the basis of literature review and our clinical experience. In the second phase, after literature review was conducted, variables were identified and scales appropriate for assessing attitude toward organ donation were collected to establish the ANP questionnaire of this study. In the third phase, the organ donation attitude questionnaire and the ANP questionnaire were combined to perform an investigation on the factors affecting the organ donation decisions.

4.3. Research Instrument. The research instrument was self-developed on the basis of the results of literature review and our own clinical experience. It addressed three aspects, namely, demographic attributes, attitude toward organ donation, and factors affecting the organ donation decision of families with a member in an ICU. The demographic attributes recorded were age, gender, marital status, educational attainment, occupation, and religious belief as well as family relations and relationship with the patient. The organ donation attitude questionnaire was compiled by Yen [33]. It was employed to gauge the participants' attitudes toward organ donation. Specifically, the questionnaire consisted of eight items covering egoism, altruism, and constraints and exhibited a Cronbach's α of 0.84. A Likert Scale was used for scoring, with a high score denoting a positive attitude.

In the third phase of the research process, we developed a questionnaire inquiring into the factors affecting the organ donation decisions of families with a member in an ICU by applying the ANP method based on experience and the results of the literature review. The 1–9 scoring system proposed by Saaty [31] was adopted for pairwise comparisons. The initial draft of the questionnaire, which consisted of 18 factors associated with 4 constructs, was reviewed by six experts, who provided valuable feedback and rated the items by their relevance and appropriateness in relation to language and scoring. The content validity index of the items averaged 0.87, exceeding the benchmark of ≥0.8 [34]. The method also calculates a consistency ratio (CR) to verify the coherence of the judgments, which must be about 0.10 or less to be acceptable. Mathematical foundations of the AHP can be found in Saaty [31]. The questionnaire was then issued to participants for data collection.

4.4. Ethical Considerations. This study was approved by the institutional review board of E-Da hospital (approval no: AF08-008). Before answering the questionnaire, the participants were comprehensively informed of their rights and the purpose of the study and were also required to sign an informed consent form. To protect the rights of the ICU patients and their families, the questionnaire was anonymous, and identifiers such as patient identification codes were encoded with serial numbers such that individuals could not be identified.

5. Results

A cross-sectional research design was adopted to explore the factors affecting the organ donation decision of families with a member in an ICU. A total of 180 participants were included; among them, 68.9% were females, 66.1% were married, and 57.8% had an educational attainment of college or higher. Service industries were the most prevalent occupation, accounting for 36.1% of the participants. The participants were aged 21–87 years, with a mean age of 40.43 years. In terms of religious belief, most participants were either Taoist (69; 38.3%) or Buddhist (49; 27.2%). Regarding their relationships with the patient, 20 (11.1%), 18 (10%), 26 (14.4%), and 75 (41.7%) participants were the patient's parent, sibling, partner, and child, respectively. Most of the participants (96.7%) reported their family relations as being close. Only 15 (8.3%) participants reported that the ICU patient from their family had signed an organ donor card. Most of the participants (117; 65%) were positive about organ donation, but only 39 (21.7%) had discussed organ donation with other family members. A total of 115 (63.9%) participants asserted that they would support their family members' decision to donate their organs at their time of death.

In the organ donation attitude questionnaire, the participants gave scores of 31.67 ± 5.34 out of 40 (a higher score

Table 3: Weights of constructs and factors of decision-making for organ donation.

Constructs	Factors	Factor's weight	Priority of factor	Construct's weight	Rank of dimension
Spiritual	Faith	0.192601	V	0.197258	3
	If mention dead issue with family	0.038521			
Physiological	Organ selection	0.31903	V	0.123799	4
	Cause of death	0.05139			
External	Transplant coordinator background	0.003948	V	0.203193	2
	Organ procurement timing	0.003190			
	Nursing care experience	0.005554			
	Physician care experience	0.008866			
	Hospital ranking	0.001857			
	Hospital reputation	0.002569			
	Privacy environment	0.001858			
Psychological	Positive attitude	0.31589	V	0.475751	1
	Negative attitude	0.31589			

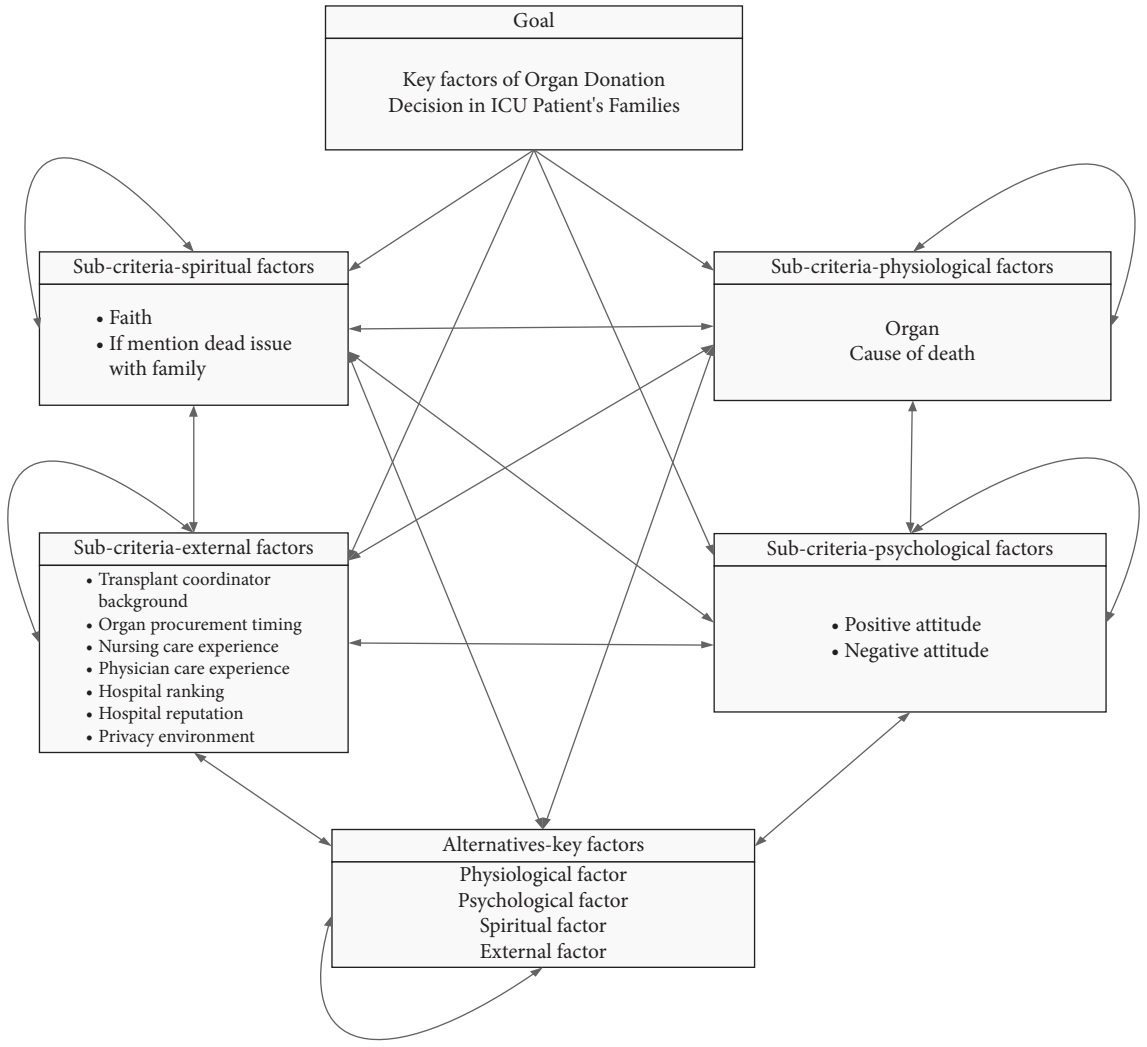


Figure 2: Network structure of ANP.

indicated a more positive attitude). Detailed information is presented in Table 2.

The results of bivariate analysis indicated that age and attitude toward organ donation were significantly negatively correlated ($r = -0.221$ and $p = 0.003$). Other parameters that significantly correlated with organ donation attitude were gender, marital status, educational attainment, family relations, and responses to the following questionnaire items:

Table 4: Unweighted supermatrix in the ANP model.

Constructs	Constructs	Spiritual		Physiological				External					Psychological	
	Factors	1	2	3	4	5	6	7	8	9	10	11	12	13
Spiritual	Faith (1)	0	1	0.75	0.8	0.8	0.75	0.8	0.833	0.8	0.833	0.833	0.857	0.8
	If mention dead issue with family (2)	1	0	0.25	0.2	0.2	0.25	0.2	0.167	0.2	0.167	0.167	0.143	0.2
Physiological	Organ selection (3)	0.197	0.197	0.875	0.197	0.125	0.197	0.197	0.197	0.197	0.197	0.197	0.197	0.197
	Cause of death (4)	0.051	0.051	0.125	0.051	0	0.051	0.051	0.051	0.051	0.051	0.051	0.051	0.051
External	Transplant coordinator background (5)	0.1	0.167	0.333	0.167	0.017	0	0.128	0.158	0.08	0.083	0.148	0.09	0.09
	Organ procurement timing (6)	0.752	0.752	0	0.752	0.875	0.752	0.752	0.752	0.752	0.752	0.752	0.752	0.752
	Nursing care experience (7)	0.108	0.167	0	0.167	0.107	0.138	0	0.179	0.146	0.229	0.269	0.148	0.149
	Physician care Experience (8)	0.136	0.167	0	0.167	0.519	0.209	0.222	0	0.22	0.145	0.413	0.209	0.21
	Hospital ranking (9)	0.186	0.167	0.333	0.167	0.035	0.055	0.054	0.147	0	0.054	0.067	0.045	0.063
	Hospital reputation (10)	0.216	0.167	0	0.167	0.071	0.085	0.08	0.103	0.051	0	0.103	0.062	0.042
	Privacy environment (11)	0.254	0.167	0.333	0.167	0.252	0.513	0.517	0.412	0.503	0.489	0	0.447	0.446
Psychological	Positive attitude (12)	0.875	0.833	1	1	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0	1
	Negative attitude (13)	0.125	0.167	0	0	0.5	0.5	0.5	0.5	0.5	0.5	0.5	1	0

Table 5: Weighted supermatrix in the ANP model.

Constructs	Constructs	Spiritual		Physiological				External					Psychological	
	Factors	1	2	3	4	5	6	7	8	9	10	11	12	13
Spiritual	Faith (1)	0	0.04	0.12	0.13	0.13	0.36	0.38	0.39	0.38	0.39	0.39	0.17	0.16
	If mention dead issue with family (2)	0.04	0	0.04	0.03	0.03	0.12	0.09	0.08	0.09	0.08	0.08	0.03	0.04
Physiological	Organ selection (3)	0.02	0.02	0.12	0.03	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	Cause of death (4)	0	0	0.02	0.01	0	0	0	0	0	0	0	0	0
External	Transplant coordinator background (5)	0.05	0.08	0.13	0.07	0.01	0	0.01	0.02	0.01	0.01	0.02	0.01	0.01
	Organ procurement timing (6)	0.07	0.07	0	0.1	0.12	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04
	Nursing care experience (7)	0.05	0.08	0	0.07	0.04	0.02	0	0.02	0.02	0.03	0.03	0.01	0.01
	Physician care Experience (8)	0.07	0.08	0	0.07	0.21	0.02	0.02	0	0.02	0.02	0.05	0.02	0.02
	Hospital ranking (9)	0.09	0.08	0.13	0.07	0.01	0.01	0.01	0.02	0	0.01	0.01	0	0.01
	Hospital reputation (10)	0.11	0.08	0	0.07	0.03	0.01	0.01	0.01	0.01	0	0.01	0.01	0
	Privacy environment (11)	0.13	0.08	0.13	0.07	0.1	0.06	0.06	0.05	0.06	0.05	0	0.04	0.04
Psychological	Positive attitude (12)	0.19	0.18	0.11	0.11	0.06	0.11	0.11	0.11	0.11	0.11	0.11	0	0.43
	Negative attitude (13)	0.03	0.04	0	0	0.06	0.11	0.11	0.11	0.11	0.11	0.11	0.43	0

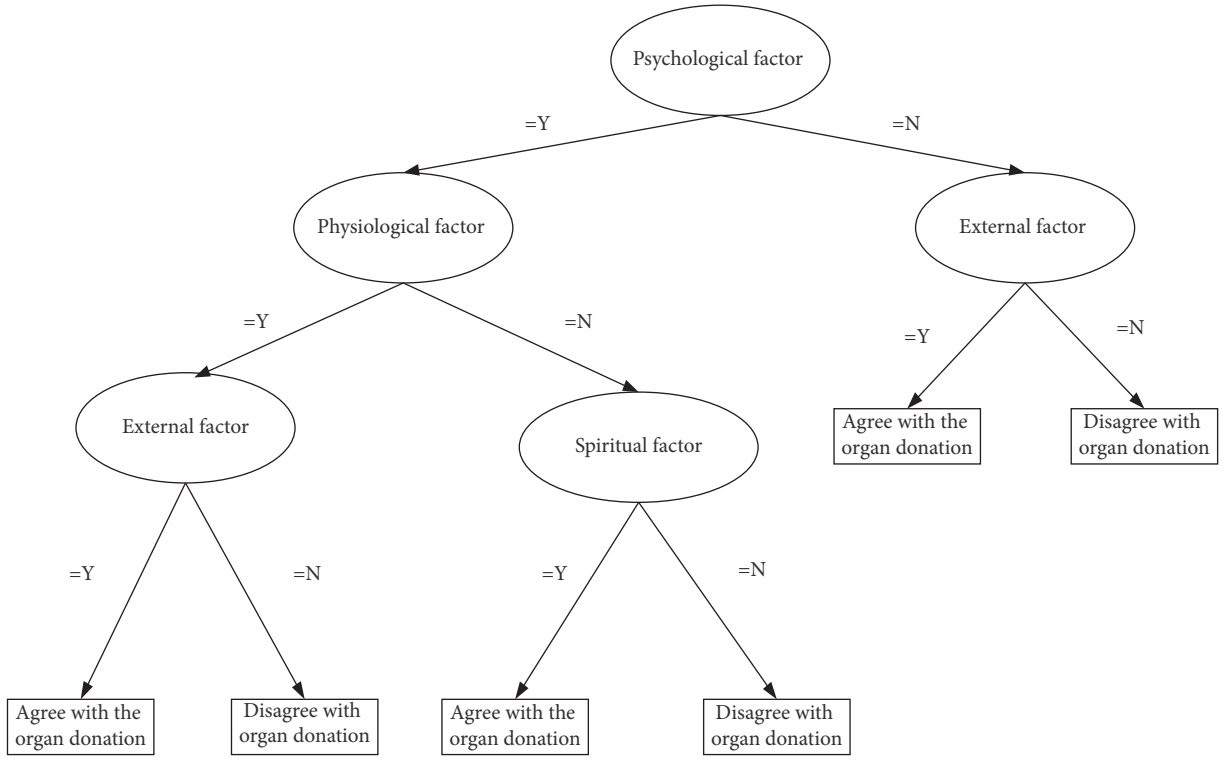


Figure 3: Decision tree for organ donation of ICU patient's family.

"*Have you signed an organ donor card?*," "*Do you agree that organ donation is a personal decision?*," "*Have you discussed organ donation with the patient?*," and "*Do you agree with the patient's organ donation intention and attitude toward organ donation?*" The results of one-way analysis of variance indicated that occupation ($F = 1.556$ and $p = .152$), religious belief ($F = 1.091$ and $p = .362$), and the relationship with the patient ($F = 1.571$ and $p = .171$) were not significantly correlated with attitude toward organ donation or organ donation intention.

Clustering was performed with the scores obtained using the organ donation attitude questionnaire. Quartile clustering was applied to the attitude scores, and the cluster with the highest score ($n = 50$) was selected for an ANP-based analysis to determine the main factors affecting the organ donation decision of the participants who held a positive attitude toward organ donation. Pairwise comparison was conducted on the factors in the four constructs to establish factor weights on the basis of the internal dependence they exhibited. Table 3 presents the priority values of the factors associated with attitude toward organ donation; the factors were ranked by importance. The table thus reveals the constructs and factors that play a major role when a patient's family must make a decision regarding organ donation.

A limit supermatrix, which was obtained from the computation results of Super Decisions, was used to determine the order of priority of the constructs and factors (Figure 2). All the criteria and subcriteria and the details are given a code letter. There, codes given in tables 4 and 5 will be used in the supermatrix. The weights in Table 3 were sorted, revealing the psychological construct (0.475751) as the most influential, followed by the external factors (0.203193) and spiritual (0.197258), and physiological (0.123799) constructs. For the psychological construct, the factor receiving the highest priority was attitude toward organ donation (0.31589). For the external factor's construct, the key factor was the physician's care experience (0.008866); for the spiritual construct, the key factor was religious belief (0.192601); for the physiological construct, the main factor was organ selection (0.31903). In summary, attitude toward organ donation (part of the psychological construct) was the most critical factor, whereas organ selection (a physiological construct) was the least critical factor.

The results of the clusters with the highest score were computed in WEKA software version 3.8.3 to investigate the classification and establishment of the decision tree by the C.4.5 [35], cart [36], and random tree [37] classifiers. The experimental results indicated that the random tree algorithm was the most suitable classifier. Therefore, random tree was used for the development of the decision tree, which could aid clinical procurement transplant coordinators in deciding when to discuss organ donation with ICU patients' families. Figure 3 illustrates the resultant decision tree, from which a set of if-then classification criteria, namely, a decision factor priority list of agreeing with organ donation, is generated as follows:

(a) If psychological factors are satisfied = Y, physiological factors are satisfied = Y, and external factors are satisfied = Y, then organ donation = agree.

(b) If psychological factors are satisfied = Y, physiological factors are satisfied = N, and spiritual factors are satisfied = Y, then organ donation = agree.

(c) If psychological factors are satisfied = N and external factors are satisfied = Y, then organ donation = agree.

6. Discussion and Conclusion

This study investigated factors affecting the attitude of families of a member in an ICU regarding organ donation. The results indicated that organ donation attitude was correlated with age. Educational attainment, whether the patient had signed an organ donor card, and attitude toward organ donation were significantly correlated with families' consent to organ donation.

This study primarily applied the ANP to identify the factors affecting the decisions of ICU patients' families regarding organ donation. This study exhibited some limitations in relation to the selection of methodology and data collection. First of all, the ANP-based questionnaire used in this study was markedly different from conventional questionnaires and required step-by-step explanation, which considerably prolonged the time required to obtain responses. Consequently, the answers may not have accurately reflected the participants' consideration in decision-making, thus compromising the results. Furthermore, because the study was cross-sectional and predictive, the results may be inapplicable to other organ donation settings. Another concern is that the participants were recruited at a single hospital, rendering the results potentially unsuitable for extrapolation to other populations. Moreover, the most obvious limitation of the research is the subjectivity due to the use of surveys. All the opinions about organ donation and relative importance are obtained from the questionnaire, and the research results are greatly dependent on the relative knowledge and experience of the patients' families. Although a scientific approach was used, only using pairwise comparison of factors compiled in advance could not thoroughly reflect the feelings of family members. Nevertheless, we contend that the key features of this study were the proposed MCDM model and preliminary understanding of the attitude of ICU patients' families. Future research should employ a qualitative method to explore the views of patients' families on the meaning and value of organ donation when their relative's organs are removed for transplantation. Through a more in-depth analysis, additional factors that directly and indirectly influence patients' families in their organ donation decision-making should be identified. In clinical settings, we advise that medical institutions should host workshops related to organ donation and transplantation more often. Such workshops would enable personnel from various backgrounds to share their experiences, facilitate discussion, and enable members of an organ procurement team to better understand each other's roles, optimize use of team resources, build a consensus, increase the likelihood of successful organ donation, and save more patients requiring organ transplant.

Data Availability

The data used to support the finding of this study are included within the article.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research was supported in part by the Ministry of Science and Technology of Taiwan under grant numbers MOST 109-2637-H-242 -001 and MOST 110-2410-H-242 -001 -MY2.

References

[1] A. M. Rosenblum, L. D. Horvat, L. A. Siminoff, V. Prakash, J. Beitel, and A. X. Garg, "The authority of next-of-kin in explicit and presumed consent systems for deceased organ donation: an analysis of 54 nations," *Nephrology Dialysis Transplantation*, vol. 27, no. 6, pp. 2533–2546, 2012.

[2] D. Li, "Effect of persuasive messages on organ donation decisions: an experimental test," *Journal of Economic Behavior & Organization*, vol. 131, pp. 150–159, 2016.

[3] J. R. Rodrigue, D. L. Cornell, and R. J. Howard, "Organ donation decision: comparison of donor and nondonor families," *American Journal of Transplantation*, vol. 6, no. 1, pp. 190–198, 2006.

[4] M. Toews and T. Caulfield, "Evaluating the "family veto" of consent for organ donation," *Canadian Medical Association Journal*, vol. 188, no. 17-18, pp. E436–E437, 2016.

[5] N. Kentish-Barnes, L. A. Siminoff, W. Walker et al., "A narrative review of family members' experience of organ donation request after brain death in the critical care setting," *Intensive Care Medicine*, vol. 45, no. 3, pp. 331–342, 2019.

[6] L. R. Sophie, J. C. Salloway, G. Sorock, P. Volek, and F. K. Merkel, "Intensive care nurses' perceptions of cadaver organ procurement," *Heart & Lung*, vol. 12, no. 3, pp. 261–267, 1993.

[7] S. A. Bidigare and M. H. Oermann, "Attitudes and knowledge of nurses regarding organ procurement," *Heart & Lung: The Journal of Critical Care*, vol. 20, no. 1, pp. 20–24, 1991.

[8] L. A. Siminoff, N. Gordon, J. Hewlett, and R. Arnold, "Factors influencing families' consent for donation of solid organs for transplantation," *JAMA*, vol. 286, no. 1, pp. 71–77, 2001.

[9] L. Siminoff, M. B. Mercer, G. Graham, and C. Burant, "The reasons families donate organs for transplantation: implications for policy and practice," *The Journal of Trauma, Injury, Infection, and Critical Care*, vol. 62, no. 4, pp. 969–978, 2007.

[10] W. Walker, A. Broderick, and M. Sque, "Factors influencing bereaved families' decisions about organ donation," *Western Journal of Nursing Research*, vol. 35, no. 10, pp. 1339–1359, 2013.

[11] C. Araujo and M. Siqueira, "Brazilian healthcare professionals: a study of attitudes toward organ donation," *Transplantation Proceedings*, vol. 48, no. 10, pp. 3241–3244, 2016.

[12] F. Goz, M. Goz, and M. Erkan, "Knowledge and attitudes of medical, nursing, dentistry and health technician students towards organ donation: a pilot study," *Journal of Clinical Nursing*, vol. 15, no. 11, pp. 1371–1375, 2006.

[13] J. Cohen, S. B. Ami, T. Ashkenazi, and P. Singer, "Attitude of health care professionals to brain death: influence on the organ donation process," *Clinical Transplantation*, vol. 22, no. 2, pp. 211–215, 2008.

[14] I. Milaniak, E. Wilczek-Rużyczka, K. Wierzbicki, J. Sadowski, B. Kapelak, and P. Przybyłowski, "Evaluation of the

motivation to consent to and to refuse organ donation among participants of educational meetings concerning organ transplantation," *Transplantation Proceedings*, vol. 48, no. 5, pp. 1332–1336, 2016.

[15] H. Jung, "Reluctance to donate organs: a survey among medical students," *Transplantation Proceedings*, vol. 45, no. 4, pp. 1303-1304, 2013.

[16] M. Pouraghaei, M. Tagizadieh, A. Tagizadieh, P. Moharamzadeh, S. Esfahanian, and K. Shahsavari Nia, "Knowledge and attitude regarding organ donation among relatives of patients referred to the emergency department," *Emergency*, vol. 3, no. 1, pp. 33–39, 2015.

[17] H. S. Akgün, N. Bilgin, I. Tokalak, A. Kut, and M. Haberal, "Organ donation: a cross-sectional survey of the knowledge and personal views of Turkish health care professionals," *Transplantation Proceedings*, vol. 35, no. 4, pp. 1273–1275, 2003.

[18] L. A. Siminoff, H. M. Traino, and N. H. Gordon, "An exploratory study of relational, persuasive, and nonverbal communication in requests for tissue donation," *Journal of Health Communication*, vol. 16, no. 9, pp. 955–975, 2011.

[19] R. Afshar, S. Sanavi, and M. R. Rajabi, "Attitude and willingness of high school students toward organ donation," *Saudi journal of kidney diseases and transplantation: An Official Publication of the Saudi Center for Organ Transplantation, Saudi Arabia*, vol. 23, no. 5, pp. 929–933, 2012.

[20] M. L. Stoeckle, "Attitudes of critical care nurses toward organ donation," *Dimensions of Critical Care Nursing*, vol. 9, no. 6, pp. 354–361, 1993.

[21] C. E. Hsieh, Y. L. Chen, H. C. Lin, Y. Y. Chang, and C. S. Hsieh, "A study of orgon donation and procurement," *Formasan Journal of Medicine*, vol. 14, no. 1, pp. 26–31, 2010.

[22] S. N. Davison and G. S. Jhangri, "Knowledge and attitudes of Canadian first nations people toward organ donation and transplantation: a quantitative and qualitative analysis," *American Journal of Kidney Diseases*, vol. 64, no. 5, pp. 781–789, 2014.

[23] A. Holman, A. Karner-Huţuleac, and B. Ioan, "Factors of the willingness to consent to the donation of a deceased family member's organs among the Romanian urban population," *Transplantation Proceedings*, vol. 45, no. 9, pp. 3178–3182, 2013.

[24] E. Witkamp, M. Droger, R. Janssens, L. van Zuylen, and A. van der Heide, "How to deal with relatives of patients dying in the hospital? Qualitative content analysis of relatives' experiences," *Journal of Pain and Symptom Management*, vol. 52, no. 2, pp. 235–242, 2016.

[25] A. L. Simpkin, L. C. Robertson, V. S. Barber, and J. D. Young, "Modifiable factors influencing relatives' decision to offer organ donation: systematic review," *BMJ*, vol. 338, p. b991, 2009.

[26] A. M. S. Wu, "Discussion of posthumous organ donation in Chinese families," *Psychology Health & Medicine*, vol. 13, no. 1, pp. 48–54, E, 2008.

[27] S. Gortmaker, C. Beasley, E. Sheehy et al., "Improving the request process to increase family consent for organ donation," *Journal of Transplant Coordination*, vol. 8, no. 4, pp. 210–217, 1998.

[28] A. L. Simpkin, L. C. Robertson, V. S. Barber, and J. D. Young, "Modifiable factors influencing relatives' decision to offer organ donation: systematic review," *BMJ British Medical Journal*, vol. 338, 2009.

[29] P. Niles and B. Mattice, "The timing factor in the consent process," *Journal of Transplant Coordination*, vol. 6, no. 2, pp. 84–87, 1996.

[30] T. L. Saaty, *The Analytic Hierarchy Process*, McGraw-Hill, New York, NY, USA, 1980.

[31] T. L. Saaty, *Decision Making with Dependence and Feedback: The Analytic Network Process: The Organization and Prioritization of Complexity*, Rws Publications, Pittsburgh, PA, Pennsylvania, 2001.

[32] F. Faul, E. Erdfelder, A.-G. Lang, and A. Buchner, "*G**power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences," *Behavior Research Methods*, vol. 39, no. 2, pp. 175–191, 2007.

[33] S. T. Yeh and C. Y. Chen, "A study of people on the wishes of the family's organ donation and related factors in keelung city," *Journal of Health Promotion and Health Education*, vol. 41, no. 1, pp. 25–46, 2014.

[34] C. F. Waltz, O. L. Strickland, and E. R. Lenz, *Measurement in Nursing Research*, Springer, Manhattan, NY, USA, 2016.

[35] J. R. Quinlan, *C4.5: Programs for Machine Learning*, Elsevier Science, Amsterdam, Netherlands, 1993.

[36] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and Regression Trees*, Wadsworth International Group, Belmont, California, 1984.

[37] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

Hindawi
Journal of Healthcare Engineering
Volume 2021, Article ID 1233483, 12 pages
https://doi.org/10.1155/2021/1233483

Research Article

Multitask Healthcare Management Recommendation System Leveraging Knowledge Graph

Wanheng Liu,[1] Ling Yin,[2] Cong Wang,[1] Fulin Liu,[3] and Zhiyu Ni[3]

[1]*Beijing University of Posts and Telecommunications, Beijing, China*
[2]*National Population Health Data Center, Changping, China*
[3]*Affliated Hospital of Hebei University, Baoding, Hebei, China*

Correspondence should be addressed to Ling Yin; lwhlcf@163.com

Received 20 May 2021; Accepted 11 September 2021; Published 5 November 2021

Academic Editor: Chi-Hua Chen

In this paper, a novel multitask healthcare management recommendation system leveraging the knowledge graph is proposed, which is based on deep neural network and 5G network, and it can be applied in mobile and terminal device to free up medical resources and provide treatment programs. The technique we applied is referred to as KG-based recommendation system. When several experiments have been carried out, it is demonstrated that it is more intelligent and precise in disease prediction and treatment recommendation, similar to the state of the art. Also, it works well in the accuracy and comprehension, which is much higher and highly consistent with the predictions of the theoretical model. The fact that our work involves studies of multitask healthcare management recommendation system, which can contribute to the smart healthcare development, proves to be promising and encouraging.

1. Introduction

Recently, intelligent healthcare recommendation system has become a hot topic in healthcare management application research. As much attention has been paid to health, people have been gradually changing from the passive response to medical treatment to the active normal prevention and performing healthcare. Driven by the need for healthcare management application, an accurate and efficient healthcare recommendation system which can be applied in terminal device is now playing an important role in healthcare, which not only can make a more comprehensive and continuous record and analysis of our health condition but also can recommend appropriate health interventions and treatment programs. It can contribute to freeing up more medical resources while improving efficiency and accuracy.

It is known that the traditional recommendation system is to solve the problem of information explosion and there are two major tasks in it, including rating prediction and CTR prediction. Previously, most of the studies on recommendation system focused on collaborative filtering [1], content-based filtering [2], and hybrid method [3]. Collaborative filtering algorithm starts from similarity measurement to consider the similarity between users or items to make recommendations; and content-based recommendation methods can learn the representation of user and item form the contents of item. However, hybrid method is designed to address the problem of cold start and sparsity in the interaction matrix data. It can combine content information of the user and the item for integration into the collaborative filtering framework to obtain better performance.

While traditional recommendation system has made great progress, the accuracy and the explainability of the system are still the problem that needs to be solved. To fill the gaps, KG-based recommendation systems were carried out, in which the knowledge graph served as auxiliary information to be integrated into the recommendation system to improve the precision, diversity, and explainablity in it. The approaches of KG-aware recommendation system can be divided into three general classes: embedding-based method [4–7], path-based method [8, 9], and the unified method

[10, 11]. The embedding-based method is to learn a low-dimensional feature for each entity and each relation in knowledge graph. In literature, Zhang et al. [4] in 2016 proposed using the heterogeneous information in knowledge to promote the quality of the recommendation system. In 2017, Wang et al. [5] put forward signing heterogeneous information network embedding for sentiment link prediction. Then, in 2018, they [6] presented a deep knowledge-aware network to integrate knowledge graph representation into news recommendation. Wang et al. [7] in 2019 proposed a multitask feature learning approach for knowledge graph enhanced recommendation. Besides, the basic idea of the path-based method is to design the connected similarity or the entity semantic similarity between user and item to improve the recommendations. In 2014, Xiao et al. [8] introduced metapath-based latent features to represent the connectivity between users and items. Hu et al. [9] came up with leveraging metapath-based context for top-N recommendation in 2018. Furthermore, the unified method is to combine the embedding method with the path-based method to fully excavate the information in both aspects. Wang et al. [10] in 2018 proposed RippleNetwork to incorporate the knowledge graph into recommendation system based on the unified method. In 2019, they raised Knowledge Graph Convolutional Networks (KGCN) [11], which is an end-to-end framework that captures interitem relatedness effectively by mining their associated attributes on the KG.

In addition, from the perspective of model structure, there are three forms in the combination of knowledge graph and recommendation system, sequential training [6], joint training [10], and alternate training [7]. Sequential training refers to the fact that the entity and relation vectors of the knowledge graph are obtained by embedding firstly. Then the recommendation system is introduced to learn the user vector and item vector for model training. In a word, it is conducted sequentially in the training of the embedding of the knowledge graph and the recommendation system. And joint training means that knowledge graph embedding and model training are simultaneous. What is more, in alternate training, it was trained alternately in the task of the feature learning in the knowledge graph and the click-through rate prediction of the model.

As healthcare management in demand in our daily life and the recommendation system develops [12–14], Zaman et al. [12] proposed a personalized healthcare recommendation system using semantic web technology and healthcare social network in 2014. Also, Ahire et al. [14] in 2015 utilized the ontology based framework for healthcare recommendation system. Paramonov et al. [13] in 2016 also presented a recommendation service for smart space-based personalized healthcare system. Furthermore, it has become a tendency in healthcare management recommendation system which can be applied in mobile or terminal device to perform smart medical services. What is more, in literature, Archenaa et al. [15] presented health recommender system based on big data analytics in 2017. Guzmán et al. [16] in 2018 proposed a collaborative framework for sensing abnormal heart rate based on a semantic recommender system for healthcare. Meanwhile, Kaur et al. [17] proposed an efficient multiparty scheme for privacy preserving collaborative filtering for healthcare recommender system. Ali et al. [18] in 2018 presented a type-2 fuzzy ontology-aided recommendation system for IoT-based healthcare. Simultaneously, Somarathna et al. [19] proposed a recommendation system for customer preferred mental healthcare facility. Then Sahoo et al. [20] used collaborative filtering to perform deep learning based health recommender system in 2019. At the same time, Roy et al. [21] demonstrated integrating wearable devices and recommendation system toward a next-generation healthcare service delivery. Meanwhile, Rathi et al. [22] proposed a mobile based healthcare tool, an integrated disease prediction and recommendation system. In 2020, Noshad et al. [23] presented the clinical recommender system, which can predict medical specialty diagnostic choices with neural network ensembles. Simultaneously, Hussein et al. [24] proposed an accurate and reliable recommender system for chronic disease diagnosis. Meanwhile, Nagaraj et al. [25] designed a framework for e-healthcare management service based on recommender system. Then, in 2021, Ochoa et al. [26] raised a medical recommender system based on continuous-valued logic and multicriteria decision operators with interpretable neural networks. At the same time, Saad et al. [27] proposed a situation-aware recommendation system for personalized healthcare applications. Also, Pitchai et al. [28] presented a generic medicine recommendation system for advanced e-healthcare based on cloud computing. Meanwhile, Ponselvakumar et al. [29] proposed combining the recommendation system with deep learning to achieve the precision quality of healthcare. Moreover, to improve the performance of the recommendation system based on knowledge graph, in 2021, Andrea et al. [30] presented a comprehensive comparison of knowledge graph embedding-based link prediction methods. Meanwhile, Qian et al. [31] utilized a knowledge-aware multimodel adaptive graph learning principal for the effective feature learning. At the same year, Saikat et al. [32] proposed relation prediction of comorbid disease using knowledge graph completion of a tensor. In 2020, Huang et al. [33] provided a knowledge-driven multimodel activity recognition framework that exploits external knowledge to fuse multimodal data and reduce the dependence on large-scale training samples. In 2019, Wang et al. [34] proposed RippleNet, an end-to-end framework that naturally incorporates the KG into recommender systems.

Great progress has been made in the performance of the recommendation system which can be applied in healthcare management, can improve the accuracy and the efficiency of the home visits and in treatment, and can also alleviate some medical resources. However, there are still some limitations in it: (1) Since the structural knowledge of the healthcare is the major information adopted in the knowledge base, it can be not comprehensive, precise, and efficient. (2) Due to single traditional recommendation system applied in healthcare management, it can lead to data sparsity, cold start problem, and overfitting. (3) Due to lack of abundant medical and healthcare data, the performance of the

recommendation system is poor. (4) The deep neural network is not incorporated into most healthcare recommendation systems, which cannot lead to intelligence and efficiency to free up more medical resources and alleviate the pressure in treatment. (5) Owing to single disease paid attention to in the recommendation system, other diseases' prediction, diagnosis, and treatment recommendation would be ignored, which is incomprehensive.

In order to overcome the limitations and the problems above, a novel multitask healthcare management recommendation system leveraging the knowledge graph and deep neural network based on 5G network applied in mobile and terminal device is presented in this paper. To make it more comprehensive and accurate, except for the structural knowledge, the textual knowledge and the visual knowledge are also contained in the knowledge base. Then the knowledge graph is utilized in our healthcare recommendation to alleviate the problems of the single systems and be more precise and explainable. What is more, to acquire better performance, 600 thousand pieces of new data can be provided to support our system. In addition, deep neural network is also adopted in our system to make it more intelligent in various diseases' prediction, diagnosis, and treatment recommendation. Finally, it can be applied in mobile and terminal device, which can serve the patients and the doctors. The contributions in our paper are as follows:

(i) To be more comprehensive in our healthcare recommendation system, structural knowledge, textual knowledge, and visual knowledge are involved in the knowledge base.

(ii) The healthcare and medical knowledge graph is leveraged in our system, which can be more precise and accurate.

(iii) A better model can be obtained due to new labelled and processed dataset.

(iv) Deep neural network is utilized in our system to contribute to intelligent various diseases' prediction, diagnosis, and treatment recommendation provided to serve patients and doctors.

The remainder of this paper is organized as follows: Section 2 introduces the materials and methods which the multimodel healthcare management recommendation system and network architecture and framework contained. In Section 3, the results and discussion are expressed. The paper is concluded in Section 4.

2. Materials and Methods

The architecture of our multitask healthcare management recommendation system leveraging the knowledge graph and deep neural network based on 5G network is shown in Figure 1. It can be applied in mobile and terminal device, which is connected to various places, such as hospitals, communities, and homes to integrate users' health information data from multiple channels. The collaborative knowledge graph is input and a multimode knowledge graph entity encoder is utilized in the knowledge graph embedding module. Then the new entity representation is used to learn knowledge graph embedding in order to represent the knowledge inference relationship. In the recommendation module, the embedding was learned by the knowledge graph, also with the collaborative knowledge graph, to enrich the expression of patients and medical items to improve the recommendation effect and the experience. Taking viral pneumonia as an example, when the user passes their data through our system, it will automatically rank the severity of the disease and advise the user on what grade of hospital they should go to and provide other treatment recommendations and so on.

2.1. Multitask Healthcare Management Recommendation System Framework. The framework of our multitask healthcare recommendation is shown in Figure 2, which consists of recommendation task module, feature learning task of the healthcare knowledge graph module, and cross compression unit. In the recommendation part, the feature representation of the user and the item is regarded as input, while the predicted click probability is regarded as output. The features of the user and item are extracted, respectively, by the multilayer perceptron and cross-compression unit and then they are fed into another multilayer perceptron. In the knowledge graph part, the head node of the triple and the relationship representation are considered as input, while the predicted trail node is regarded as output. When a couple of the head and relation is input, the features of the head and relation are also extracted, respectively, by multilayer perceptron and cross-compression unit. Then the representation of predictive tail can be calculated on the basis of the head and the relation; also the similarity between the predicted tail and the actual tail can also be calculated by function f. In other words, it is the knowledge graph embedding capability score of the link prediction. However, in practical application, the recommendation task and the feature learning of the knowledge graph are not independent of each other, on account of the overlap between the items in the recommendation system and the entities in the knowledge graph. The cross feature-sharing unit is designed as a connected band between the two tasks, which is the key to connect the recommendation module and the knowledge graph embedding module and can automatically learn the high-level interaction features of recommendation item and knowledge graph entity.

The information can be exchanged in the cross-compression unit in Figure 3. It is seen that the item vector and the entity vector are the two descriptions of the same object actually and the cross-sharing of the information between them allows each to get additional information from each other to make up for the lack of information sparsity. However, it is noted that the cross-compression unit can exist at the lower level of the system. It is known that, along the network, the features can be translated from generic to specific, while the feature transferability is decreased significantly at higher levels with the increased variability in tasks, which can lead to a negative shift in the high-level sharing, especially the heterogeneous task. Besides, in the high layer, the mixed features of item, user, entity, and relation are not suitable for sharing due to none being explicitly related.

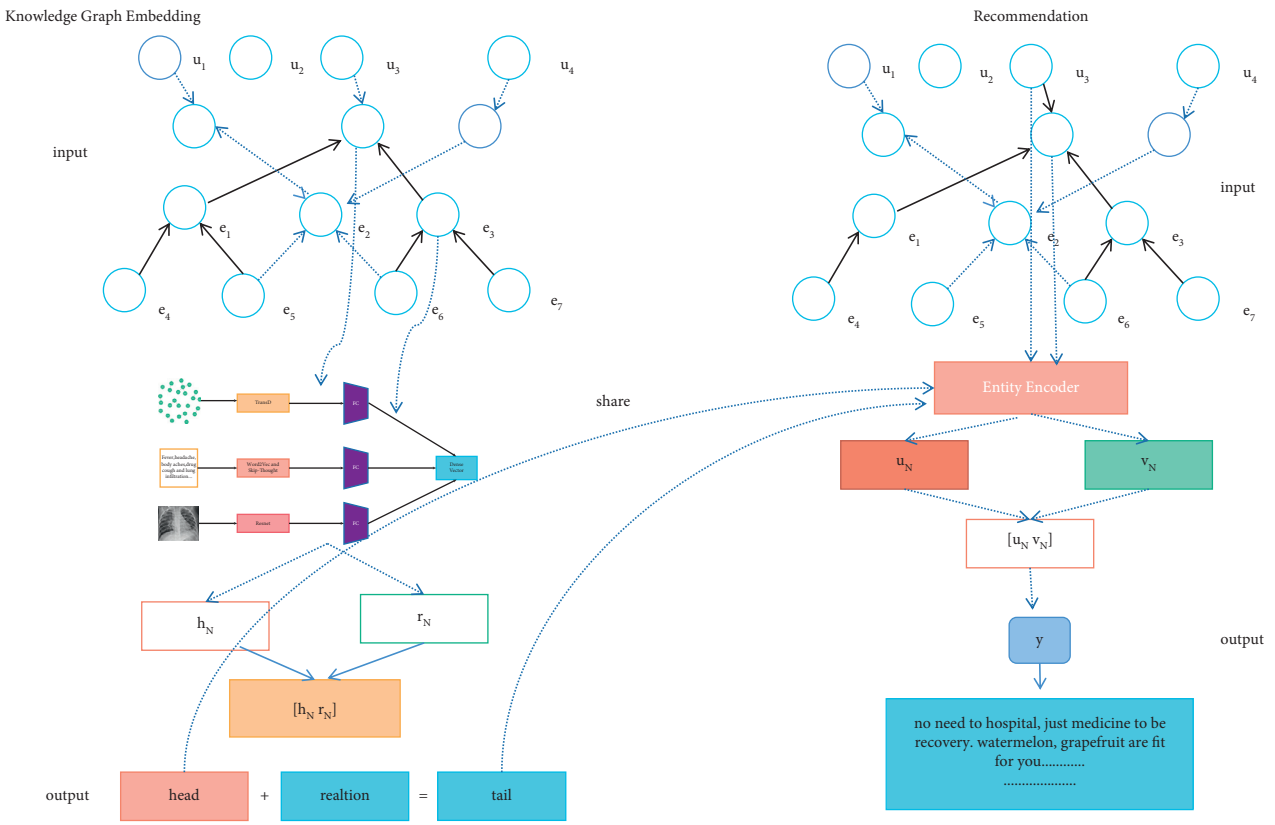


Figure 1: The architecture of our multitask healthcare management intelligent recommendation system.

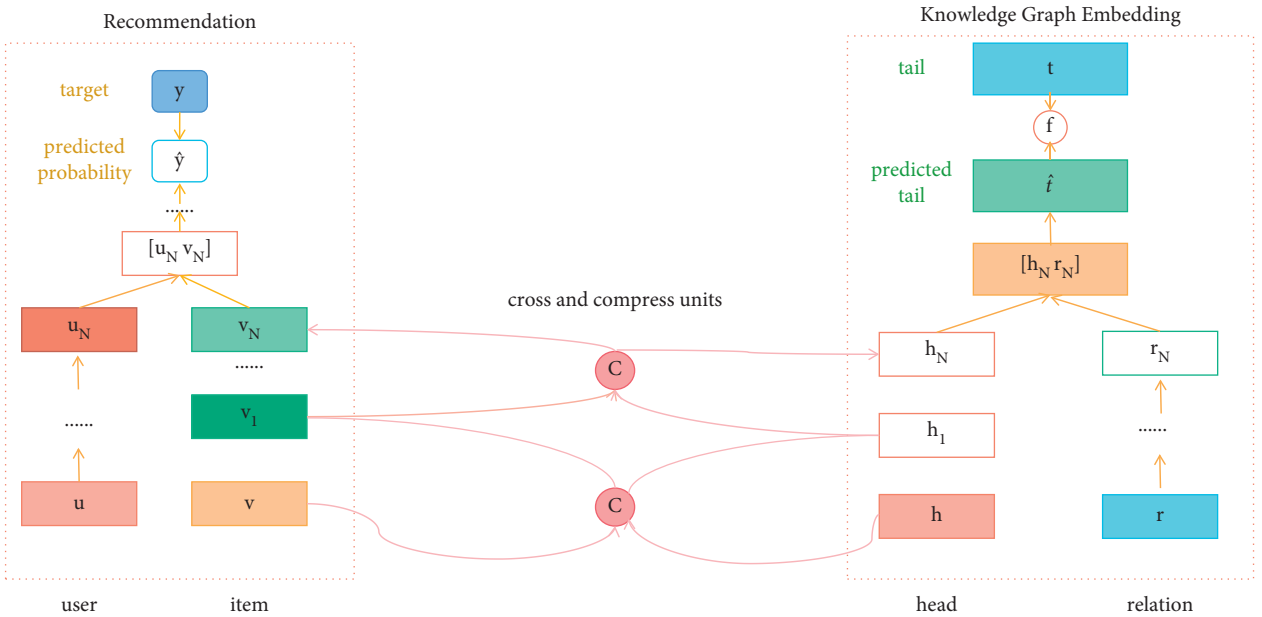


Figure 2: The framework of multitask healthcare management recommendation system.

2.2. Structural Knowledge. As shown in Figure 4, our healthcare knowledge graph is constructed of various entities and links, which can be considered as a heterogeneous network. Our healthcare knowledge graph model is based on 600 thousand pieces of training data, which has a good performance on department classification, food and drug recommendation, treatment recommendation, and major item examination according to the disease. While this representation method works well for the structured data, it is difficult to process the healthcare knowledge graph.

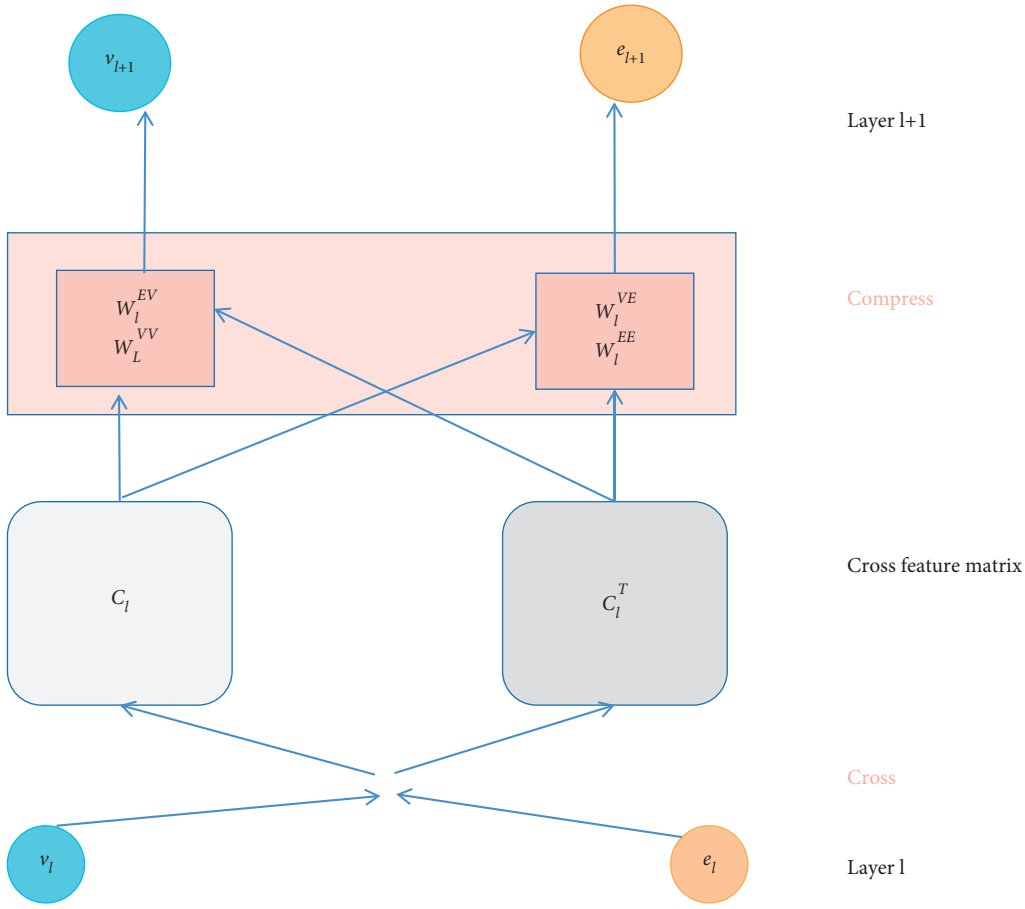


Figure 3: The cross-compression unit.

Therefore, the knowledge graph embedding approach is utilized to address the above problem, especially the embedded components, which not only can transform the relationships between entities into a continuous vector space to simply the operation but also can retrain the original structure of the knowledge graph. In our system, the translation distance model is adopted as the knowledge embedding method, in which the distance-based scoring function is utilized and the reasonableness of a fact is measured by the distance between two entities. In general, there are several types of translation distance models including TransE, TransH, TransD, and TransR. It is known that an entity is a complex of properties and different attributes of the entity are focused on by different relationships. In our work, TransD shown in Figure 5, which is the improvement of TransR, is used and it is considered that different entities should be mapped into different semantic spaces. Also, it overcomes the shortcomings of computational complexity and model parameter. Owing to fewer parameters and no matrix vector multiplication in TransD, it can be suitable for our large-scale healthcare knowledge graph.

Two vectors were used to represent each entity and relationship in TransD, and the two mapping matrices are defined as formulas (1) and (2), where the first vector represents the meaning of the entity or relationship and the second vector would be utilized to construct the mapping matrix. It is found that the mapping matrix is defined by entities and relations and I represents the identity matrix. In formulas (3) and (4), it is seen that $h_\perp$ and $t_\perp$ are projected vectors of entities. The loss function of knowledge representation and training method is shown in formula (6) and the function of f_r is defined in formula (5).

$$M_{rh} = r_p h_p^T + I, \tag{1}$$

$$M_{rt} = r_p t_p^T + I, \tag{2}$$

$$h_\perp = M_{rh} h, \tag{3}$$

$$t_\perp = M_{rt} t, \tag{4}$$

$$f_r(h,t) = \left\| h_\perp + r - t_\perp \right\|_2^2, \tag{5}$$

$$L = \sum_{(h,r,t)\in S} \sum_{(h',r,t')\in S'} \max\left(0, f_r(h,t) + \gamma - f_r(h',t')\right). \tag{6}$$

2.3. Textual Knowledge. In our healthcare management system, a large amount of text data on symptoms would be recorded and updated every day. Therefore, the textual

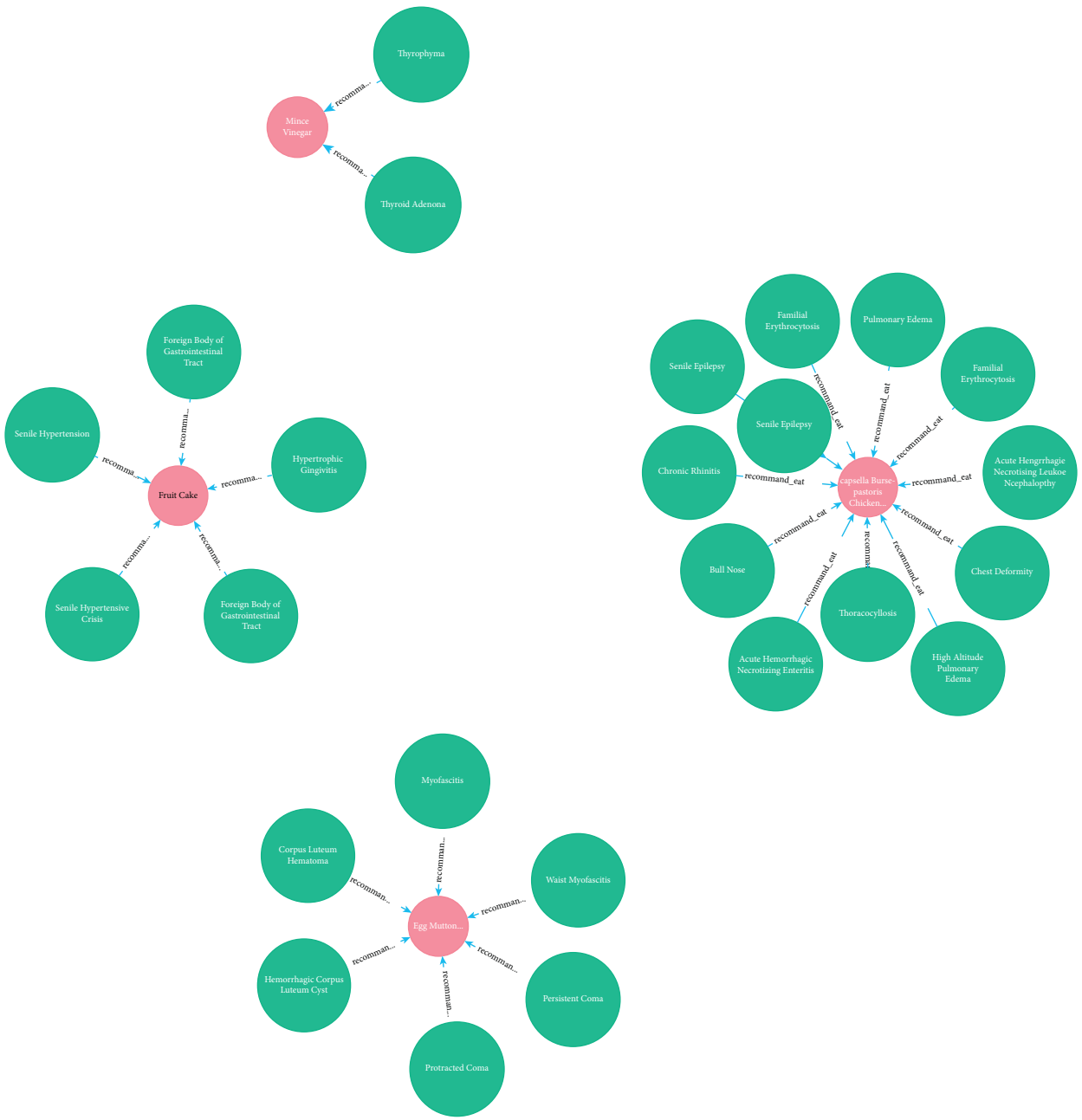


Figure 4: Healthcare knowledge graph.

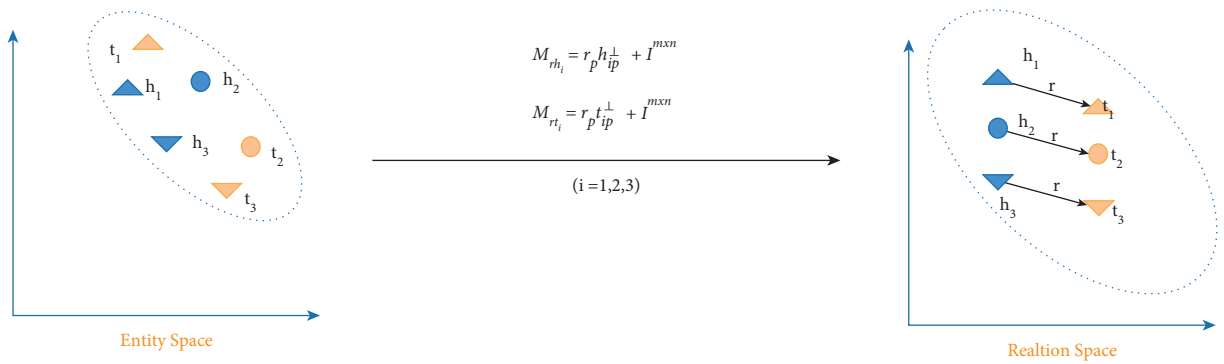


Figure 5: The illustration of TransD.

knowledge can contribute to the accuracy and efficiency of the analysis of the symptoms and text representation is of great significance, which can convert the symbols of human language into numbers that the machines can compute. In our work, word embedding method of Word2Vec is adopted, which is one of the most commonly used word embedding models and neural network models. Besides, the semantic information of the word can be represented in the form of word vector, and semantically similar words are close together in the space through an embedded space. The word embedding method utilized is shown in Figure 6 and it is similar to the idea of autoencoder. It is constructed based on the training data and then the model learned parameters can be obtained. It is shown that $w(t)$ is the input word and in the hidden layer it can perform the dot product between matrix and the input vector $w(t)$ and then the outcome is passed to the output layer. In the output layer, it can calculate the dot product between the output vector of the hidden layer and the weight matrix of the output layer. Then the activation function of softmax is used to calculate the probability that a word appears in the context of $w(t)$ at a given context location. Finally, Skip-Thought method is utilized as a sentence vector to acquire the textual features representation, which is the encoder-decoder architecture and the GRU model. There are several advantages in Word2Vec and Skip-Thought adopted in our system. Firstly, it can be suitable for any medical and healthcare original text to complete modeling of the relationships between sentences, since it is an unsupervised learning technique. Secondly, it takes less memory which can be applied in mobile terminal device. Finally, it has fewer dimensions in the weight matrix, which leads to fewer computation, which can be fit for healthcare management application.

The probability is shown in formula (7), where $w(c, j)$ represents the predicted j word at the c context location, $w(O, c)$ represents the actual word that appears at the c context position, $w(I)$ represents the only input word, and $u(c, j)$ represents the j value of the U vector when a word is predicted at the c context position. The loss function is expressed in formula (8), where the probability can be maximized when predicting $w(c, j)$ at the c context position.

$$p\left(w_{c,j} = w_{O,c} \mid w_I\right) = \frac{\exp u_{c,j}}{\sum_{j-1}^{V} \exp u_j}, \tag{7}$$

$$\begin{aligned} L &= -\log P\left(w_{c,1}, w_{c,2}, \ldots, w_{c,C} \mid w_o\right) \\ &= -\sum_{c=1}^{C} u_{c,j^*} + \sum_{c=1}^{C} \log \sum_{j=1}^{V} \exp\left(u_{c,j}\right). \end{aligned} \tag{8}$$

2.4. Visual Knowledge. It is demonstrated that visual knowledge also plays an important role in our healthcare management recommendation system; and there are many diseases that need medical images to perform further diagnosis, for example, lung disease, liver disease, and femoral head disease. In our healthcare management system, a large number of medical images also would be uploaded and updated. In order to obtain the semantic information and the embedded representation of the image, the backbone network ResNet would be adopted in our system. The structure of the residual learning block is shown in Figure 7. It is obvious that ResNet is a residual network which can be stacked to form a deep network, which contains a natural identity mapping that can solve the problem of network degradation to some extent. What is more, the mapping with residuals is more sensitive to changes in output. Also, in forward propagation, the input signal can propagate directly from any low level to the high level and error signals can be propagated directly to the lower level without any intermediate weight matrix transformation. Therefore, to some degree, the problem of gradient dispersion can be alleviated, which allows information to propagate back and forth more smoothly in residual connection. The residual unit is expressed in formula (9), where x_l and x_{l+1} represent the input and the output of the residual unit of l, respectively. In addition, F is the residual function, which represents the learned residual and h (x_l) represents the identity mapping. In formula (10), f is the ReLU activation function. The learning feature from shallow layer l to deep layer L is expressed in formula (11). The gradient of the reversed process can be obtained by the chain rule in formula (12), where 1 shows that the short-cut mechanism can propagate the gradient without loss and in the other item the residual gradient passes through a layer with weights.

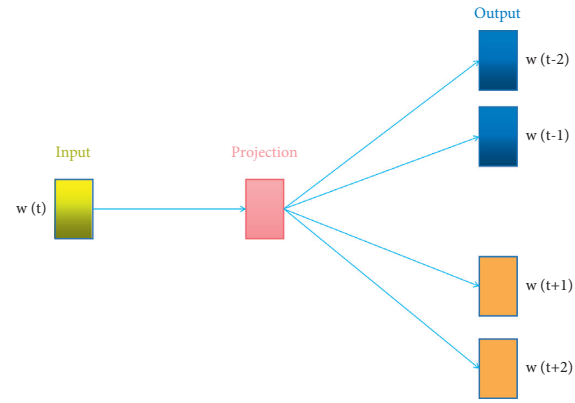


Figure 6: The architecture of Skip-Gram.

$$y_l = h(x_l) + F(x_l, W_l), \tag{9}$$

$$x_{l+1} = f(y_l), \tag{10}$$

$$x_L = x_l \sum_{i=1}^{L-1} F(x_i, W_I), \tag{11}$$

$$\frac{\partial \text{loss}}{\partial x_L} = \frac{\partial \text{loss}}{\partial x_l} \cdot \frac{\partial \text{loss}}{\partial x_L} = \frac{\partial \text{loss}}{\partial x_L} \cdot \left(1 + \frac{\partial}{\partial x_L} \sum_{i=1}^{L-1} F(x_i, W_i)\right). \tag{12}$$

The architecture of the backbone network ResNet is shown in Figure 8, in which each residual block can

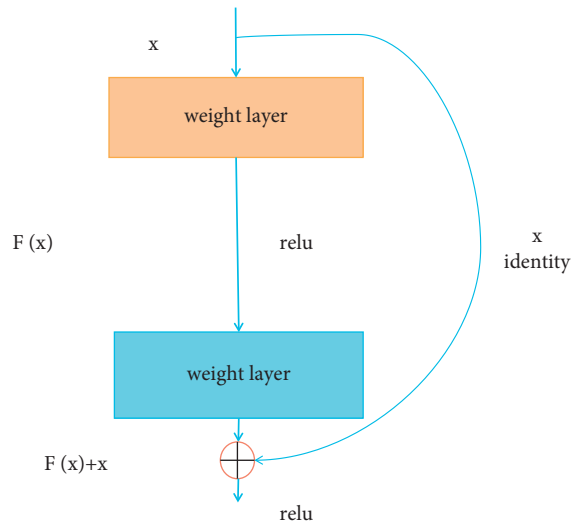


Figure 7: The structure of residual learning block.

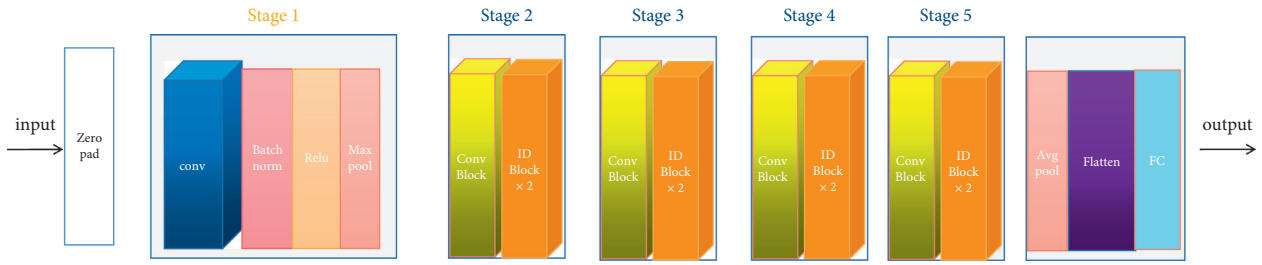


Figure 8: The architecture of ResNet.

constitute a residual network and convolution layer, Batchnorm, ReLU, Maxpool, Convblock, IDblock, and so on involved to acquire the medical images representation. The ID block (Identity block) is shown in Figure 9 and the convolutional block is shown in Figure 10.

2.5. Healthcare Management Knowledge Graph Encoder. Our multitask healthcare knowledge graph encoder is shown in Figure 11. The healthcare structural knowledge was embedded by translation distance model TransD, the healthcare textual knowledge was embedded by word embedding method Word2Vec and a sentence vector Skip-Thought method to acquire the textual features representation, and the visual knowledge was embedded by ResNet backbone network to obtain the semantic information and the embedded representation of the image. They were regarded as tail to be input, and then they passed through the dense layer to make the features extracted above nonlinearly changed to acquire the correlation between these features and then mapped to the output space to obtain the dense vector.

The loss function of our healthcare management system is shown in formula (13), where the first term is the cross-entropy loss of the recommendation module and u and v are the the users and items collections which have been traversed. What is more, the second term is the loss of knowledge graph embedding module, the aim of which is to increase the score of the correct triple and reduce the score of the error triples. In addition, the third item is regular item to prevent overfitting, where λ_1 and λ_2 are the equilibrium constants.

$$\begin{aligned} L &= L_{\mathrm{RS}} + L_{\mathrm{KG}} + L_{\mathrm{REG}} \\ &= \sum_{u\in U, v\in V} \mathcal{J}\left(\hat{y}_{uv}, y_{uv}\right) - \lambda_1 \left(\sum_{(h,r,t)\in G} \mathrm{score}\left(h,r,t\right) - \sum_{\left(h',r,t'\notin G\right)} \mathrm{score}\left(h',r,t'\right) + \lambda_2 \|W\|_2^2 \right. \end{aligned} \tag{13}$$

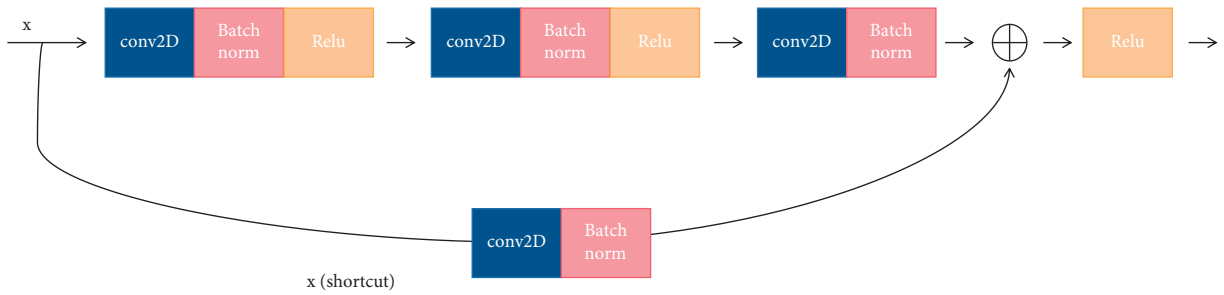


Figure 9: The architecture of ID block.

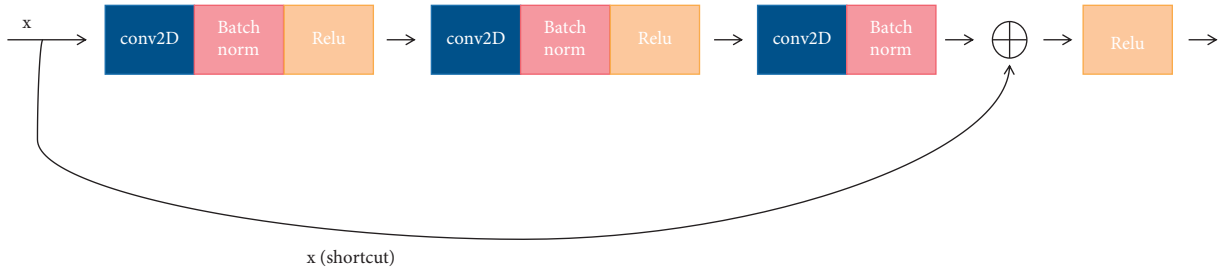


Figure 10: The architecture of convolutional block.

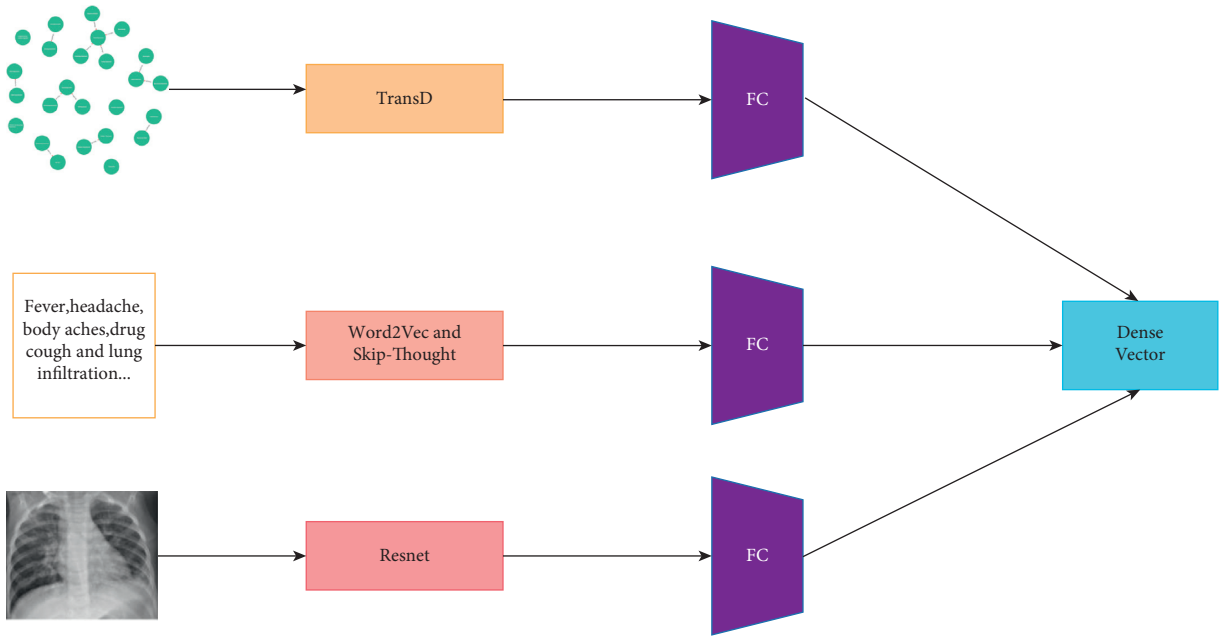


Figure 11: Multitask healthcare knowledge graph encoder.

3. Results and Discussion

Table 1 shows the recall (state of the art) in different models for our dataset. It is obvious that our multitask healthcare management recommendation system has better performance compared to other common recommendation systems. What is more, compared with base models, our system still performs better than text-based and image-based models in Table 2. The bar chart is shown in Figures 12 and 13. It is seen that our multitask healthcare management recommendation system is encouraging, which is more convenient to serve the patients and doctors to perform disease prediction and provide treatment recommendation. In addition, it can free up and alleviate medical resources to make it balanced in medical field. However, there exist some limitations in this system to recommend the most suitable doctor corresponding to the disease due to lack of the doctors data, which needs to be solved in the future.

Table 1: The comparison of recall (state of the art) in different models for our dataset.

Model	Recall
NFM	0.2595
CKE	0.2719
RippleNet	0.2791
MKR	0.2812
NGCF	0.2819
KGAT	0.2827
MMGCN	0.2979
Ours	0.3126

Table 2: The comparison of recall (state of the art) in our models and base models for our dataset.

Model	Recall
Base	0.0712
Base + image	0.0755
Base + text	0.0729
Ours	0.0771

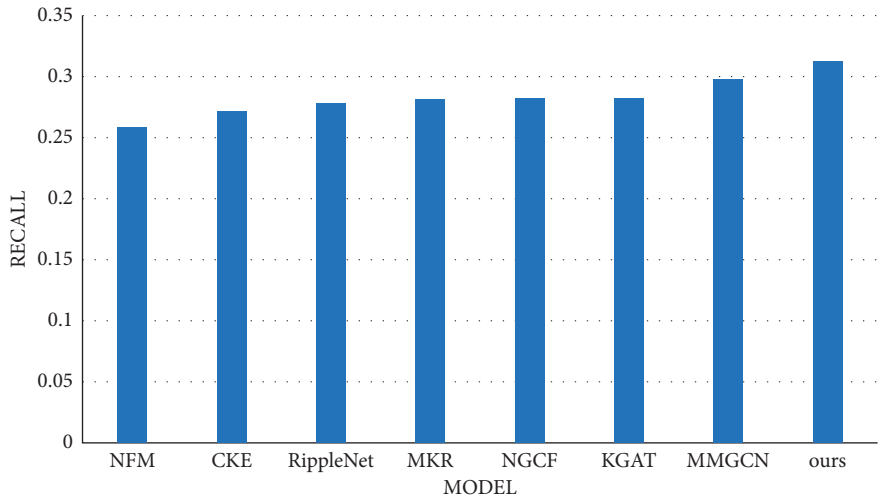


Figure 12: The comparison of recall in different models for our dataset.

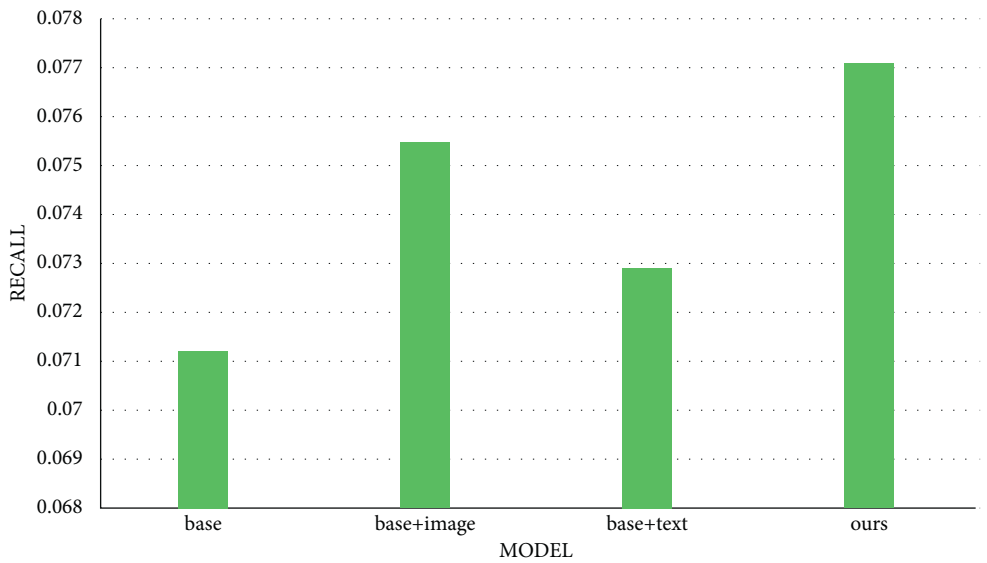


Figure 13: The comparison of recall in our models and base models for our dataset.

4. Conclusions

In conclusion, it is shown that our multitask healthcare management recommendation system is intelligent and promising, which can free up and alleviate the medical resources to be more convenient and efficient. In addition, it has advantages compared to other common methods in propagating and generalizing the data to obtain the high-value and useful information for the patients and the doctors. However, due to lack of the doctors data, our system is not able to recommend the appropriate doctor corresponding to the disease. Besides, there are still some limitations in comprehensive professional guidance in our recommendation system. In the future, abundant doctors data should be added and processed, accompanied by more professional guidance to support out healthcare management recommendation system to make it more comprehensive.

Data Availability

The data used are available and can be accessed to perform multitask healthcare management recommendation system based on knowledge graph. Part of the data are available from the corresponding author upon request (liuwanheng301@163.com).

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the study of this work and publication of this paper.

Acknowledgments

This work was supported by Key R&D Program of Hebei Province, People's Livelihood Science and Technology Project, Project no. 20277716D (Project name: Study on the pathological characteristics and pathogenic mechanism of Covid-19).

References

[1] B. Sarwar, "Item-based collaborative filtering recommendation algorithms," in *Proceedings of the the 10th International World Wide Web Conference (WWW10)*, Hong Kong, China, May 2001.

[2] P. Lops, M. D. Gemmis, and G. Semeraro, "Content-based recommender systems: state of the art and trends," *Recommender Systems Handbook*, Springer, Boston, MA, USA, 2011.

[3] R. Burke, "Hybrid recommender systems: survey and experiments," *User Modeling and User-Adapted Interaction*, vol. 12, no. 4, pp. 331–370, 2002.

[4] F. Zhang, N. J. Yuan, and D. Lian, "Collaborative knowledge base embedding for recommender systems," in *Proceedings of the the 22nd ACM SIGKDD International Conference*, ACM, New York, NY, USA, August 2016.

[5] H. Wang, F. Zhang, M. Hou, X. Xie, M. Gao, and Q. Li, "Shine: signed heterogeneous information network embedding for sentiment link prediction," in *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*, CA, USA, February 2017.

[6] H. Wang, F. Zhang, and X. Xing, "DKN: deep knowledge-aware network for news recommendation," in *Proceedings of the The 2018 World Wide Web Conference*, Lyon, France, April 2018.

[7] H. Wang, F. Zhang, and M. Zhao, "Multi-task feature learning for knowledge graph enhanced recommendation," in *Proceedings of the The 2019 World Wide Web Conference*, San Francisco, CA, USA, May 2019.

[8] Y. Xiao, R. Xiang, and Y. Sun, "Personalized entity recommendation: a heterogeneous information network approach," in *Proceedings of the 7th ACM International Conference on Web Search and Data Mining*, ACM, New York, NY, USA, February 2014.

[9] B. Hu, C. Shi, W. X. Zhao, and S. Y. Plillips, "Leveraging meta-path based context for top- n recommendation with a neural co-attention model," in *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1531–1540, London, UK, August 2018.

[10] H. Wang, F. Zhang, and J. Wang, "RippleNet: propagating user preferences on the knowledge graph for recommender systems," in *Proceedings of the 27th ACM International Conference*, Turin, Italy, October 2018.

[11] H. Wang, M. Zhao, X. Xie, M. Gao, and W. Li, "Knowledge graph convolutional networks for recommender systems," in *Proceedings of the The 2018 World Wide Web Conference*, San Francisco, CA, USA, May 2019.

[12] N. Zaman and J. Li, "Semantics-enhanced recommendation system for social healthcare," in *Proceedings of the 2014 IEEE 28th International Conference on Advanced Information Networking and Applications*, pp. 765–770, IEEE, Victoria, BC, Canada, May 2014.

[13] I. Paramonov and A. Vasilyev, "Recommendation service for smart space-based personalized healthcare system, open innovations association," in *Proceedings of the 2016 19th Conference of Open Innovations Association (FRUCT)*, IEEE, Jyvaskyla, Finland, November 2016.

[14] S. B. Ahire and H. Khanuja, "HealthCare recommendation for personalized framework," *International Journal of Computer Application*, vol. 110, no. 1, pp. 24–26, 2015.

[15] J. Archenaa and E. Anita, "Health recommender system using big data analytics," *Journal of Management Science and Business Intelligence*, vol. 2, 2017.

[16] G. Guzmán, M. R. Torres, and V. Tambonero, "A collaborative framework for sensing abnormal heart rate based on a recommender system: semantic recommender system for healthcare," *Journal of Medical and Biological Engineering*, vol. 38, 2018.

[17] H. Kaur, N. Kumar, and S. Batra, "An efficient multi-party scheme for privacy preserving collaborative filtering for healthcare recommender system," *Future Generation Computer Systems*, vol. 86, pp. 297–307, 2018.

[18] F. Ali, S. M. R. Islam, D. Kwak, P. Khan, N. Ullah, and S. Yoo, "Type-2 fuzzy ontology-aided recommendation systems for IoT-based healthcare," *Computer Communications*, vol. 119, pp. 138–155, 2018.

[19] U. Somarathna, S. Walia, and L. Manchuri, "Recommendation system for customer preferred mental healthcare facility," in *Proceedings of the 2018 IISE Annual Conference*, Orlando, FL, USA, May 2018.

[20] A. K. Sahoo, C. Pradhan, R. K. Barik, and H. Dubey, "DeepReco: deep learning based health recommender system

using collaborative filtering," *Computation*, vol. 7, no. 2, p. 25, 2019.

[21] S. N. Roy, S. K. Srivastava, and R. Gururajan, "Integrating wearable devices and recommendation system: towards a next generation healthcare service delivery," *Journal of Information Technology Theory and Application*, vol. 19, no. 4, 2019.

[22] M. Rathi and V. Pareek, "Mobile based healthcare tool an integrated disease prediction & recommendation system," *International Journal of Knowledge and Systems Science*, vol. 10, no. 1, pp. 38–62, 2019.

[23] M. Noshad, I. Jankovic, and J. H. Chen, "Clinical recommender system: predicting medical specialty diagnostic choices with neural network ensembles," 2020, https://arxiv.org/abs/2007.12161.

[24] A. S. Hussein, W. M. Omar, and X. Li, "Accurate and reliable recommender system for chronic disease diagnosis," in *Proceedings of the First International Conference on Global Health*, pp. 113–118, Venice, Italy, 2012.

[25] P. Nagaraj and P. Deepalakshmi, "A framework for e-Health care management service using recommender system," *Electronic Government an International Journal*, vol. 16, no. 1, 2020.

[26] J. Ochoa, O. Csiszár, and T. Schimper, "Medical recommender systems based on continuous-valued logic and multi-criteria decision operators, using interpretable neural networks," *BMC Medical Informatics and Decision Making*, vol. 21, no. 186, 2021.

[27] A. Saad, H. Fouad, and A. A. Mohamed, "Situation-aware recommendation system for personalized healthcare applications," *Journal of Ambient Intelligence and Humanized Computing*, no. 2, pp. 1–15, 2021.

[28] R. Pitchai, S. Anjanayya, and M. Maravarman, "Cloud computing based generic medicine recommendation system for advanced E-Healthcare," *Materials Today: Proceedings*, no. 10, , 2021.

[29] A. P. Ponselvakumar, S. Anandamurugan, and K. S. Logeswaran, "Advancement in precision medicine and recommendation system for clinical trials using deep learning methods," *IOP Conference Series: Materials Science and Engineering*, vol. 1055, no. 1, Article ID 012110, 2021.

[30] A. Rossi, D. Barbosa, D. Firmani, A. Matinata, and P. Merialdo, "Knowledge graph embedding for link prediction: a comparative analysis," *ACM Transactions on Knowledge Discovery from Data*, vol. 15, no. 2, p. 49, 2021.

[31] S. Qian, J. Hu, F. Quan, and C. Xu, "Knowledge-aware multimodal adaptive graph convolutional networks for fake news detection," *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 17, no. 3, p. 23, 2021.

[32] S. Biswas, P. Mitra, and K. S. Rao, "Relation prediction of Comorbid diseases using knowledge graph completion," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 18, no. 2, pp. 708–717, 2021.

[33] Y. Huang, X. Yang, J. Gao, J. Sang, and C. Xu, "Knowledge-driven egocentric multimodal activity recognition," *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 16, no. 4, 2020.

[34] H. Wang, F. Zhang, J. Wang et al., "Exploring high-order user preference on the knowledge graph for recommender systems," *ACM Transactions on Information Systems*, vol. 37, no. 3, 2019.

Hindawi
Journal of Healthcare Engineering
Volume 2021, Article ID 2828141, 12 pages
https://doi.org/10.1155/2021/2828141

Research Article

Enhancing the Health and Well-Being of People with Chronic Diseases: Assessment and Sustainable Development Planning for Therapeutic Landscapes after Urban Expansion

Lin Mei,[1] Kun Liu,[2] and Bo-Wei Zhu[3]

[1]*School of Art and Design, Wuhan University of Technology, Wuhan, Hubei 430070, China*
[2]*School of Art and Design, Quanzhou Normal University, Feng ze, Quanzhou, Fujian 362000, China*
[3]*Faculty of Humanities and Arts, Macau University of Science and Technology, Macau 999078, China*

Correspondence should be addressed to Bo-Wei Zhu; bwzhu@must.edu.mo

Received 21 May 2021; Revised 21 August 2021; Accepted 21 September 2021; Published 8 October 2021

Academic Editor: Hao-Chun Lu

Under the influence of economic, environmental, and social structural changes, urban space expands and contracts to varying degrees and the everyday urban landscape changes in response. Over the past 20 years, a large number of cities in China have undergone a brief but rapid urban expansion and are moving toward shrinking cities. Most of these cities are now facing social problems such as an aging population and a high prevalence of chronic diseases. Therefore, the "therapeutic" role and impact of everyday landscapes in these cities need to be examined in the context of urban development processes through appropriate assessment methods. Therefore, this study applies the ANP-mV model to examine the therapeutic nature of everyday urban landscapes in different development periods, with the aim of enhancing the health and well-being of people with chronic diseases. Firstly, this study uses the city of Jinzhou in Northeast China as an example to develop a framework for assessing the therapeutic nature of everyday urban landscapes based on the health care needs of people with chronic diseases; secondly, it examines the therapeutic nature of the former Jinzhou Suburban Riverfront Forest Park as it has developed and evolved over the past 16 years; finally, it explores place-making and regeneration strategies for therapeutic landscapes from the perspectives of dynamic impact and sustainable development to enhance chronic illness patients' well-being. At the theoretical level, this study contributes by providing a methodology and research ideas for examining the "therapeutic" nature of everyday urban landscapes and proposing further development plans for renewal, constructing a framework for assessing therapeutic landscapes, and elucidating the relationship between networks of influence and the relative importance of various assessment dimensions/elements. At the practical application level, the contribution of this study is to provide local policymakers with a key decision basis for the future development planning of the East Lake Forest Park. The aim is to explore landscape creation and regeneration strategies for the East Lake Forest Park in the context of Jinzhou's progressive move toward a shrinking city, in order to sustain the well-being of the chronically ill.

1. Introduction

Previous research has repeatedly documented the restorative effects of place on health and well-being [1, 2] and has developed a number of framework concepts such as the biophilia hypothesis, the Attention Restoration Theory, and the concept of healing landscapes [3]. Under the concept of therapeutic landscapes, scholars have examined many types of landscapes associated with healing or rehabilitation, including natural landscapes such as villages, mountains, and lakes that have a reputation for healing [4, 5]. There are also everyday landscapes, such as places and residential areas where medical services are provided, and libraries [6, 7]. Then, there are urban landscapes, such as urban public green spaces and streets, and social networks [8, 9]. In humanist and cultural ecology theories, the formation of therapeutic landscapes is dynamic and the landscape can be seen as an evolving process where "therapeutic" is examined in the

context of changing environmental, social, and economic conditions [10]. Yan and He [8] argue that it is important to explore the evolution of therapeutic landscapes, i.e., how therapeutic landscapes change over time.

Realistic experience shows that the development of local productivity, changes in social structure, and the introduction of macrogovernance policies can have a vital impact on the evolution of the urban landscape. When the landscape style changes, it will inevitably lead to discussions about endowing the landscape with healing properties or maintaining the healing properties of the place. Examples include exploring and understanding the creation of restorative and therapeutic spatial places for refugees or dislocated farmers in urban distribution and resettlement planning [9, 11, 12], exploring longevity villages as tourist destinations as the health tourism industry grows [8], and exploring design strategies for healing and therapeutic gardens in senior communities or medical buildings as we move toward a healthy aging society [13]. It is worth noting that as urban economies develop, populations grow, and the quality of life improves, cities at all levels in many regions experience varying degrees of expansion, and the landscape of former suburban areas will evolve significantly under the influence of multiple factors. For local urban dwellers, villages, green spaces, forests, and parks in the suburbs are often seen as therapeutic and healing landscapes that combine physical and nonphysical levels [14, 15].

Due to the economic downturn of cities, industrial transformation, and an aging population, the rate of urban expansion into suburban areas has generally slowed down in recent years, and more emphasis has been placed on the planning concept of "transformation" rather than "new construction/rebuilding," with more emphasis on the inner development of cities, microrenewal, and adaptive improvement of urban space. Even in China, where the built-up area has grown exponentially over the past 30 years, spatial expansion is no longer the dominant form of urban development, and many cities are facing a shift from "incremental planning" to "stock planning" [16]. In China, in particular, a large number of small and medium-sized cities are now moving toward shrinking cities after urban expansion, with serious urban population loss and an increasing trend toward aging [17].

Concerned with the development of these cities, scholars and practitioners have proposed a range of urban regeneration and development strategies that can be summarized in three main development planning directions: regrowth, urban islands, and dedensification and greening [18]. Regardless of the development strategy, what needs to be acknowledged is the irreversibility of urban shrinkage, and that urban growth and decline, like life cycles, are seen as natural processes of urban change, requiring planning managers to shift a commonly accepted perception [19]. Scholars have suggested that local policymakers should examine whether sprawling urban landscapes meet the real needs of the current population and improve urban green space networks to increase the livability and attractiveness of cities, thereby mitigating population loss and enhancing economic vitality [20–22].

In contrast to China's Tier 1-2 cities, there is a large demand for healthcare in small and medium-sized cities that have undergone a brief and rapid urban expansion and are now gradually shrinking. Perhaps due to its unique natural environment and dietary habits, Northeast China has long been a region with a high prevalence of chronic diseases such as cardiovascular disease, diabetes, and gout [23, 24]. In many cities in Northeast China, there is widespread and strong awareness of the need for chronic disease patients to go outdoors for physical, mental, and spiritual healing [25]. Through ongoing observations and interviews over many years, this study considers that people with chronic illnesses in the region value rely more on outdoor blue, green, and white spaces for healing and therapy in appropriate seasonal and climatic conditions. In the cities of Jinzhou in Liaoning Province and Jiamusi in Heilongjiang Province, for example, outside of the extreme weather days in the northeast, large numbers of people regularly enter urban squares, parks, and waterfront streets where people can gather and relax on an almost daily basis to enjoy the healing and therapeutic effects of the landscape through a variety of health behaviors and activities. Therefore, as cities shift from sprawl to contraction and respond to the current needs of the population to continue to enhance the health and well-being of people with chronic diseases, the former suburban landscape needs to be reexamined not only for its therapeutic qualities but also for new place-making strategies in urban regeneration. However, much of the previous research has focused on examining the positive effects of landscape environments on patients [26, 27] and explaining the formation and evolution of therapeutic landscapes. Few scholars have integrated urban development and planning concepts in a public health context, examining the therapeutic nature of an evolving landscape and exploring how to shape therapeutic landscapes for sustainable well-being in urban regeneration.

In summary, this study uses the city of Jinzhou in Liaoning Province, China, as a case study. The city has experienced a brief and rapid urban expansion and is now shifting to a shrinking city with significant population loss and an aging trend. The city has a high prevalence of chronic diseases in the northeast and is generally representative of healthcare resources and the development of the built-up areas. The purpose of this study is to develop a framework for assessing the therapeutic nature of everyday urban landscapes based on the health care needs of people with chronic diseases, examine the therapeutic nature of the former Jinzhou Suburban Riverfront Forest Park as it has developed and evolved over the past 16 years, and explore strategies for place-making and regeneration of therapeutic landscapes for the well-being of people with chronic diseases from a dynamic impact and continuous development perspective. The design of this study is shown in Figure 1. Firstly, this study uses a literature review to initially extract the elements of therapeutic assessment for everyday urban landscapes and then, through focus group interviews, constructs a framework for assessing the therapeutic landscape of the East Lake Forest Park in Linghe District, Jinzhou City, based on the health care needs of people with chronic illnesses. Secondly, a network analysis method (ANP) was

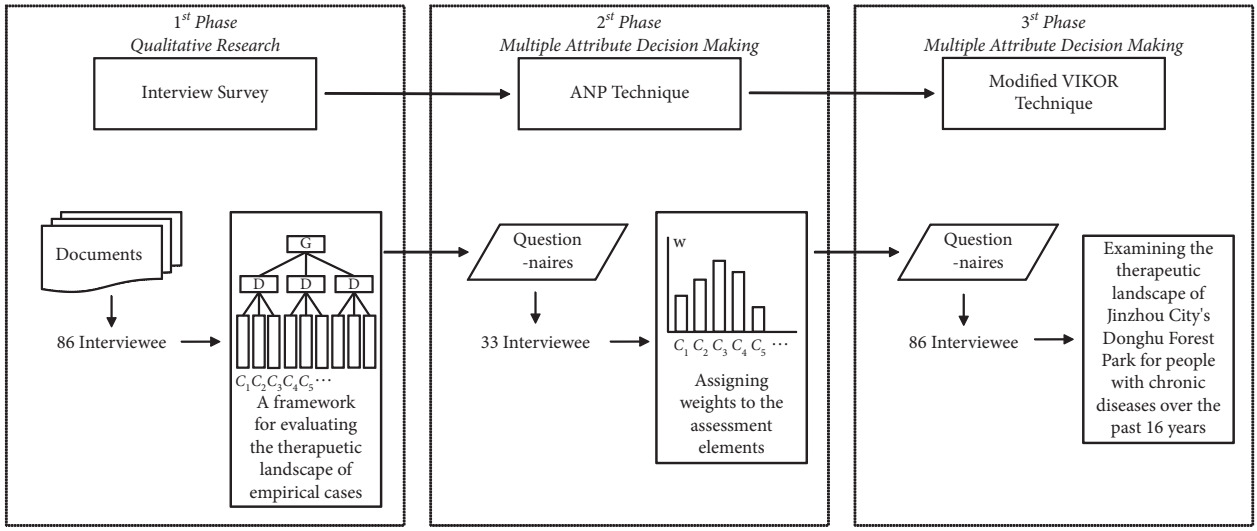


Figure 1: The process and design of this study.

applied to assign weights to the assessment elements based on expert opinion, and a modified VIKOR technique was applied to examine the therapeutic landscape of Jinzhou City's Donghu Forest Park for people with chronic diseases over the past 16 years, in conjunction with interviews and questionnaires administered by the public. Finally, based on the results of the assessment and analysis, strategies for the creation and renewal of the therapeutic landscape of the East Lake Forest Park are explored in the context of Jinzhou's progressive move toward a shrinking city, in order to sustain the well-being of people with chronic diseases.

The remaining sections of this study are organized as follows. A review of the relevant literature is presented in Section 2. The research methodology and steps are clarified in Section 3. The findings and discussion are presented in Section 4, and the study is summarized in Section 5.

2. Literature Review: Healing and Therapeutic Landscape Design

The term "therapeutic landscape" is often used to describe this type of place space and has been defined by scholars as "a place that sustains physical and psychological recovery and comfort" [10]. Conradson [4] states that therapeutic landscapes involve a complex reciprocal relationship between a person and the broader social context of his or her environment. From the 1920s onwards, with the rapid development of modern medical technology in the form of drugs and medical diagnostic equipment, which revolutionized the way humans were treated, the potential of the landscape environment to regulate and restore disease has rarely been mentioned or studied. It was not until the late 1980s that scholars argued that medicine should be based on harmony between people and their environment, and numerous studies have since confirmed that natural landscapes can promote human health in both physical and psychological terms, with the therapeutic effects being more spiritual in nature. Duff [28] explicitly recommends that those suffering from mental illness go to places with proven therapeutic properties for ongoing healing. A "therapeutic" landscape is not just one type of space; that is, a landscape environment that can have a therapeutic effect includes many different types of spaces, such as waterfront spaces [7, 29], villages [4], horticulture and gardens [30], natural forests [5], and courtyards [31]. Due to the more complex components of the landscape environment and the service functions it can provide, its therapeutic function is not only spiritual. Moreover, the healing of landscape environments is not only for people with illnesses but also for people from all walks of life and all age groups [26]. Finlay et al. [14] divide the therapeutic landscape into two components: blue space and green space, and discuss the different effects they have on older people in their old age. Blue space is particularly important for mental recovery and building mental health, while green space can have a positive impact on older people's social activities. As such, research on the therapeutic nature of such spaces is often focused on spatial environments such as natural landscapes, landscape gardens, residential courtyards, and built environments [32–34].

Gesler [35] describes the sense of place in therapeutic landscapes and the four main dimensions: the natural environment, the built environment, the symbolic environment, and the social environment. In the context of related theoretical perspectives (e.g., Aesthetic Affective Theory (AAT) and spiritual evolutionary theory), scholars have summarized three main characteristics of therapeutic landscapes and therapeutic gardens: relief of physical symptoms, illness or trauma, reduction of stress for individuals dealing with emotionally and/or physically stressful experiences, and improvement of overall well-being [36, 37]. Therapeutic landscapes within healthcare buildings include places where horticultural therapeutic activities exist such as therapeutic gardens, healing gardens, meditation gardens, and memorial gardens [38]. However, this relational perspective on therapeutic landscapes makes it difficult to know how to shape places to enhance well-being, and how place experiences may be healing or therapeutic has been undertheorized [39–41]. The research carried out by

Doughty [42] and Pitt [41] highlights that exploring the role of the moving body in interacting with the natural environment provides valuable insight into how places act therapeutically, and they both argue for a more dynamic understanding of Butterfield and Martin [43], which also discussed how the sensory richness of place provides opportunities for therapeutic effects to emerge when studying the environment of cancer support centers. Gorman [44] provided more insight into exactly how these sensory experiences become therapeutic by focusing on one particular sense. He explored the role of smell within Community Supported Agricultural projects. Belčáková et al. [13] proposed a strategy for greening hospital facilities and external spaces and adding elements that highlight and support the therapeutic effects of the spatial environment, creating specific therapeutic landscape design forms that create the required physical environmental base for patients.

3. Research Methodology and Steps

This study uses a multicriteria decision-making approach (MCDM) to structure the research design. Through a review of the relevant literature, an initial assessment framework for assessing the therapeutic nature of everyday urban landscapes was extracted. A focus group approach was then used to collect qualitative data and to identify the assessment framework (dimensions/elements) in this study through qualitative inductive analysis based on the well-being and health of people with chronic diseases and to clarify the interactions between the factors in the structure of the assessment framework. The ANP technique was then used to assign weights to the assessment factors through the administration of a questionnaire. In combination with the people's interviews and questionnaires, the modified VIKOR technique was applied to examine the therapeutic landscape of Jinzhou East Lake Forest Park for people with chronic illnesses over the last 16 years of its development and evolution. In summary, the results of each phase of the study were integrated to explore the evolution of the therapeutic landscape in Jinzhou East Lake Forest Park and to explore strategies for the creation and renewal of the therapeutic landscape in East Lake Forest Park in order to continue to promote the well-being of patients with chronic illnesses.

3.1. Focus Group Interview Method and Data Collection. Focus group interviewing originated in 1941 at Columbia University's Division of Radio Studies, when Paul Lazarsfeld invited Robert Merton to help assess listeners' reactions to radio programs. The method is designed to gather information of interest to the researcher and tends to uncover a wide range of opinions from a variety of types of people. It provides a more natural context than individual interviews, where the researcher is the presenter, the listener, the observer, and finally the analyst who processes the data using an inductive approach. It is in the focus group that the researcher gains understanding based on the content of the discussion, rather than aiming to test preconceived theories or hypotheses. Importantly, the focus group approach allows the researcher to observe the interaction of the participants and to capture the substance of the views, opinions, experience, and attitudes expressed in the spoken word, as well as the design of the discussion questions to assess the issues related to the research topic.

This study considers the behavioral and perceptual impact of respondents' own chronic disease categories, as well as the slow walking distance from the forest park due to the development and expansion of the urban built environment, on their behavior and perception in the therapeutic landscape. Before sampling, the chronic diseases that are most prevalent in the area were classified into three categories: cardiovascular, metabolic, and respiratory diseases [45]. The three types of sampling areas were divided into a radius of 5-, 15-, and 30-minute walking distance from each park entrance. Respondents were recruited anonymously in these three types of areas, and those who entered the focus group interviews were required to be chronically ill, have outdoor health habits, and enter East Lake Forest Park at least three days a week on average to receive healing and therapeutic influences from the daily landscape. Respondents were recruited and placed into focus groups according to their subdistrict and chronic disease category. A total of 9 groups of 8 to 12 people each were interviewed, making a total of 86 people. After all the qualitative data had been collated and analyzed, it was found that sufficient information had been obtained on the research themes, so the number of focus groups was not increased (information on the respondents is shown in Table 1).

3.2. ANP Technique and Data Collection. Given that the elements used to examine the everyday therapeutic nature of urban landscapes have different attribute categories, the analytical approach is based on the Analytic Network Process (ANP) used in multicriteria decision analysis [46]. Compared to older generations of multicriteria decision analysis (MCDA) methods (e.g., AHP), the ANP method is more accurate because it has the ability to examine relationships between variables. In addition to its multicriteria analysis capabilities, the ANP technique does not require independence among the criteria and factors; thus, it can be used as an efficient method for cases wherein the criteria are interdependent and their factors affect the final decision-making goal. The procedure for the ANP technique consists of a series of following steps:

Step (1): model (network) construction.

Step (2): design questionnaire and survey.

Step (3): (1) construct pairwise comparison weight; (2) calculate criteria weight, using (1) and (2); (3) do testing of consistency; the value of consensus can be estimated by (3).

Table 1: Basic information about participants in the focus group interviews.

Respondent source	Chronic disease category			Age (year)		
	Cardiovascular	Metabolic	Respiratory	20–35	35–50	50–65
A. Residents within a 5-minute slow walking distance	8	10	10	7	15	6
B. Residents within a 15-minute slow walking distance	8	12	8	5	14	9
C. Residents within a 30-minute slow walking distance	12	10	8	4	11	15
Total (people)	28	32	26	16	40	30

$$\mathbf{A} = \left[a_{ij}\right]_{n\times n} = \begin{array}{c} \\ C_1 \\ \vdots \\ C_i \\ \vdots \\ C_n \end{array} \begin{array}{c} \begin{array}{cccc} C_1 & \cdots & C_j & \cdots C_n \end{array} \\ \begin{bmatrix} \frac{w_1}{w_1} & \cdots & \frac{w_1}{w_j} & \cdots & \frac{w_1}{w_n} \\ \vdots & \vdots & \vdots & & \\ \frac{w_i}{w_1} & \cdots & \frac{w_i}{w_j} & \cdots & \frac{w_i}{w_n} \\ \vdots & \vdots & \vdots & & \\ \frac{w_n}{w_1} & \cdots & \frac{w_n}{w_j} & \cdots & \frac{w_n}{w_n} \end{bmatrix} \end{array}, \tag{1}$$

$$r_i = \left(\prod_{j=1}^{n} a_{ij}\right)^{1/n}, \qquad w_i = \frac{r_i}{\sum_{i=1}^{n} r_i}, \tag{2}$$

$$C.I. = \frac{(\lambda_{\max} - n)}{(n-1)}, \qquad C.R. = \frac{C.I.}{R.I.}. \tag{3}$$

Step A4: construct an unweighted supermatrix.

Step A5: weighted supermatrix.

Step A6: limited supermatrix.

In this study, the ANP questionnaire was administered to people with chronic illnesses who had previously participated in focus group interviews and who had lived in Jinzhou for over16 years and entered East Lake Forest Park almost daily. A total of 33 respondents completed the questionnaire, and 28 valid questionnaires were returned.

3.3. Modified VIKOR Technique and Data Collection. VIKOR (VlseKriterijumska Optimizacija I Kompromisno Resenje) is a multicriterion decision-making method proposed by Professor Serafim Opricovic and Professor Gwo-Hshiung Tzeng in 1998 whereas VIKOR is one of the optimal compromises of multicriteria decision-making methods, whose basic idea is to first define the positive-ideal solution (optimal solution) and the negative-ideal solution (the worst solution).

Step V1: determine the best values and the worst values (very bad← 0, 1, 2, . . ., 10 ⟶ very good).

Step V2: compute the gap values of group utility and the gap values of individual regret.

Step V3: summarize the gap between group utility and individual regret. These values can be calculated using

$$s_k = \sum_{j=1}^{n} w_j r_{kj} = \sum_{j=1}^{n} w_j \frac{\left(\left|f_j^{\text{aspired}} - f_{kj}\right|\right)}{\left(\left|f_j^{\text{aspired}} - f_j^{\text{worst}}\right|\right)}. \tag{4}$$

Step V4: rank the alternatives.

A therapeutic landscape performance assessment questionnaire was administered to respondents who participated in focus group interviews in Jinzhou City's Donghu Forest Park. Respondents were asked to evaluate the therapeutic landscape performance of the park based on their knowledge and perceptions of the park and the development of the park over the past 16 years, as collected by this study. Finally, a total of 86 questionnaires were distributed, and 81 valid questionnaires were returned.

4. Results and Discussion

4.1. Constructing Therapeutic Landscape Assessment Frameworks and Influence Network Relationships. This study explores the role of sensory and embodied place experiences as an area of significant contribution, using the elements of the everyday urban landscape that can contribute to the well-being and therapeutic care of people with chronic illness as the basic questions for semistructured interviews in focus groups. Sensory and embodied experiences of place contribute to an understanding of how a place may act therapeutically [34, 41, 43]. Group discussions in the focus group interviews have been centered on the questions "how does East Lake Forest Park heal and treat you" and "what elements of the landscape are healing in." The interviews were transcribed into verbatim transcripts, and qualitative data was summarized and analyzed in Nvivo12 software. This study used the general qualitative inductive analysis method proposed by David R Thomas in 2006 to analyze the data. The analysis procedure of this method can be described as coding of qualitative information into a number of "concept nodes," which are "categorized" and formed organizationally. The expected outcome of this process is the creation of three to eight summary categories that capture the key aspects of the themes identified in the raw data. The inductive coding that identifies more than eight main categories is incomplete, in which case certain categories are combined or the assessor identifies which themes or categories are most important.

A memo has been written for each category, such as association, connection, and meaning. Each category was revised and improved through continuous comparative analysis by searching for new concepts and contradictions. After the results were organizationally formed to be sufficiently stable and saturated, the results of the inductive analysis are shown in Figure 2. Four main categories were identified, and the relationships of their interaction networks were clarified. Each category has a self-influencing relationship due to the interactions between the concept nodes under the category. Within this network of influence, there are three sets of two-way influence relationships within the main categories named "body practice and physical movement" and "access and bond to nature," "body practice and physical movement" and "sociality and symbolic environment," and "sociality and symbolic environment" and "sensory experience." Unidirectional influences include the following: both "body practice and physical movement" and "access and bond to nature" have an impact on "sensory experience." "Sociality and symbolic environment" receive the influence from "access and bond to nature." As shown in Figure 2, the 14 conceptual nodes fall under four main categories and can be considered as a hierarchical structure. In this study, the main categories obtained through the inductive analysis are considered as assessment dimensions, and the coded conceptual nodes are considered as assessment elements, thus building an assessment framework for the therapeutic landscape of Jinzhou East Lake Forest Park.

4.2. Weighting Analysis of Assessment Elements in the Healing Landscape. In this study, the ANP technique is applied to calculate the relative weights of each assessment element based on the examination of the influence network relationships between the assessment constructs. After passing the consistency check, the unweighted supermatrix constructed in this study is shown in Table 2. After integrating the empirical judgments of 28 respondents and the limited supermatrix, the weights of each assessment dimension/element are shown in Table 3. The results show that sociality and symbolic environment (D_3) and sensory experience (D_4) are the two constructs with the highest relative importance. The remaining levels of importance are, in order of importance, D_1 and D_2. Under the D_3 dimension, the assessment element with the highest local weight is promoting a wide range of social interactions (C_8). In this assessment framework, the relative weights of the assessment elements are closer to each other under D_1, the lowest relative weight.

This means that the outdoor social opportunities, symbolic environment, and range of sensory experiences that East Lake Forest Park can provide are of paramount importance in enhancing the well-being and health of local people with chronic illnesses. As suggested by scholars in previous studies, healing and therapeutic landscapes should be integrated with active living in the course of urban development, linking them to social interactions and together creating a supportive environment that is essential for healthy living [47, 48]. In addition, positive and influential social culture, people, or a spiritual intention related to the therapeutic nature of the landscape can also have an important healing and therapeutic effect on groups of people with chronic illnesses. In certain contexts, people with chronic illnesses may even emulate the behavior of the symbolic environment and actively restore their self-confidence [8]. Some of the sensory experiences that the East Lake Forest Park offers to the public that are different from everyday urban life are also important landscape elements for people with chronic illnesses, for example, fresh scents [44], culturally appropriate touch [49], and nature-derived sound and visual stimuli [50].

4.3. Landscape Evolutionary Development and Therapeutic Assessment in Jinzhou East Lake Forest Park. With the rapid urbanization of the whole of China, the built environment of Jinzhou City has been upgraded, and new residential developments have been built on both sides of the Xiaoling River, the main river flowing through the built-up area of the city. Over the past 20 years, a number of high-quality residential projects have been launched in the area, and local policymakers have set the long-term goal of creating a "livable and healthy" residential area. In this context, the slow walking trail system in the East Lake Forest Park and on both sides of the Xiaoling River has developed over the years into a major urban park landscape, enhancing the well-being of Jinzhou residents and providing daily outdoor fitness and social interaction for a wide range of people. According to the site survey, the East Lake Forest Park has gone through three key stages of development, from the initial construction in 2004 to the gradual shaping of the landscape in 2012 and then to the completion and refinement in 2020 (as shown in Table 4). Therefore, this study examines the development and evolution of the therapeutic landscape of the East Lake Forest Park in Jinzhou City by using these three development periods as time points and assesses the therapeutic nature of this everyday urban landscape with the goal of continuously enhancing the well-being and health care of people with chronic diseases.

In 2004, Jinzhou East Lake Park was only constructed with walking paths, the trees in the park were not yet cultivated or were still in the growing stage, and the infrastructure was lacking. In 2012, the third year prior to the completion of the construction work on the East Lake Forest Park, the vegetation coverage was much higher than when construction began in 2004, the roads were better, and the infrastructure was evenly distributed throughout the park. From the west side of the park to the middle of the park, three areas have been built 200 meters apart for the public to relax and enjoy. In the middle of the park, starting from the riverfront on the southern side of the site, three more viewing spots have been built along the riverbank to the east. By 2020, five years after the completion of East Lake Park, the car park and cultural square on the west side of the park had been built, the road leading from the entrance to the river had been widened, the size of each recreational space in the park had been increased, the central rest area had been transformed into a comprehensive sports ground with football, basketball, and badminton courts, and the park had

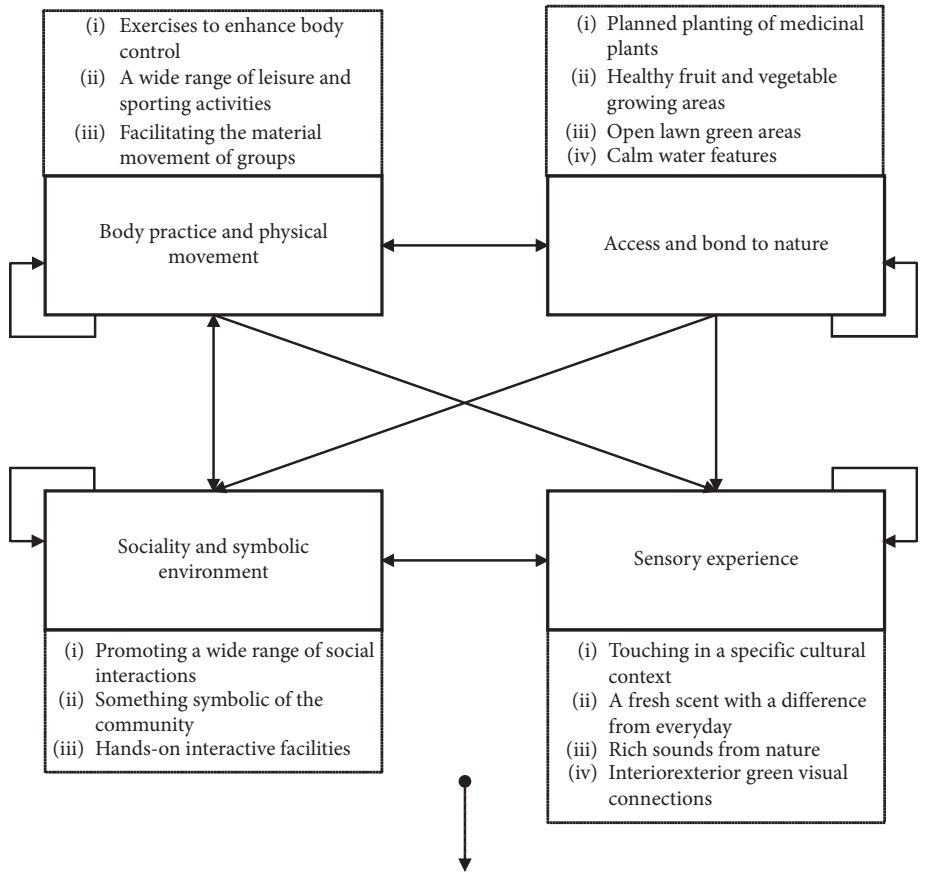


Figure 2: Assessment framework and influence network relationships.

Table 2: Construct unweighted supermatrix.

	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9	C_{10}	C_{11}	C_{12}	C_{13}	C_{14}
C_1	1.00	0.00	0.00	0.20	0.30	0.18	0.35	0.20	0.17	0.47	0.00	0.00	0.00	0.00
C_2	0.00	1.00	0.00	0.15	0.53	0.33	0.40	0.24	0.44	0.15	0.00	0.00	0.00	0.00
C_3	0.00	0.00	1.00	0.65	0.17	0.49	0.25	0.56	0.39	0.38	0.00	0.00	0.00	0.00
C_4	0.13	0.15	0.05	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
C_5	0.25	0.10	0.17	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
C_6	0.36	0.48	0.55	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
C_7	0.26	0.27	0.23	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
C_8	0.28	0.45	0.69	0.25	0.45	0.50	0.47	1.00	0.00	0.00	0.57	0.25	0.36	0.45
C_9	0.15	0.33	0.25	0.63	0.18	0.16	0.38	0.00	1.00	0.00	0.21	0.55	0.54	0.28
C_{10}	0.57	0.22	0.06	0.12	0.37	0.34	0.15	0.00	0.00	1.00	0.22	0.20	0.10	0.27
C_{11}	0.42	0.25	0.08	0.22	0.30	0.22	0.23	0.09	0.25	0.55	1.00	0.00	0.00	0.00
C_{12}	0.08	0.10	0.35	0.45	0.42	0.25	0.24	0.25	0.40	0.03	0.00	1.00	0.00	0.00
C_{13}	0.15	0.27	0.25	0.06	0.12	0.15	0.40	0.41	0.30	0.15	0.00	0.00	1.00	0.00
C_{14}	0.35	0.38	0.32	0.27	0.16	0.38	0.13	0.25	0.05	0.27	0.00	0.00	0.00	1.00
SUM	4.00	4.00	4.00	4.00	4.00	4.00	4.00	3.00	3.00	3.00	2.00	2.00	2.00	2.00

introduced the water of the Xiaoling River through the entire park, further enhancing its ornamental qualities. The park has been further enhanced by the introduction of the Xiaoling River. The east side of the park is less covered with vegetation than the west side, and the area between the roads is mostly grass. The distribution of trees in the park is more rational than in 2012, with different types of trees planted depending on their location, giving the park a more structured appearance overall.

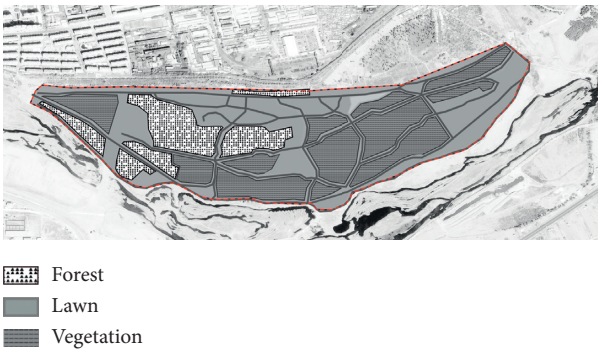
The results of the analysis of the therapeutic landscape assessment in different development and construction periods show that the healing and therapeutic effects of Jinzhou East Lake Forest Park for the local chronic disease patient group are continuing to improve year by year as the quality of the built environment improves. However, at the level of the assessed components, the performance of the D_1 and D_3 components improved significantly; however, the D_2 and D_4 components showed varying degrees of decline in points.

Table 3: Weighting of the assessed dimensions/elements.

Dimensions	Local weights	Elements	Local weights	Global weights
Body practice and physical movement (D_1)	0.188	Exercises to enhance body control (C_1)	0.250	0.046
		A wide range of leisure and sporting activities (C_2)	0.300	0.056
		Facilitating the material movement of groups (C_3)	0.450	0.085
Access and bond to nature (D_2)	0.062	Planned planting of medicinal plants (C_4)	0.100	0.006
		Healthy fruit and vegetable growing areas (C_5)	0.170	0.011
		Open lawn green areas (C_6)	0.480	0.030
		Calm water features (C_7)	0.250	0.016
Sociality and symbolic environment (D_3)	0.375	Promoting a wide range of social interactions (C_8)	0.430	0.162
		Something symbolic of the community (C_9)	0.360	0.135
		Hands-on interactive facilities (C_{10})	0.210	0.078
Sensory experience (D_4)	0.375	Touching in a specific cultural context (C_{11})	0.240	0.090
		A fresh scent with a difference from every day (C_{12})	0.250	0.091
		Rich sounds from nature (C_{13})	0.280	0.106
		Interior exterior green visual connections (C_{14})	0.230	0.088

Table 4: Assessment of therapeutic landscapes based on well-being and health care for people with chronic diseases.

2004



Dimensions	Body practice and physical movement (D_1)			Access and bond to nature (D_2)				Sociality and symbolic environment (D_3)			Sensory experience (D_4)			
Gap ratio	0.68			0.57				0.86			0.44			
Elements	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9	C_{10}	C_{11}	C_{12}	C_{13}	C_{14}
Gap ratio	0.93	0.43	0.71	0.97	0.74	0.38	0.68	0.78	0.93	0.92	0.96	0.24	0.20	0.40
Total performance: 3.44														

2012

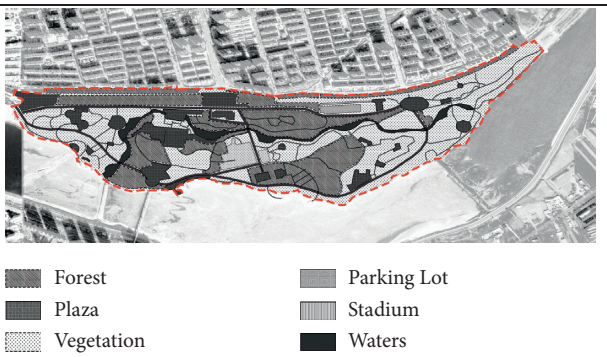
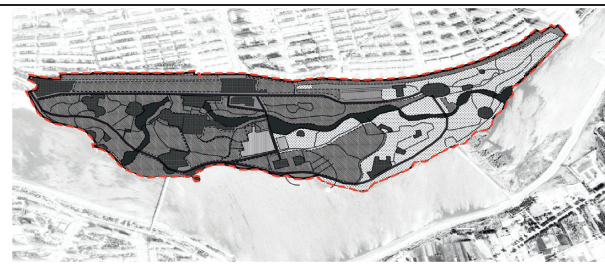


Table 4: Continued.

Dimensions	Body practice and physical movement (D_1)			Access and bond to nature (D_2)				Sociality and symbolic environment (D_3)			Sensory experience (D_4)			
Gap ratio	0.54			0.60				0.55			0.58			
Elements	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9	C_{10}	C_{11}	C_{12}	C_{13}	C_{14}
Gap ratio	0.60	0.68	0.42	0.68	0.57	0.69	0.43	0.44	0.64	0.59	0.88	0.56	0.50	0.41
Total performance: 4.35														
2020														
	Forest							Parking Lot						
	Plaza							Stadium						
	Vegetation							Waters						
Dimensions	Body practice and physical movement (D_1)			Access and bond to nature (D_2)				Sociality and symbolic environment (D_3)			Sensory experience (D_4)			
Gap ratio	0.36			0.56				0.42			0.46			
Elements	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9	C_{10}	C_{11}	C_{12}	C_{13}	C_{14}
Gap ratio	0.43	0.36	0.31	0.53	0.37	0.72	0.39	0.31	0.56	0.42	0.80	0.42	0.38	0.28
Total performance: 5.64														

This implies that the access and exposure to the natural environment dimension in the East Lake Forest Park have not improved with the built environment from the perspective of the chronically ill, and even the open green space is decreasing, and medicinal plants are not being used sufficiently in the vegetation community planning. Therefore, this study suggests that the current East Lake Forest Park needs to stimulate and improve the senses of people with chronic illnesses who enter the site through certain colors, rhythmic sounds, directional smells, operable interactive facilities, etc., so that they can detach themselves from their negative or undesirable state. In addition, this study suggests an attempt to make medicinal plants the basic plants of the entire planting community. Using a large number of plants to create visual amenities, public services can be placed near the medicinal planting community. In addition, referring to the influence networks in Figure 1, this study argues that the social and symbolic nature of the urban landscape can be used to achieve some kind of spiritual healing through activities, services, and even rituals, creating a so-called therapeutic encounter, which is a potential function not revealed in past concepts of community design, a process of psychological well-being and health through "participation." This also responds to the complex process of healing taking place in the aforementioned contexts mentioned by W. M. Gesler, which are not just spatially generated therapeutic outcomes but are more often dependent on the relational interactions that arise in these contexts, some kind of empowerment, inclusion or, as mentioned above, collective social behavior, and the following of symbolic imagery.

On the other hand, this study recommends that local policymakers continue to promote the regeneration and quality of the East Lake Forest Park by designing different types of privacy and social spaces and turning the unused green space into an accessible rehabilitation garden that can be used by the elderly and chronically ill to provide appropriate physiotherapy and even light exercise, for example, by providing gravel paths to balance the rehabilitation therapy walking exercises. Appropriately placed fruit and vegetable growing areas within the grounds also allow for the production of safe, nutritious vegetables. This study proposes a reconstruction of the building with greenhouses. Its function should be updated. In the past, the greenhouse had everything people need for herbs. It could also be utilized during wintertime. After renovation, chronic disease patients can cultivate healthy fruits and vegetables suitable for nutrition here. Fig trees or citrus plants can also be cultivated along with different kinds of herbs having a variety of smells and tastes (for example, basil, horsemint, bee balm, basil thyme, chive, or meadow sage). While growing herbs, patients can improve their management and responsibility skills as well as their interest in the natural environment. These activities can also enable them to gain better social skills and communication in a team. Additionally, such activity makes patients less

stressed during their rehabilitation therapy, and they can feel themselves be more efficient and successful. What needs to be maintained and further enhanced is the development of the Jinzhou East Lake Forest Park as an area that offers good group mobility to a wide range of people. The concept of continuous "flow" is important in a therapeutic landscape where one may be so fully engaged in physical activity that different concerns and stresses may be temporarily forgotten. As Doughty [42] found, group walking can be seen as a supportive social space through shared movement and social relationships within the environment.

5. Conclusions

This study provides useful insights for examining the therapeutic nature of everyday urban landscapes as they evolve over the course of urban development and for planning landscape development to enhance people's well-being. Firstly, based on a review of the relevant literature, a systematic assessment framework was developed that can be used to assess the therapeutic nature of everyday urban landscapes, and then, a hierarchical assessment framework was constructed based on the qualitative data collected through focus group interviews and inductive analysis. The ANP technique was then applied to train the relative weights of the components/elements in the assessment framework. Finally, the modified VIKOR technique was applied to determine the value of the gap between the actual level and the desired level of Jinzhou East Lake Forest Park in 2004, 2012, and 2020 as the assessment cases. This study found that the healing and therapeutic effect of Jinzhou's East Lake Forest Park on the local chronic disease patient population has been increasing year by year. The park has been developed over the years and is suitable for a variety of recreational sports, group exercise, and mobility and a wide range of social interaction activities. In addition to this, the healing and positive therapeutic benefits that the case can unleash on people with chronic illnesses and its landscape nodes already have some degree of symbolic impact. However, compared to 16 years ago, the level of access and contact with the natural environment in the East Lake Forest Park has not improved with the built environment, and even the open green spaces are declining, and medicinal plants are not being used enough in the vegetation community planning. Therefore, this study suggests that future development planning for Jinzhou East Lake Forest Park should focus on a more microscopic landscape design level, designing different types of open green spaces, improving the utilization of unused green spaces, and providing the public with a richer sensory experience through certain colors, rhythmic sounds, and directional smells. For local policymakers, this study provides an effective way to examine the therapeutic nature of everyday urban landscapes as urban sprawl shifts to contraction and to explore planning for the continued development of therapeutic landscapes in relation to networks of influence between landscape elements.

The limitations of this study should be acknowledged, as they may provide guidance for future research. Respondents recruited for this study were from a broad group of people with chronic diseases earning socially acceptable wages, and the analysis does not apply to people with chronic diseases who are of higher socioeconomic status or who live in different geographical locations and climates. In addition, elements of the therapeutic landscape of the East Lake Forest Park in 2004 and 2012 can only be discussed and generalized based on recollections of recruited respondents and personal experience judgments in focus group interviews. The quantitative techniques applied in this study were additive in nature during the performance evaluation phase of the empirical cases, which may have detracted from reality in the ranking and selection of performance. In future research, nonadditive quantitative techniques (e.g., fuzzy integral method) could be attempted to conduct performance assessment analysis of multiperiod therapeutic landscapes. It is important to acknowledge that the therapeutic landscape development planning strategies explored in this study are specific to Jinzhou City Forest Park and are not applicable to everyday urban landscapes in other areas. However, the assessment methodology provided in this study and the assessment framework constructed within a decision management perspective may provide a theoretical basis and a key reference for future examination of the therapeutic nature of urban landscapes in other regions.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

The authors are grateful to all the landscape designers who took part in this study. The authors would like to thank Yu-Feng Luo, the research assistant, for all the hard work and effort put in to make the questionnaire survey a success. The authors are also immeasurably grateful to the alumni association of the environmental design, School of Architecture, Harbin Institute of Technology (HIT), for the support during the research. The authors are grateful to the General Project of Philosophy and Social Sciences in Jilin Province (2019B104) and the Specialized Subsidy Scheme for Macao Higher Education Institutions in the Area of Research in Humanities and Social Sciences (and Specialized Subsidy Scheme for Prevention and Response to Major Infectious Diseases) (HSS-MUST-2020-9) for their financial support.

References

[1] T. Hartig and H. Staats, "Restorative environments," *Journal of Environmental Psychology*, vol. 23, no. 2, 2003.

[2] K. Nilsson, C. Baines, and C. C. Konijnendijk, *Health and the Natural Outdoors-Final Report of the COST Strategic Workshop*, COST, Brussels, Belgium, 2007.

[3] W. M. Gesler, "Therapeutic landscapes: medical issues in light of the new cultural geography," *Social Science & Medicine*, vol. 34, no. 7, pp. 735–746, 1992.
[4] D. Conradson, "Landscape, care and the relational self: therapeutic encounters in rural England," *Health & Place*, vol. 11, no. 4, pp. 337–348, 2005.
[5] A. Williams, "Spiritual therapeutic landscapes and healing: a case study of St. Anne de Beaupre, Quebec, Canada," *Social Science & Medicine*, vol. 70, no. 10, pp. 1633–1640, 2010.
[6] L. Brewster, "The public library as therapeutic landscape: a qualitative case study," *Health & Place*, vol. 26, pp. 94–99, 2014.
[7] S. L. Bell, C. Phoenix, R. Lovell, and B. W. Wheeler, "Seeking everyday wellbeing: the coast as a therapeutic landscape," *Social Science & Medicine*, vol. 142, pp. 56–67, 2015.
[8] X. Yan and S. He, "The co-evolution of therapeutic landscape and health tourism in bama longevity villages, China: an actor-network perspective," *Health & Place*, vol. 66, Article ID 102448, 2020.
[9] J. Biglin, "Embodied and sensory experiences of therapeutic space: refugee place-making within an urban allotment," *Health & Place*, vol. 62, Article ID 102309, 2020.
[10] W. Gesler, "Therapeutic landscapes: an evolving theme," *Health & Place*, vol. 11, no. 4, pp. 295–297, 2005.
[11] R. Sampson and S. M. Gifford, "Place-making, settlement and well-being: the therapeutic landscapes of recently arrived youth with refugee backgrounds," *Health & Place*, vol. 16, no. 1, pp. 116–131, 2010.
[12] D. Alaazi, *Aboriginality, Homelessness, and Therapeutic Landscapes of home: Mapping the Experiences of Aboriginal Housing First Participants in Winnipeg*, University of Manitoba, Winnipeg, Canada, 2013.
[13] I. Belčáková, P. Galbavá, and M. Majorošová, "Healing and therapeutic landscape design–examples and experience of medical facilities," *ArchNet-IJAR: International Journal of Architectural Research*, vol. 12, no. 3, p. 128, 2018.
[14] J. Finlay, T. Franke, H. McKay, and J. Sims-Gould, "Therapeutic landscapes and wellbeing in later life: impacts of blue and green spaces for older adults," *Health & Place*, vol. 34, no. 97, pp. 97–106, 2015.
[15] A. E. Cheesbrough, T. Garvin, and C. I. J. Nykiforuk, "Everyday wild: urban natural areas, health, and well-being," *Health & Place*, vol. 56, pp. 43–52, 2019.
[16] B.. Zou, "Increment planning, inventory planning and policy planning," *City Planning Review*, vol. 2, pp. 35–37, 2013.
[17] H. Li, K. Lo, and M. Wang, "Economic transformation of mining cities in transition economies: lessons from Daqing, Northeast China," *International Development Planning Review*, vol. 37, no. 3, pp. 311–328, 2015.
[18] M. Busa, *Designing for the Shrinking City: Re-imagining Burke Lakefront Airport in Cleveland*, OH (Doctoral Dissertation, The Digital Repository at the University of Maryland (DRUM), College Park, MD, USA, 2013.
[19] P. Zhou, J. Wu, and X. Wu, "Research on shrinking cities' renewal strategy based on green infrastructure construction," *Urban Planning International*, vol. 32, no. 1, pp. 91–98, 2017.
[20] M. Ročak, G. J. Hospers, and N. Reverda, "Searching for social sustainability: the case of the shrinking city of Heerlen, The Netherlands," *Sustainability*, vol. 8, no. 4, p. 382, 2016.
[21] F. Ortiz-Moya, "Green growth strategies in a shrinking city: t," *Journal of Urban Affairs*, vol. 42, no. 3, pp. 312–332, 2020.
[22] Y. Chen and D. Zhang, "Evaluation and driving factors of city sustainability in Northeast China: an analysis based on interaction among multiple indicators," *Sustainable Cities and Society*, vol. 67, Article ID 102721, 2021.
[23] L. G. Ma, *Analysis of the Regional Balance of Health Status and Health Service Utilization Among the Elderly in China and its Changing Trends*, Anhui Medical University, Hefei, China, 2012.
[24] Y. You, L. Ning, M. Wu et al., "Analysis of the current prevalence and influencing factors of major chronic diseases among urban residents in Liaoning Province," *Preventive Medicine*, vol. 30, no. 1, pp. 35–41, 2018.
[25] Y. Ye, T. Fei, and H. Mei, "The relationship between walkability and environment characteristics in cold region cities: case study in Harbin," *IOP Conference Series: Earth and Environmental Science*, vol. 63, no. 1, Article ID 012053, 2017.
[26] P. Liamputtong and D. Suwankhong, "Therapeutic landscapes and living with breast cancer: the lived experiences of Thai women," *Social Science & Medicine*, vol. 128, pp. 263–271, 2015.
[27] C. Kershaw, B. Marques, and J. McIntosh, "Rehabilitating healthcare: healthcare landscapes a catalyst for health, well-being and social equity," in *Proceedings of the 52nd International Conference of the Architectural Science Association*, Melbourne, Australia, November 2018.
[28] C. Duff, "Exploring the role of 'enabling places' in promoting recovery from mental illness: a qualitative test of a relational model," *Health & Place*, vol. 18, no. 6, pp. 1388–1395, 2012.
[29] J. A. Hipp and O. A. Ogunseitan, "Effect of environmental conditions on perceived psychological restorativeness of coastal parks," *Journal of Environmental Psychology*, vol. 31, no. 4, pp. 421–429, 2011.
[30] H. Kamioka, K. Tsutani, M. Yamada et al., "Effectiveness of horticultural therapy: a systematic review of randomized controlled trials," *Complementary Therapies in Medicine*, vol. 22, no. 5, pp. 930–943, 2014.
[31] C. Maller, M. Townsend, A. Pryor, P. Brown, and L. St Leger, "Healthy nature healthy people: 'contact with nature' as an upstream health promotion intervention for populations," *Health Promotion International*, vol. 21, no. 1, pp. 45–54, 2006.
[32] M. D. Velarde, G. Fry, and M. Tveit, "Health effects of viewing landscapes–landscape types in environmental psychology," *Urban Forestry and Urban Greening*, vol. 6, no. 4, pp. 199–212, 2007.
[33] V. J. Wood, W. Gesler, S. E. Curtis et al., "'' and the importance of nostalgia, solastalgia, salvage and abandonment for psychiatric hospital design," *Health & Place*, vol. 33, pp. 83–89, 2015.
[34] R. Foley and T. Kistemann, "Blue space geographies: enabling health in place," *Health & Place*, vol. 35, pp. 157–165, 2015.
[35] W. M. Gesler, *Healing Places*, Rowman & Littlefield, Lanham, MD, US, 2003.
[36] R. S. Ulrich, "Biophilia, biophobia, and natural landscapes," *The biophilia hypothesis*, vol. 7, pp. 73–137, 1993.
[37] C. C. Marcus and N. A. Sachs, *Therapeutic Landscapes: An Evidence-Based Approach to Designing Healing Gardens and Restorative Outdoor Spaces*, John Wiley & Sons, New York, NY, USA, 2013.
[38] X. B. Wang and J. F. Li, "Analysis of the healing landscape and its relevant conceptions," *Journal of Beijing University of Agriculture*, vol. 27, no. 2, pp. 71–73, 2012.
[39] C. Duff, "Networks, resources and agencies: on the character and production of enabling places," *Health & Place*, vol. 17, no. 1, pp. 149–156, 2011.

[40] E. Rose, "Encountering place: a psychoanalytic approach for understanding how therapeutic landscapes benefit health and wellbeing," *Health & Place*, vol. 18, no. 6, pp. 1381–1387, 2012.
[41] H. Pitt, "Therapeutic experiences of community gardens: putting flow in its place," *Health & Place*, vol. 27, pp. 84–91, 2014.
[42] K. Doughty, "Walking together: the embodied and mobile production of a therapeutic landscape," *Health & Place*, vol. 24, pp. 140–146, 2013.
[43] A. Butterfield and D. Martin, "'s as therapeutic landscapes," *Landscape Research*, vol. 41, no. 6, pp. 695–706, 2016.
[44] R. Gorman, "Smelling therapeutic landscapes: embodied encounters within spaces of care farming," *Health & Place*, vol. 47, pp. 22–28, 2017.
[45] Q.-B. Song, Y. Zhao, Y.-Q. Liu, J. Zhang, S.-J. Xin, and G.-H. Dong, "Sex difference in the prevalence of metabolic syndrome and cardiovascular-related risk factors in urban adults from 33 communities of China: the CHPSNE study," *Diabetes and Vascular Disease Research*, vol. 12, no. 3, pp. 189–198, 2015.
[46] T. L. Saaty and L. G. Vargas, *Decision Making with the Analytic Network Process*, Vol. 282, Springer Science+ Business Media, LLC, Berlin, Germany, 2006.
[47] L. Xu, "Restorative environment, health and green urbanism," *Southern Architecture*, vol. 3, pp. 101–107, 2016.
[48] B.-W. Zhu, J.-R. Zhang, G.-H. Tzeng, S.-L. Huang, and L. Xiong, "Public open space development for elderly people by using the DANP-V model to establish continuous improvement strategies towards a sustainable and healthy aging society," *Sustainability*, vol. 9, no. 3, p. 420, 2017.
[49] K. Wang, Q. Cui, and H. Xu, "Desert as therapeutic space: cultural interpretation of embodied experience in sand therapy in Xinjiang, China," *Health & Place*, vol. 53, pp. 173–181, 2018.
[50] M. Q. Li and Q. L. Xu, "Research and practice of healing environment from public health perspective," *Journal of Human Settlements in West China*, vol. 35, no. 5, pp. 39–47, 2020.

Hindawi
Journal of Healthcare Engineering
Volume 2021, Article ID 2858536, 10 pages
https://doi.org/10.1155/2021/2858536

Research Article

Event-Related Potential Sensing Analysis on the Risk Perception and Decision-Making by Grassroots Managers in Different Fatigue States

Yixin Huang,[1,2] **Hongxia Li**,[3] **Yanling Yang**,[4] **Jin Wang**,[1,2] **and Janis Jansz**[5]

[1]*School of Economics and Management, Yan'an University, Yan'an 716000, China*
[2]*Soft Science Research Base for Green and Low Carbon Development of Energy Industry in Shaanxi Province, Yan'an University, Yan'an 716000, China*
[3]*Graduate School, Xi'an University of Science and Technology, Xi'an 710054, China*
[4]*Medical School, Yan'an University, Yan'an 716000, China*
[5]*School of Mines: Minerals, Energy and Chemical Engineering, Faculty of Science and Engineering, Curtin University, Perth 6945, Australia*

Correspondence should be addressed to Yixin Huang; huangyixin01@126.com

Received 16 April 2021; Accepted 8 September 2021; Published 24 September 2021

Academic Editor: Chia-Yu Lin

The risk perception and decision-making ability of grassroots managers is the key to the normal operation of enterprises. This study used event-related potential indicators (ERPs) to reveal the process of risk perception and decision-making behaviour of coal mine grassroots managers in different fatigue states. The ERP components, such as CNV, P300, MMN, and FRN, during risk perception, decision-making, and postperception periods were obtained and evaluated. The peak value and variation characteristics of ERP components of grassroots managers under fatigue and nonfatigue conditions were analysed. Accordingly, the effectiveness of decision-making behaviour in different periods was determined. The results showed that the P300 component is a key indicator in measurements of the deviation of grassroots managers' decision-making behaviour, and FRN could reflect the negative emotions in the decision-making process and reflect the sensitivity of the risk perception of grassroots managers. There was a significant difference between the peak voltages of the ERP components of the grassroots managers in fatigue and nonfatigue states. The peak voltage of the ERP components of the grassroots managers in a fatigue state was generally greater than 10 μV; therefore, the quality of decision-making by the grassroots managers could be evaluated according to the characteristics of the ERP components. This study provides a risk decision-making reference for grassroots managers of coal mine enterprises.

1. Introduction

In recent years, infrastructure construction in developing countries has been in full swing. A large number of large geotechnical engineering projects in the fields of water conservancy, mining, and underground engineering are under construction [1–4]. In this context, the contradiction between the construction needs of geotechnical engineering and the relatively insufficient technology or lagging management of engineering in developing countries has become prominent [5]. This contradiction leads to frequent accidents in the process of engineering construction in developing countries. Taking China as an example, engineering accidents in China has been highly detrimental in recent years, resulting in high economic losses and casualties. Of these, the proportion of accidents involving geotechnical engineering has been as high as over 10% [6]. Among these engineering accidents, coal mine accidents in China have always been the focus of society as a whole [7]. According to statistics in the first half of 2020, in Shanxi Province alone, there were 5 coal mines with fatal accidents with 8 deaths, and 150 mines were suspended for rectification. Ensuring the safety of coal mine workers is the key to ensuring the healthy and sustainable development of mining enterprises.

Risk decision-making analysis is a key link in enterprise quality management and operational risk assessment [8, 9].

To perform risk analysis and decision-making processes, a multiattribute analysis method must be established, and the most important influencing factors affecting enterprise operational risks must be screened out [10, 11]. Xiong et al. analysed the causes of 50 typical engineering accidents and found that 55.6% of the accidents were caused by human factors [6]. Long-term monotonous labour and harsh environments can easily lead to heavy mental stress and physical or mental fatigue in miners [12, 13]. For some grassroots managers, long-term fatigue could cause them to have a negative and insensitive attitude to coping with work. Therefore, some grassroots managers could lose the sensitivity of risk perception in daily work, and they would exhibit incorrect behaviours when coping with risk decisions. This would directly cause production accidents [14]. As a result, poor decision-making behaviours would occur when dealing with risks, which would directly cause accidents. Hence, grassroots managers' perception of risks and their ability to make correct decisions are important factors in the prevention of safety accidents. An increasing number of research institutes of coal enterprises are committed to integrating the physiological response of managers under specific risk stimulus conditions with the evaluation system of decision-making and to studying the physiological components and change indicators of managers in the process of risk decision-making [15]. Accordingly, the work status and behaviour of managers are identified, and the decision-making behaviours of managers are scientifically evaluated.

It is beneficial to use modern experimental research techniques, such as event-related potentials (ERPs) [16], magnetoencephalograms [17], and functional magnetic resonance imaging [18], to analyse the physiological indicators of the decision-making process of human beings. Among these technical methods, ERP research has penetrated many fields, such as psychology, physiology, medicine, neuroscience, and artificial intelligence and is increasingly recognized by people [19–21]. As early as the 1970s, Squires et al. proposed the concept of ERPs, which recorded brain evoked potentials from the surface of the skull using averaging and superposition technology to reflect the neuroelectrophysiological changes in the brain during the cognitive process [22]. Given the close relationship between ERPs and the cognitive process, ERPs are considered to be windows into the mental activities of people. With the continuous application of ERP experimental research, an increasing number of ERP components have been discovered. Walter et al. first proposed the contingent negative variation (CNV) component during the ERP testing process in 1964 [23]. In addition to the CNV component, components such as MMN, P300, and FRN are also widely used in ERP tests [24–27]. In recent decades, an increasing number of people have applied ERP methods to the field of engineering safety management. These studies mainly focused on the unsafe behaviour of subjects and the identification of risk patterns of physiological indicators. For instance, ERP tests are used to evaluate the physiological indicators of the driver in each stage of "cognition-perception-response" [28]. ERP experimental research in the field of coal mine safety management mainly focuses on the safety psychology and unsafe behaviour of miners, discussing the influence of the safety psychological capital of miners on unsafe behaviour from the perspective of brain nerve mechanisms [29].

Although some research progress has been made in engineering safety analysis by ERP technology, there are still few studies on risk perception and decision-making with grassroots managers as the subjects. Especially in coal mining engineering, various coal mine dynamic disasters, such as coal bumps [30] and gas outbursts [31] seriously threaten the personal safety of workers. Therefore, it is of great significance to evaluate the risk perception and decision-making process of coal mine grassroots managers. In this study, several grassroots managers of a coal mine are selected as subjects, and some representative risk decision-making cases in daily work are selected as stimulus conditions. The ERP tests of risk decision-making under fatigue and nonfatigue states of grassroots managers are carried out in groups. The experimental process is divided into the risk perception period, decision period, and postperception period. The characteristics of ERP components, such as CNV, P300, MMN, and FRN, of the participants of the three periods are studied, and the peak voltage and change rule of different ERP components are statistically analysed. Accordingly, ERP component index thresholds that can evaluate the risk decision-making behaviour of grassroots managers in coal mines are proposed. This study provides a reference for the risk decision-making of grassroots managers of coal mine enterprises.

2. Experimental Design and Data Processing

2.1. Experimental Design and Preparation. Due to the complexity of mining work, grassroots managers in coal mines have a low sense of risk perception and a high rate of decision errors. The decision-making level of grassroots managers is directly related to the safety of miners and the production efficiency of coal mines. Evaluating the risk perception and decision-making level of grassroots managers is important to ensure the safe and efficient production of coal mines.

ERPs are special brain evoked potentials that use the potential of the brain caused by multiple or diverse stimuli by deliberately giving stimuli a special psychological meaning. It reflects the changes in the neurophysiology of the brain during the cognitive process. It is also called cognitive potential, which refers to the brain potential recorded from the surface of the head when people perform cognitive processing on a certain subject. ERP experiments are conducted to clarify the risk perception and decision-making level of grassroots managers in coal mines during the decision-making process. Grassroots managers such as trade union chairmen, mining team captains, and ventilating team captains of multiple coal mines are selected as the subjects. The ERP experimental system of the risk perception and decision-making of coal mine managers is constructed as follows: the ERP test machine (Version: B-Alert 3.0) is connected in parallel with the sensor electrodes, and the

terminals are connected to the computer. The E-Prime 2.0 program platform is used to collect and analyse ERP signals, and the cases matching the complex coal mine project decision events are selected as the stimulus conditions. The ERP experiment test system for risk perception and decision-making is shown in Figure 1, and the ERP experiment equipment, material library, and stimulus conditions required for the specific decision and perception process are shown in Figure 2.

This experimental study is divided into ERPs testing of risk perception and risk decision-making. The characteristics of ERP signal components were analysed during risk perception and decision-making. This experiment was divided into risk perception and risk decision-making ERP testing and analysed the characteristics of ERP signal components in the early stage of risk perception and the risk decision-making process. The design experiment was divided into two groups: the risk decision ERP test group and the risk perception ERP test group. The risk perception ERP experiment was based on the completion of the risk decision experiment. The decision-making ERP experiment was divided into 9 control groups, and 12 grassroots managers with experience in large-scale coal mine project construction were selected. The stimulation conditions were divided into positive and negative groups (control group), with 6 subjects in each group. The test results are divided into three categories: "correct," "positive meaning shift," and "negative meaning shift." Take the ERPs experiment in the process of risk decision-making under fatigue stimulus conditions as an example. Twelve subjects in each group will be divided into groups to carry out "positive" and "negative" stimulus experiments; that is, the subjects will make decisions during the exhausted state and the good state to test fatigue. Auditory response and visual induction response under loading conditions such as stress, simultaneous handling of multiple events, and crisis attributes comprehensively evaluate the risk decision-making and perception level of coal mine managers in the risk decision-making process.

The ERP signals were analysed through the E-prime 2.0 program platform. The important nodes and fluctuation characteristics of the signal data are marked. According to the decision-making process, the experiment is divided into three stages: consciousness period (Block 1), decision period (Block 2), and postperception period (Block 3). To exclude the influence of time factors on risk perception and changes in ERP composition characteristics, the experiment time was limited to 60 s. The duration of the consciousness period (Block 1), decision period (Block 2), and postperception period (Block 3) were 20 s, 30 s, and 10 s, respectively. To ensure the correlation between the two experiments and the validity of the data of the subjects, the test subjects rested for 5 minutes after the first experiment. To ensure the relevance of the two experiments and the validity of the test data, the rest period between groups was set to 5 minutes. During the experiment, key nodes and abnormal signal zones are manually selected, and different types of ERP data are recorded at different times. This is helpful for the follow-up observation of the ERP component characteristics of grassroots managers during the risk perception stage and decision-making stage.

2.2. ERP Signal Filtering and Analysis Method. When the brain performs tasks or receives external stimuli, it generates electrical signals (usually 1.0 to 100 mV). By arranging potential sensors on different parts of the brain surface, the weak electrical signals can be filtered and amplified, and a continuous waveform can be obtained. These electrical signals could be divided into electroencephalograms (EEGs) and event-related potentials (ERPs). EEG scans show the continuous brain electrical signal waveform obtained by the electrode sensor, while ERPs depict the electrical signal waveform of the evoked potentials before and after the trigger event. This paper carried out an ERP experiment on the risk decisions and perceptions of grassroots managers in fatigue and nonfatigue states. Waveforms of ERP components, such as contingent negative variation (CNV), P300, mismatch negativity (MMN), and feedback-related negativity (FRN), are obtained during the risk decision-making process of grassroots managers in coal mines. The peak voltage of different components of coal mine grassroots managers during the decision-making process was extracted. The level of risk perception and decision-making quality of coal mine grassroots managers are judged. This provides support for the construction of an intelligent risk management and control system for complex coal mine projects. Figure 3 shows the experimental methods of ERPs and waveforms of ERP signals.

During the ERP test of coal mine grassroots managers, the frequency of ERP components fluctuated frequently due to the stimulation of external conditions and events. ERP components are different under loading conditions, such as fatigue, stress, multievent handling, aural induction, and visual induction. In coal mine production, the decision-making events of grassroots managers include the replacement and maintenance dates of common equipment, turn-off dates of miners, safety qualification checks for miner positions, and daily workload arrangements. Different decision events correspond to different test potentials, so the ERP tests are performed on the subjects using the one-by-one measurement method. The extracted signal components mainly include CNV, MMN, and P300. In addition, the N4 component (related to semantic understanding matching), N2 component (related to auditory induction), and FRN component (related to intention processing) are observed. Published ERP research shows that the early CNV component mainly determines the concentration and response level of the subject in the early stage of decision-making. N2 reflects the early visual image of decision-making by the subject. MMN reflects the auditory image in the decision-making process of the subject. P300 reflects the concentration and response level during the decision period. The late CNV component can measure the level of alertness of the subject. This paper mainly analyses the characteristics of ERP signals in the risk perception and decision-making process of coal mine grassroots managers. The ERP components in the awareness period (Block 1), decision period (Block 2), and postperception period (Block 3) during the experiment are compared. The feature of voltage change reveals

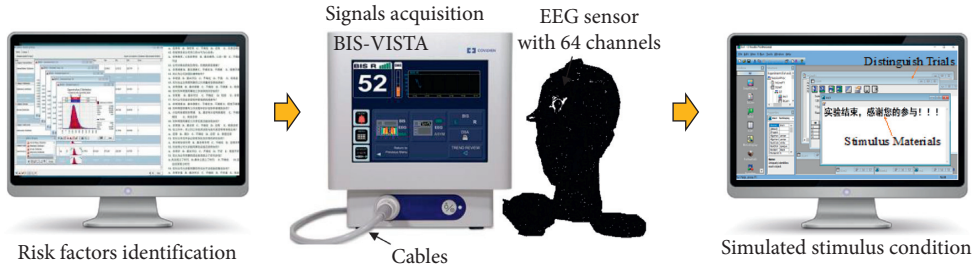


Figure 1: ERP experimental equipment and simulated stimulation conditions.

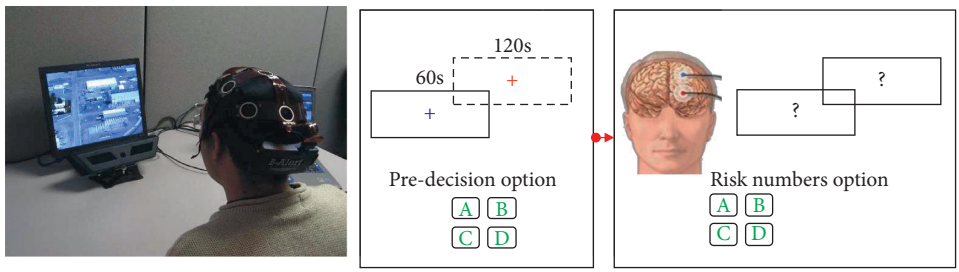


Figure 2: Experimental stimulation conditions and risk decision process.

the risk perception and decision-making of coal mine grassroots managers to reveal the risk perception and decision-making rules of coal mine grassroots managers.

3. ERP Characteristics of Decision-Making by Grassroots Managers

3.1. Peak Voltage Variations of CNV and MMN during the Risk Perception Period. To explore the influence of negative stimuli on the risk decision-making of grassroots managers in coal mines, the experimental grassroots managers were divided into two groups: a nonfatigue group and a fatigue group. The following takes equipment replacement as a decision event as an example to explain the test process. First, a predictive signal *S*1 (conditional stimulus) is applied to the subject before providing a decision-making event; then, the command signal *S*2 is applied when the decision-making event is executed. The two stimuli are separated by 1 to 2 seconds. The recorder will obtain ERP signals, such as CNVs and MMNs, during the process. In the tests, subjects numbered *P*1 to *P*6 belong to the nonfatigue group, and subjects *P*7 to *P*12 belong to the fatigue group. The test process is divided into three periods: Block 1 (risk perception period, 0–220 s), Block 2 (decision-making period, 20–550 s), and Block 3 (postperception period 50–660 s). The changes in CNV and MMN were analysed during the risk perception period and the results are provided below.

The CNV change characteristics and peak voltage results of the grassroots managers during the risk perception period (Block 1) are shown in Figure 4. The 6 subjects in the nonfatigue group all produced CNV components. The peak voltage time ranges 9.15~11.23 s, and the peak value ranges 6.5~10.1 μV. For the fatigue group, only 4 subjects produced CNV components, the peak time ranged from 11.26 to 11.86 s, and the peak value ranged 5.5 μV~5.9 μV. Taking Subject 5 in the nonfatigue group and Subject 10 in the fatigue group as examples, the impact of fatigue on the decision-making behaviour of grassroots managers in the coal mine is analysed. During the risk perception period, the CNV component induction time of Subject 5 is 11230 ms with a duration of 50 ms, and the peak value reaches 10.1 μV with an amplitude of 12.1 μV. CNV gradually stabilized after the peak. The changing trend of CNV of Subject 10 is consistent with that of Subject 5. However, the CNV component induction time is 11260 ms with a duration of 50 ms, and the peak value is 5.8 μV with an amplitude of 7.7 μV. The CNV component of Subject 10 induces a longer latency period and lower peak voltage, which indicates that the negative wave of the CNV component is more active in the nonfatigue condition during the risk perception period (Block 1). Therefore, the characteristics of the CNV component can be used to judge whether the grassroots managers in coal mines are in a state of fatigue during the risk perception period. During the construction of risk management and intelligent control systems for coal mines, it is possible to consider designing and installing matching potential sensors in miners' helmets to monitor the characteristics of the CNV components in real-time and establish scientific risk management and early warning platforms based on this.

The MMN component is related to the early preprocessing activities of auditory information, reflecting the activation process of the primary auditory cortex and adjacent superior temporal gyrus. The peak MMN component results of 6 nonfatigue subjects during the risk perception period are shown in Figure 5. In the tests, the subjects in the

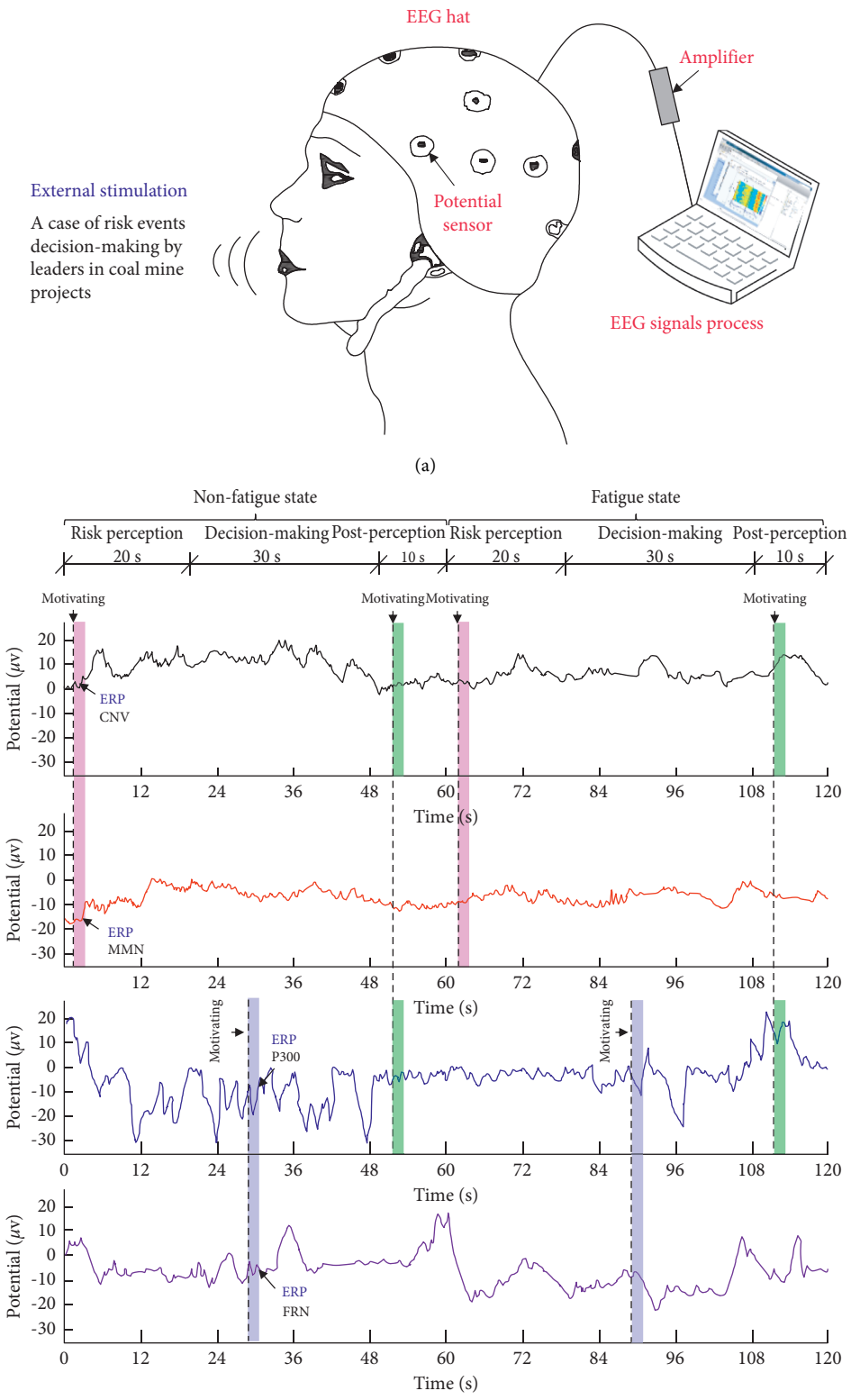
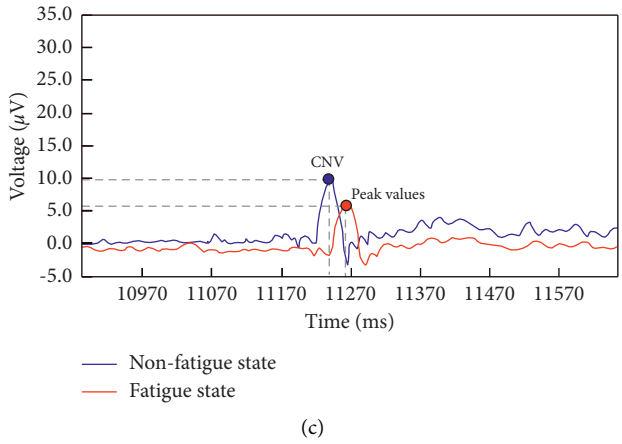


Figure 3: Continued.



(c)

Figure 3: Experimental methods of ERPs and waveforms of ERP signals.

fatigue group produced no MMN components during the perception period, while 4 subjects in the nonfatigue group produced MMN components (the peak voltage was negative). This indicates that the nonfatigue group has significantly stronger anti-interference ability than the fatigue group in terms of auditory evoked potential during the tests. Therefore, it can be considered to judge the working status of grassroots managers in coal mines during the risk perception period with the characteristics of MMN components.

3.2. Peak Voltage Variations of P300 and FRN during the Decision-Making Period. It is of great significance to study the characteristics of ERP components in the process of risk decision-making of grassroots managers in coal mines. The P300 component was discovered and proposed by Sutton et al. in 1965 and belongs to the third positive wave. P300 usually appears approximately 300 ms. Generally, P300 is widely used in the judgement, early diagnosis, and adjuvant treatment of various diseases caused by brain injury. At present, most studies focus on the use of P300 to study dementia diseases. Therefore, the P300 component characteristics in the decision-making period of grassroots managers in coal mines are the key information indicators for judging the biased stimulus of decision-making behaviour. (Figure 6(a)) shows the characteristics of the P300 component during the decision-making period of the subjects, and the statistical analysis of the peak voltage of the P300 component during the decision-making period of 12 subjects is shown in (Figure 6(b)). The results show that the P300 component induction levels of the 6 subjects in the nonfatigue group are relatively consistent, with an average induction time of 300.16 ms and a peak value of 8.491 μV. The P300 component induction level of the 6 subjects in the fatigue group was different. The earliest and latest induction times are 300 and 300.45 ms, respectively. The average induction time was 300.22 ms, and the average peak voltage was 13.19 μV. Therefore, the P300 component of the test subject has relatively high voltage feedback in the fatigue state, which indicates that the fatigue test subject has weak anti-interference ability, insufficient concentration, and weak decision-making willingness. The feedback of coal mine managers in the state of fatigue to emotional motivation is more obvious. In the construction of the coal mine risk management and intelligent perception system platform, the peak voltage of the measured P300 component can be considered as the threshold to determine whether the risk decision of coal mine grassroots managers under fatigue and nonfatigue is effective.

The FRN component is the feedback negative wave in the midline area of the forehead of the scalp. The negative wave component appears within approximately 200–400 ms after the feedback stimulus is presented. FRN reflects the difference between expected and actual results. It can represent a negative wave component induced by negative feedback stimuli such as behaviour errors. Gehring et al., 2002, proposed the emotional motivation hypothesis that FRN reflects the evaluation process of the meaning of emotional motivation. Therefore, FRN can reflect the negative emotions of coal mine managers when they make the wrong risk decisions in this study. Figure 7 shows the FRN component signal characteristics and peak voltage comparisons of 12 subjects in the risk decision period. The results show that subject P4 in the nonfatigue group fails to induce FRN components. The FRN component induction time of the remaining five subjects in the nonfatigue group was concentrated at 312 ms, with an average peak voltage of -6.34 μV. Six subjects in the fatigue group all induced FRN components, and the induction time was 320 ms after stimulation with an average peak voltage of −2.67 μV. In the fatigue state, the FRN component induction time of subjects was relatively late. The above research shows that, during the decision-making period of the nonfatigue group, the subjects were more focused and motivated to make decisions, and they were less willing to make gambling decisions. These fatigued subjects have inattention, weak decision-making motivation, and strong decision-making behaviour and have strong negative feedback on incorrect decision-making behaviour. Therefore, the FRN component can be considered the risk feedback signal of the wrong decision in the risk management and intelligent perception system of complex projects in coal mines. According to the test results, the peak

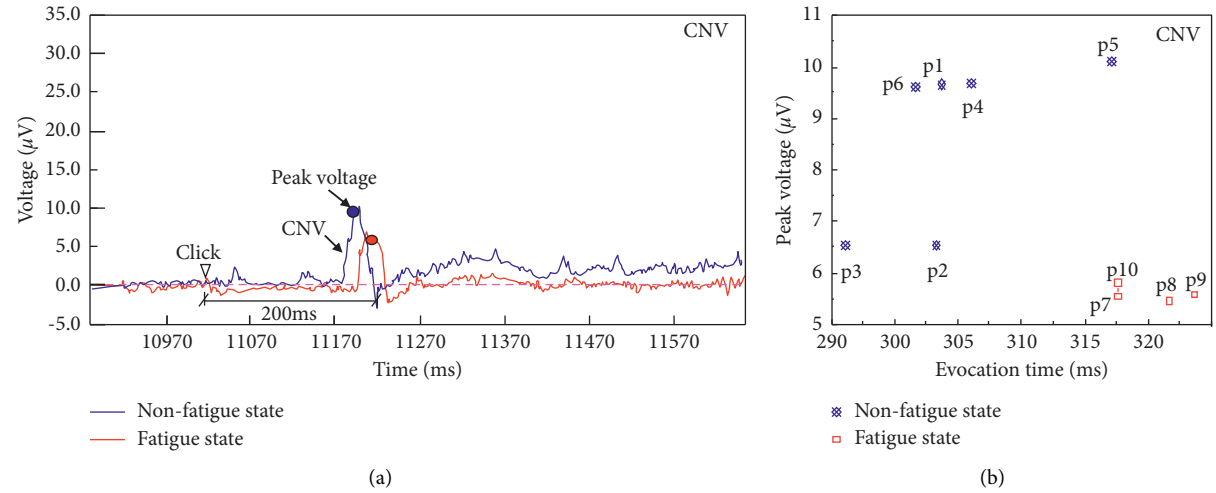


Figure 4: Changing characteristics and peak voltages of the CNV components of grassroots managers in coal mines during the risk perception period.

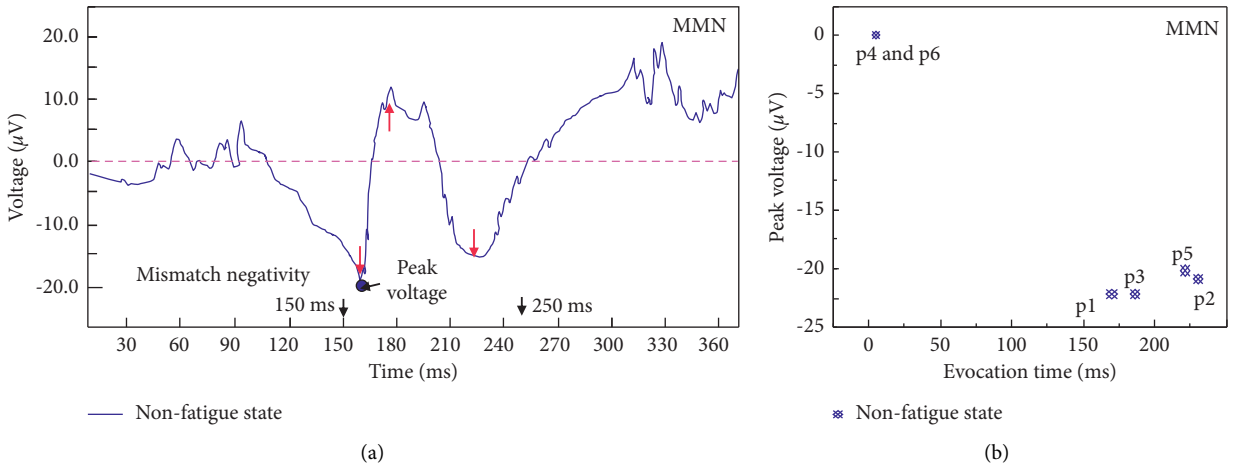


Figure 5: Changing characteristics and peak voltages of the MMN component of grassroots managers in coal mines during the risk perception period.

value of the FRN component voltage of coal mine managers during the risk decision period can be used as the critical value. This can verify the negative sentiment of the wrong decision-making behaviour of decision-makers.

3.3. Peak Voltage Variations of CNV and P300 during the Postperception Period. The feedback period after risk decision-making is the postperception process of decision-making behaviour. The periods of 50~60 s and 110~120 s in the experiment were the postperception period of the nonfatigue group participants and the postperception period of the fatigue group participants. The process is the process of thinking and feedback on risk decision awareness in the early stage of decision-making, and it is also the process of rethinking to improve cognitive risk perception ability. Therefore, in this experiment, the CNV component and the P300 component of the primary management subjects were the main analysis targets. During the postsensing period, the peak voltage and change process of the primary coal mine managers' components are shown in Figures 8 and 9. In the test results of the nonfatigue group, 5 subjects induced CNV components in the late stage, and the voltage peak was at 11.05–12.43 μV. The CNV induction time of 6 subjects in the fatigue group was generally later than that in the nonfatigue group. The CNV component of the nonfatigue group participants in the postrisk perception period had the characteristics of short latency and high potential peak, indicating the decision-making efficiency of the managers in the nonfatigue group was high during the decision recall process; also, the ERPs information response was relatively rapid, while the fatigue group ERPs information response was delayed and managers exhibited a lack of concentration.

The P300 component characteristics in the postperception period are the process of rethinking the decision-making behaviour deviation stimulus. Figure 9 shows the P300 component characteristics of the 12 subjects in the postperception period. Two subjects in the nonfatigue group

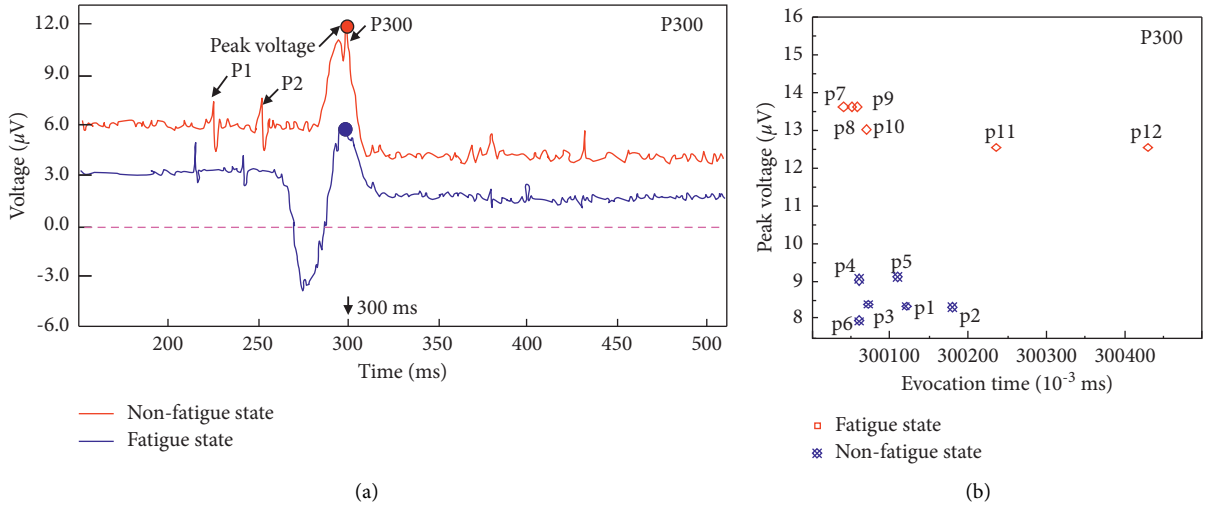


Figure 6: P300 component characteristics and peak voltage comparison of coal mine grassroots managers during the risk decision period.

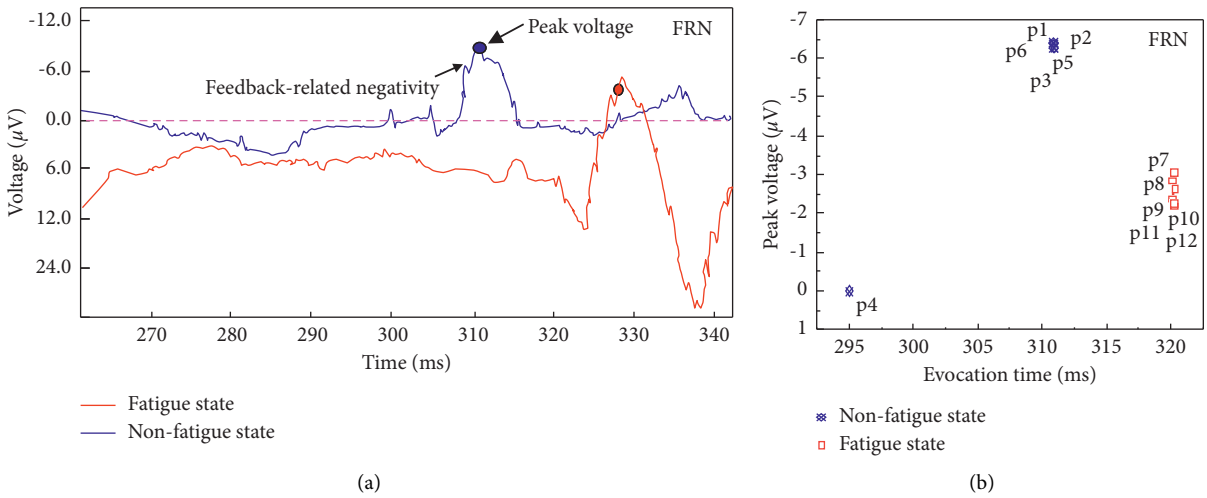


Figure 7: FRN component signal characteristics and peak voltage comparison of coal mine grassroots managers during the risk decision period.

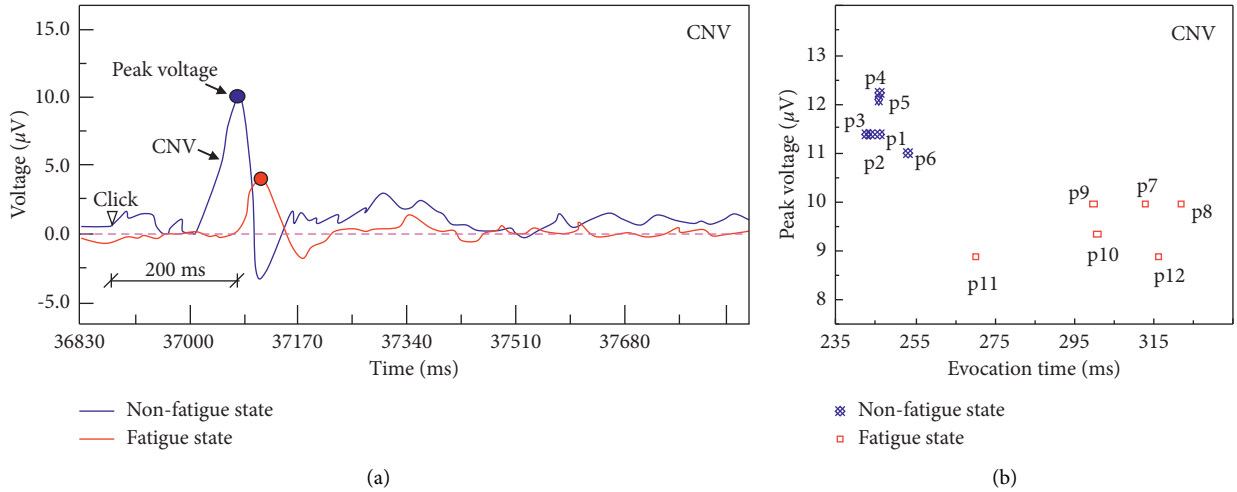


Figure 8: Comparison of CNV component signal characteristics and peak voltages of coal mine managers in the postsensing period.

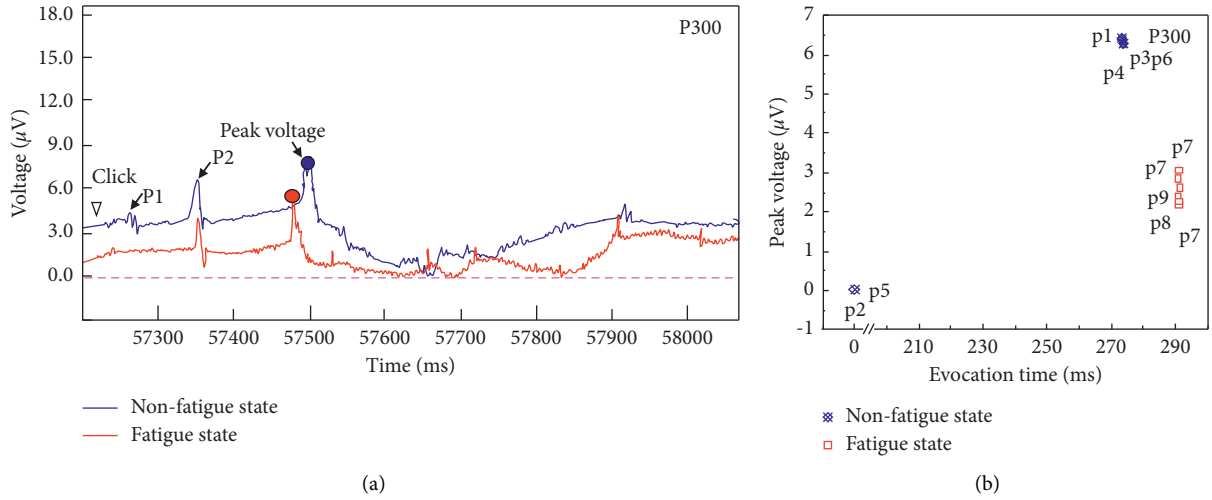


Figure 9: P300 component signal characteristics and peak voltage comparison of coal mine grassroots managers in the postperception period.

did not induce the P300 component (peak voltage is 0), and the remaining 4 subjects all obtained the P300 component and average peak voltage in approximately 277 ms. The P300 component of the fatigue group was later than that of the nonfatigue group, which shows that the grassroots coal mine managers significantly stimulated the decision-making behaviour deviation under the fatigue state.

4. Conclusions

ERP tests on the risk perception and decision-making behaviour of grassroots managers in coal mines were conducted in this paper. The characteristics of ERP components, such as CNV, P300, MMN, and FRN, of the participants were studied in three periods; also, the peak voltage and pattern of change in different ERP components were statistically analysed. The main conclusions are described below.

(1) The risk decision-making process of coal mine grassroots managers can be divided into perception periods, decision-making periods, and postperception periods according to the characteristics of ERP information. CNV and MMN components in the perception period can reflect the sensitivity of the decision-making subject's risk perception. P300 and FRN components in the decision period can reflect the auditory induction, anti-interference ability, and negative emotional feedback of the decision-making process. The CNV and P300 component characteristics in the postperception period can verify the decision-making behaviour deviation through the state of physiological indicators.

(2) There is a significant difference in the peak voltage of the risk decision ERPs of coal mine managers under fatigue and nonfatigue states. The results show that the P300 component is a key information indicator reflecting the deviation stimulus of decision-making behaviour during the decision-making period. The peak voltage of the P300 component under the fatigue state is generally greater than 10 μV. The peak value of the FRN component voltage in the fatigue state is generally less than −6.0 μV. Therefore, the P300 component and FRN component of the grassroots managers can be set as the respective thresholds to determine whether the decision-making behaviour is reasonable and effective.

(3) The CNV component and P300 component in the postperception period reflect the manager's process of rethinking the deviation in risk perception and decision-making behaviour. Therefore, the corresponding threshold can be set in the coal mine enterprise risk management and intelligent decision-making system platform. After the person completes the decision-making behaviour, the person's fatigue state can be judged by the induction time of these two ERP components, and the deviation in decision-making deviation can be verified.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare no conflicts of interest in connection with the paper submitted.

Acknowledgments

This study was funded by the Joint Funds of the National Natural Science Foundation of China (nos. U1904210, 61763045, and 81760235).

References

[1] Q. Qian and P. Lin, "Safety risk management of underground engineering in China: progress, challenges and strategies," *Journal of Rock Mechanics and Geotechnical Engineering*, vol. 8, no. 4, pp. 423–442, 2016.

[2] G. Wang, Y. Xu, and H. Ren, "Intelligent and ecological coal mining as well as clean utilization technology in China: review and prospects," *International Journal of Mining Science and Technology*, vol. 29, no. 2, pp. 161–169, 2019.

[3] K. Hong, "Typical underwater tunnels in the mainland of China and related tunneling technologies," *Engineering*, vol. 3, no. 6, pp. 871–879, 2017.

[4] J. Jia, "A technical review of hydro-project development in China," *Engineering*, vol. 2, no. 3, pp. 302–312, 2016.

[5] R. Mohanty, "Project selection by a multiple-criteria decision-making method: an example from a developing country," *International Journal of Project Management*, vol. 10, no. 1, pp. 31–38, 1992.

[6] Z. Xiong, H. Lu, M. Wang, Q. Qian, and X. Rong, "Research progress on safety risk management for large scale geotechnical engineering construction in China," *Rock and Soil Mechanics*, vol. 39, 2018.

[7] Y. Xian-ping, Y. A. N. Yong-sheng, and Z. Jin-suo, "Overview of researches on coal accident in China," *China Safety Science Journal*, vol. 24, no. 8, pp. 132–138, 2014.

[8] Z. Jiang, G. Wei, J. Wu, and X. Chen, "CPT-TODIM method for picture fuzzy multiple attribute group decision making and its application to food enterprise quality credit evaluation," *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 5, pp. 10115–10128, 2021.

[9] M. Zhao, G. Wei, C. Wei, J. Wu, and Y. Wei, "Extended CPT-TODIM method for interval-valued intuitionistic fuzzy MAGDM and its application to urban ecological risk assessment," *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 3, pp. 4091–4106, 2021.

[10] J. Lu, S. Zhang, J. Wu, and Y. Wei, "COPRAS method for multiple attribute group decision making under picture fuzzy environment and their application to green supplier selection," *Technological and Economic Development of Economy*, vol. 27, no. 2, pp. 369–385, 2021.

[11] C. Wei, J. Wu, Y. Guo, and G. Wei, "Green supplier selection based on CODAS method in probabilistic uncertain linguistic environment," *Technological and Economic Development of Economy*, vol. 27, no. 3, pp. 530–549, 2021.

[12] L. Wang, Y.-P. Cheng, and H.-Y. Liu, "An analysis of fatal gas accidents in Chinese coal mines," *Safety Science*, vol. 62, pp. 107–113, 2014.

[13] W. Yin, G. Fu, C. Yang, Z. Jiang, K. Zhu, and Y. Gao, "Fatal gas explosion accidents on Chinese coal mines and the characteristics of unsafe behaviors: 2000-2014," *Safety Science*, vol. 92, pp. 173–179, 2017.

[14] J. Zhang, D. Cliff, K. Xu, and G. You, "Focusing on the patterns and characteristics of extraordinarily severe gas explosion accidents in Chinese coal mines," *Process Safety and Environmental Protection*, vol. 117, pp. 390–398, 2018.

[15] J. Wang and Z. Huang, "The recent technological development of intelligent mining in China," *Engineering*, vol. 3, no. 4, pp. 439–444, 2017.

[16] K. Wang, C. Cai, M. Yamamoto, and T. Uchiyama, "Real-time brain activity measurement and signal processing system using highly sensitive MI sensor," *AIP Advances*, vol. 7, no. 5, Article ID 056635, 2017.

[17] G. Chen, X. Zhang, Y. Sun, and J. Zhang, "Emotion feature analysis and recognition based on reconstructed EEG sources," *IEEE Access*, vol. 8, pp. 11907–11916, 2020.

[18] X. Chen, B. Lu, and C.-G. Yan, "Reproducibility of R-fMRI metrics on the impact of different strategies for multiple comparison correction and sample sizes," *Human Brain Mapping*, vol. 39, no. 1, pp. 300–318, 2018.

[19] D. Tanner and J. G. Van Hell, "ERPs reveal individual differences in morphosyntactic processing," *Neuropsychologia*, vol. 56, pp. 289–301, 2014.

[20] T. D. Sambrook and J. Goslin, "A neural reward prediction error revealed by a meta-analysis of ERPs using great grand averages," *Psychological Bulletin*, vol. 141, no. 1, pp. 213–235, 2015.

[21] M. J. Larson, P. E. Clayson, and A. Clawson, "Making sense of all the conflict: a theoretical review and critique of conflict-related ERPs," *International Journal of Psychophysiology*, vol. 93, no. 3, pp. 283–297, 2014.

[22] K. C. Squires, E. Donchin, R. I. Herning, and G. McCarthy, "On the influence of task relevance and stimulus probability on event-related-potential components," *Electroencephalography and Clinical Neurophysiology*, vol. 42, no. 1, pp. 1–14, 1977.

[23] W. G. Walter, R. Cooper, V. J. Aldridge, W. C. Mccallum, and A. L. Winter, "Contingent negative variation: an electric sign of sensori-motor association and expectancy in the human brain," *Nature*, vol. 203, no. 4943, pp. 380–384, 1964.

[24] M. Sams, P. Paavilainen, K. Alho, and R. Näätänen, "Auditory frequency discrimination and event-related potentials," *Electroencephalography and Clinical Neurophysiology: Evoked Potentials Section*, vol. 62, no. 6, pp. 437–448, 1985.

[25] K. Sowndhararajan, M. Kim, P. Deepa, S. J. Park, and S. Kim, "Application of the P300 event-related potential in the diagnosis of epilepsy disorder: a review," *Scientia Pharmaceutica*, vol. 86, no. 2, 2018.

[26] Q. Wu, Y. Zhang, J. Liu, J. Sun, A. Cichocki, and F. Gao, "Regularized group sparse discriminant analysis for P300-based brain-computer interface," *International Journal of Neural Systems*, vol. 29, no. 6, Article ID 1950002, 2019.

[27] J. Jung, D. Morlet, B. Mercier, C. Confavreux, and C. Fischer, "Mismatch negativity (MMN) in multiple sclerosis: an event-related potentials study in 46 patients," *Clinical Neurophysiology*, vol. 117, no. 1, pp. 85–93, 2006.

[28] G. Rupp, C. Berka, A. H Meghdadi et al., "EEG-based neurocognitive metrics may predict simulated and on-road driving performance in older drivers," *Frontiers in Human Neuroscience*, vol. 12, pp. 1–14, 2019.

[29] H. Li, Y. Fan, and L. Shi, "Research of decisional behavior on miners' emergency response based on ERPs," in *Preceedings of the International Conference on Mechatronics, Materials, Chemistry and Computer Engineering ICMMCCE*, pp. 1796–1801, Atlantis press, Xi'an, China, 2015.

[30] W. Du, K. Zhang, and H. Sun, "Mechanical properties and energy development characteristics of impact-prone coal specimens under uniaxial cyclic loading," *AIP Advances*, vol. 9, no. 11, Article ID 115114, 2019.

[31] K. Gao, Z. Qi, J. Jia, S. Li, Z. Liu, and Z. Liu, "Investigation of coupled control of gas accumulation and spontaneous combustion in the goaf of coal mine," *AIP Advances*, vol. 10, no. 4, Article ID 45314, 2020.

Hindawi
Journal of Healthcare Engineering
Volume 2021, Article ID 3831453, 10 pages
https://doi.org/10.1155/2021/3831453

Research Article

A Fact-Finding Procedure Integrating Machine Learning and AHP Technique to Predict Delayed Diagnosis of Bladder Patients with Hematuria

Chia-Lun Lo,[1] Ya-Hui Yang,[1] and Hsiao-Ting Tseng[2]

[1]*Department of Health-Business Administration, Fooyin University, Kaohsiung 83102, Taiwan*
[2]*Department of Information Management, National Central University, Taoyuan 32001, Taiwan*

Correspondence should be addressed to Hsiao-Ting Tseng; appleapple928@gmail.com

Received 21 May 2021; Revised 23 July 2021; Accepted 12 August 2021; Published 21 August 2021

Academic Editor: Daniel Espino

Bladder cancer, the ninth most common cancer worldwide, requires fast diagnosis and treatment to prevent disease progression and improve patient survival. However, patients with bladder cancer often experience considerable delays in diagnosis. One reason for such delays is that hematuria, a major symptom of bladder cancer, has a high probability of also being a warning sign for urinary tract diseases. Another reason is that the sensitivity of the body parts affected by bladder cancer deters patients from undergoing cystoscopy and influences patients' "physician shopping" behavior. In this study, the analytic hierarchy process was used to determine critical variables influencing delayed diagnosis; moreover, the variables were used to construct models for predicting delayed diagnosis in patients with hematuria by using several machine learning techniques. Furthermore, the critical variables associated with delayed diagnosis of bladder cancer in patients with hematuria were evaluated using GainRatio technology. The study sample was selected from a population-based database. The model evaluation results indicated that the prediction model established using decision tree algorithms outperformed the other models. The critical risk factors for delayed diagnosis of bladder cancer were as follows: (1) cystoscopy performed 6 months after hematuria diagnosis and (2) physician shopping.

1. Introduction

Over the past 20 years, advances in technology and the popularization of social networks have facilitated the transmission of medical knowledge and information; this has enabled the public to acquire medical-related knowledge online and thus reduced medical information asymmetry [1]. Many patients actively seek additional medical opinions from multiple physicians to obtain more information on the medications or diagnoses they have obtained from the Internet. However, the increasing complexity of disease treatment strategies has increased the likelihood of physician misjudgment. According to a Harvard University study, 3.7% of hospitalized patients experienced medical injury, 27.6% experienced medical negligence, 69% experienced a human error, 2.6% experienced permanent disability, and 13.6% died due to medical errors [2]. Accordingly, reducing medical errors is a crucial issue in the medical field.

Most medical errors can be prevented, and the prevention of diagnostic delay can serve as a starting point to effectively reduce such errors [3–6]. Diagnostic delay not only causes poor prognosis but also increases medical expenses and affects quality of life. However, diagnostic delay is associated with the autonomy and professionalism of the physician; hence, this topic has rarely been explored. Several relevant medical studies have discussed major diseases that are difficult to diagnose, such as bladder cancer.

The incidence of bladder cancer has increased with urbanization and industrialization. Bladder cancer is the most common malignant neoplasm in the urinary system and the ninth most common cancer worldwide, causing more than 165,000 deaths each year [7]. The incidence of

bladder cancer is three times higher among men than among women [8, 9]. The most notable symptom of bladder cancer, namely, hematuria, is not unique to the disease, causing difficulty in diagnosis. In addition, because bladder cancer has a 50%–70% chance of recurrence and a 30%–40% chance of increased malignancy after recurrence, accurate diagnosis without delay is crucial [9].

In general, the main symptom of bladder cancer is hematuria. Therefore, the primary method of diagnosing bladder cancer entails visually checking for hematuria symptoms or performing a urine test for detecting blood reactions or cancer cells [10]. This method is simple and fast. However, hematuria is caused by various factors, and situations with unknown causes also exist. Moreover, the fact that hematuria is a painless symptom, that early symptoms of bladder cancer may be unapparent, and that the condition can be overlooked by the patient may result in misdiagnosis.

Hematuria is often discovered inadvertently, but it has a high probability of being a warning sign for urinary tract diseases. Hematuria may be caused by numerous other factors, such as lithiasis, urinary tract infection, and malignant tumors. However, even if patients are detected to have hematuria, a timely follow-up examination and tracking may not be conducted because of numerous influencing factors. For example, patients with bladder cancer may not have visible hematuria symptoms. Even if patients discover the symptoms, they may visit the nephrology or gynecology department rather than the urology department. In particular, female patients may initially visit the gynecology department (gynecological diseases can also cause hematuria). Bladder cancer may not be considered a possibility by gynecologists, and the lack of immediate judgment and further tests leads to diagnostic delays. When cancer is finally detected, the treatment may have been delayed beyond the optimal time. Such misdiagnoses can cause considerable physical and mental harm to patients and their family members.

The secondary method of diagnosing bladder cancer is cystoscopy. However, the examination of sensitive body parts, requirement of anesthesia, and physical discomfort after examination may deter patients from undergoing cystoscopy, eventually causing diagnostic delays. Studies have highlighted that delays can be caused by numerous factors, including limited awareness of guidelines, variations in recommendations by different guidelines, low perceived yield of cystoscopy in patients with hematuria, urgency, poor communication within and between clinical teams, and failures in patient adherence to prescribed plans [11–18]. Therefore, satisfactory noninvasive examination methods for the early diagnosis and accurate and timely monitoring of recurrence of bladder cancer are unavailable. Diagnostic delay implies a physician's negligent behavior and disregard for relevant patient symptoms, which may cause harm to patients. Nevertheless, if patients exhibit different symptoms for the same disease, physicians face higher risks and pressure during diagnosis. Thus, a predictive tool to facilitate early diagnosis would be valuable for clinical physicians.

The possible influence of diagnostic delay on survival and the risk factors for diagnostic delay in patients with cancer have been subjects of considerable interest and controversy for several years. Clinicians have traditionally been concerned with cancer-related research; nevertheless, most patients (especially those with cancer) face an unexpected or ambiguous situation and are generally eager to seek a second opinion from another physician to confirm their initial diagnosis. Even in such circumstances, patients and their family members face difficulties in choosing a treatment method and are faced with the following questions: "Why me?" or "Is the diagnosis real?" or "Is there a better treatment strategy?" Lower compliance with physicians' orders and chaotic "physician shopping" behaviors leads to the possibility of delayed diagnosis [19]. However, establishing a systematic support method for patients' decision-making regarding treatment choices is difficult because patients have a complicated mindset that creates difficulties in decision-making; in addition, patients who do not make follow-up visits to consult the same physician are difficult to locate. Therefore, to address the problem of delayed diagnosis, determining how to track patients' physician shopping is crucial.

Accurately identifying patients' visiting conditions can be difficult, even in a single hospital. Accordingly, in this study, we first used the analytic hierarchy process (AHP) to investigate the criteria and priorities for delayed diagnosis of bladder cancer. We subsequently selected cases of hematuria with delays in bladder cancer diagnosis from Taiwan's nationwide population-based database established using information obtained from the National Health Insurance (NHI) system, which contains complete medical care visit information. Because this database is an observational database, its data reflect real-world medical care behavior patterns. Delayed diagnosis is a sensitive topic for clinicians; hence, it is not easily identified in practice. Consequently, establishing an accurate risk model for delayed diagnosis of bladder cancer remains challenging. To address this challenge, we used artificial intelligence (AI) methods to identify the factors causing diagnostic delays and establish prediction models for delayed diagnosis in patients with bladder cancer.

AI technology and its applications are prominent research areas. In recent years, an increasing number of AI applications have been introduced in the medical field. AI programs can perform clinical diagnostic procedures and recommend treatment suggestions. Numerous successful AI applications have been reported [20–29]. These applications use AI methods such as decision trees (DTs), support vector machines (SVMs), multilayer perceptron (MLP), and logistic regression (LGR). Moreover, these applications can help physicians in analyzing and understanding complex clinical data and help improve diagnosis and medical quality. Therefore, we used three types of AI methods—namely, decision tree-based classifiers (C4.5 and random forest), functional-based methods (SVM and logistic regression), and an integrated method (MLP)—to construct prediction models for delayed diagnosis and compared their performance in identifying delayed bladder cancer diagnosis. The following model performance measures were assessed in this study: accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve (AUC).

2. Relevance of Bladder Cancer Review

Annually, more than 300,000 people are diagnosed as having bladder cancer worldwide. Bladder tumors rank as the seventh most common tumors and eighth most common cause of tumor-induced deaths. Bladder cancer is the most common malignant urinary tumors; 90% of the cases are transitional cell carcinoma, which renders early diagnosis difficult. Moreover, 5% of patients with bladder cancer develop metastasis by the time they are diagnosed. Bladder cancer typically occurs in people aged 50–70 years and is associated with the environment, smoking, and exposure to chemical substances. Studies have shown that 30%–50% of bladder cancers are caused by smoking and that smokers are 2–4 times more likely to develop bladder cancer compared with nonsmokers. Diagnosis and staging are performed through analysis of patients' medical history, urinalysis, cystoscopy, and urine cytology [8, 9].

Frequent urination, urgent urination, and pain during urination are initial symptoms of bladder cancer. The most common primary symptom in the early stage is hematuria, particularly manifested as repeated occurrence of blood in urine without pain during urination, which can be observed by either the naked eye or a microscope. Men and women should pay attention to the presence of blood in urine. Hematuria can be divided into initial, terminal, and total hematuria. However, hematuria is caused by various factors, such as urinary tract infection, urinary tract stones, urinary tract cancer, benign prostatic hyperplasia, kidney diseases, coagulation disorders, and medication; therefore, the diagnosis of bladder cancer is difficult [9, 10].

Bladder cancer is diagnosed using methods and tools such as routine urine tests, intravenous pyelography (IVP), ultrasound examination, urine cytology, and cystoscopy. However, IVP and ultrasound examination cannot detect small tumors or foreign bodies. The cell staining technique employed in urine cytology is Papanicolaou staining, which exhibits high specificity but is insensitive to urothelial carcinoma with low malignant potential, resulting in a high rate of false negative results. Therefore, in clinical practice, other tests are conducted in conjunction with this method. Alternatively, pathologists can determine the presence of cancer cells according to the cell types and characteristics. Even if the presence of a tumor is confirmed by X-ray and ultrasound tests, these tests cannot reveal whether the tumor is benign or malignant. Bladder cancer is not detected in the first test in many cases; this is either typical or because cancer cannot yet be detected using equipment. Therefore, patients with hematuria who are determined as exhibiting no bladder cancer in their tests are generally recommended to have a follow-up examination within 3–6 months. Nevertheless, in clinical practice, the cause of repeat hematuria in some patients cannot be identified [8, 9].

Therefore, cystoscopy is generally the primary means of examining bladder cancer in clinical practice. Cystoscopy is conducted to detect overall changes in the bladder, ureteral orifice, prostate gland, and urinary tract. Other symptoms in patients, such as difficulty urinating, narrowing of the prostate or urinary tract, or hematuria of unknown causes, can be examined through cystoscopy for further information. When the presence of a tumor is confirmed, its appearance and characteristics can be observed with the naked eye through cystoscopy. Subsequently, a biopsy can be performed to determine the stage of cancer and facilitate accurate diagnosis.

However, because cystoscopy is expensive and involves sensitive, invasive, and uncomfortable procedures, it is often avoided by patients and is usually not the first choice for physicians, resulting in delayed diagnosis of bladder cancer. Occasionally, physicians are misled by the self-reported symptoms of patients in their diagnostic decision-making. They occasionally use patients' vague self-reported symptoms as a clue for diagnosis; lack a holistic, systematic, or comprehensive analysis; or fail to consider a scientific basis as necessary for a diagnostic decision. These factors often lead to time delays and hence to delayed diagnosis.

3. Materials and Methods

3.1. Database and Ethical Consideration. For tracking patient physician shopping, population-based health data were used in this study. This case-control study used data retrieved from Taiwan's National Health Insurance Research Database (NHIRD) for the period of 2005–2013. Data in the NHIRD are derived from medical claims records of the Taiwan NHI program and include original medical claims and registration files for 1,000,000 enrollees of the NHI program. Taiwan's National Health Research Institutes randomly selected these 1,000,000 enrollees from all enrollees listed in the 2005 Registry of Beneficiaries ($n = 23.72$ million). The NHIRD is one of the largest and most comprehensive population-based datasets in the world. Previous studies have demonstrated the high validity of the data derived from the NHI program. In our empirical analysis, we used a large dataset sourced from Taiwan's NHIRD for the years 2005–2013.

The Institutional Review Board of Fooyin University Hospital approved this study (protocol number: FYH-IRB-106-06-06-02-A). Written consent from the study patients was not obtained because the NHIRD consists of deidentified secondary data for research purposes, and the Institutional Review Board of Fooyin University Hospital issued a formal written waiver regarding the need for consent.

3.2. Structure of the Decision-Making Model. To resolve the complications and confusing alternatives, the AHP was used to decide which input variable would be suitable for use in the models. The AHP is a research methodology developed by Thomas L. Saaty in 1971. It is mainly applied to uncertain situations and decision-making problems with multiple evaluation criteria [30]. The basic concept of AHP theory is the pairwise comparison derived from the mechanism of idea formation in the human brain. The human brain can easily make adequate judgments in a pairwise comparison but tends to become muddled in the case of multiple alternatives. In the AHP, the opinions of experts and decision-makers are collected, and through consistency verification, the experts' comparison results on each dimension are

presented logically and coherently. A decision-making problem is decomposed into a hierarchical decision-making process, and each element that constitutes the hierarchy is compared in a pairwise manner to set the priority scale. The following steps are involved in the AHP: (1) define the problem and determine the goal and (2) construct the hierarchical structure. The top level of the construction hierarchy is the goal of the problem, the middle level is the criterion, and the bottom level is the alternative. Therefore, we used the AHP to systematize the research questions. Subsequently, we applied an AHP expert questionnaire to collect and analyze the opinions of various experts in order to determine the factors influencing physician shopping in patients and consequently resulting in delayed diagnosis. After the assessment, 10 influential factors identified by the experts were included in the models for predicting the possibility of delayed diagnosis of bladder cancer in patients with hematuria. According to the experts' opinions, age of the patient and physician, seniority of the physician, hospital level and location, physician shopping, and cystoscopy record within 6 months were possible influential factors.

In general, AHP structures for two decision steps are similar. The AHP structure in this study had three levels. The first level pertained to the delayed diagnosis of bladder cancer in patients with hematuria, the second level involved the classification of the criteria, and the third level involved the subcriteria. The consistency index for the criteria of the AHP structure was 0.084, and the random index (RI) was 1.49. Accordingly, the consistency ratio (CR) was calculated as 0.056. According to Saaty's suggestion, when the CR is ≤ 0.10, the matrix is consistent and the experts' opinions are acceptable. The ranks of the variables are presented in Table 1.

Table 1: The importance of variable ranking by AHP.

Rank	Dimension
1	Characteristics of patients
2	Cystoscopy after hematuria record for half year
3	Characteristics of hospital
4	Visiting behavior
5	Characteristics of physicians

3.3. Study Population Selection and Controls. Regarding our study sample, we selected patients who were newly diagnosed as having hematuria (International Classification of Diseases, Ninth Revision, Clinical Modification (ICD9-CM) code 599.7) or blood in the stool (ICD9-CM code 578.1) between January 1, 2005, and December 31, 2013, in the NHIRD. For this retrospective case-control study, cases were included without any recruitment restrictions on age, sex, ethnicity, or cancer stage.

In the preprocessing stage, to identify patients with an actual diagnosis of a malignant neoplasm of the bladder, we selected those who were diagnosed twice as having a malignant neoplasm of the bladder (ICD9-CM code 188.9) from CD (i.e., "ambulatory care expenditures by visits") files in the NHIRD. Thus, we identified 607 patients. Patients who were suspected to have a malignant neoplasm of the bladder were not included. In addition, two patients with unknown date of birth and without sex information and 58 patients without any cystoscopy record were excluded. Moreover, 14 patients with bleeding in the digestive tract (ICD9-CM code 578.1) were removed to avoid confusion and to maintain the quality of the samples. The final sample comprised 535 patients with consistent information. For each patient, data such as physician shopping (including visits to surgery, gynecology, nephrology, Chinese medicine, and gastroenterology departments), frequency of visit, age and sex of the patient and physician, and region and accreditation of the hospital visited were collected from the database as predictors to determine whether the diagnosis was delayed. The outcome variable was delayed diagnosis; patients who were diagnosed as having a malignant neoplasm of the bladder at least 3 months after hematuria was recorded were defined as having delayed diagnosis and were thus assigned to the delayed diagnosis group ($n = 210$); otherwise, they were defined as not having delayed diagnosis and were thus assigned to the nondelayed diagnosis group ($n = 325$). 41% of patients were operated or administered other aggressive treatment 3 months after diagnosis of bladder cancer and were considered as having delayed diagnosis. The sample selection process is shown in Figure 1.

3.4. Classification Techniques. After feature selection according to experts' opinions in the AHP and the data retrieval process, prediction models for delayed diagnosis of bladder cancer were established using maximum likelihood (ML) technology. We applied several well-known ML-based single classification techniques, namely, DT, SVM, logistic regression (LGR), and MLP neural network with back-propagation classifiers.

A DT algorithm applies classification and induction methods to generate a tree-like decision structure that is learned by the inductive method of the known examples of each class. A DT is a useful ML model; it can process complex data and is not affected by linear regression and interactions between independent variables. DT nodes consist of branches and leaves; the decision node indicates the test to be performed. To classify the input data, each DT node is a predicate, and each predicate can determine whether the variable is greater than, equal to, or less than a prespecified value. During data analysis, if the selected data variable belongs to categorical data, it is called a classification tree; if the selected data variable belongs to the continuous pattern, it is called a regression tree. Data classification using a DT algorithm is a two-step process. The first step involves a learning process, wherein the training data are analyzed by the DT algorithm to create a model that is presented as classification rules or a DT. The next step involves determining the accuracy of the classification rules or DT. If the accuracy is acceptable, rules can be reused to classify new data in the same scenario of the practical field [29]. C4.5 and random forest are the two most commonly used DT-based learning techniques. A DT is similar to the clinical decision-making process of a physician. After a DT is modelized, it

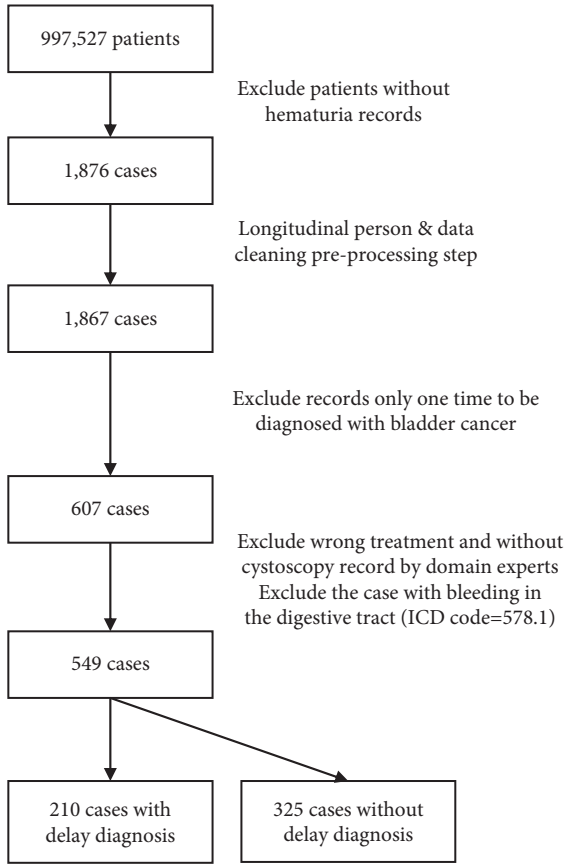


Figure 1: Flowchart showing derivation of the study sample.

can provide a suitable method for explaining the problem at the hand. Therefore, we selected a DT algorithm in this study.

Vanik's research team at the AT&T Laboratory developed the SVM algorithm [31], which is a controlled classification algorithm based on statistical learning techniques. It devises a computationally efficient method of learning to separate hyperplanes in a high-dimensional feature space based on statistical learning theory. The SVM algorithm first projects the training instances into a high-dimensional vector space and then determines the separating hyperplane exhibiting a maximal margin (i.e., the distance between the separating hyperplane and the closest sample). To reduce the generalization error of the classifier, the SVM algorithm determines an optimal hyperplane or a set of optimal hyperplanes (i.e., a hyperplane with a maximal margin) to separate training instances into two or more classes. This hyperplane is then used to determine the class label of unknown instances. The SVM algorithm can be divided into two types: linear and nonlinear. The operating principle of the SVM is based on the principle of predicting the most appropriate decision function that separates two classes in the most appropriate way to achieve the best classification effect. Its major function is to process the classification problems encountered during the data mining process. LGR is also a widely used statistical technique for forecasting the value of a binary or ordinal variable. LGR predicts the probability of occurrence of an event by fitting data into a logistic function, thereby allowing inputs with any values to be transformed and confined to a value between 0 and 1. Each regression coefficient represents the corresponding variable's degree of contribution. A positive regression coefficient increases the probability of the output, whereas a negative regression coefficient decreases the probability of the output. Both SVM and LGR algorithms are functional-based learning techniques [32].

MLP is a mathematical model that imitates the functionality of biological neural systems [33, 34]. It consists of an input layer, an output layer, and one or more hidden layers. Neurons are organized and fully connected between two adjacent layers, and the output layer is responsible for producing estimated outputs. Each layer receives inputs from the previous layer and converts them into a higher level of combinations by using combination and transfer functions. A high learning rate may result in achieving the minimum error quickly but may lead to an MLP model periodically fluctuating around the solution without being able to converge. However, a low learning rate may result in a local minimum or a long time to converge [35].

3.5. Performance Measures. In this study, WEKA 37.3, an open-source ML program, was used to establish DT (C4.5, RF), SVM, LGR, and MLP prediction models for classification. Because the predictive performance of classifiers can be considerably influenced by the parameter settings, the CV parameter selection metalearner module implemented in WEKA was used to optimize the predictive performance of the selected classifiers. The specific values of the various parameters were combined for each classifier; subsequently, the optimal parameter setting was automatically determined based on the best prediction results obtained using the validation strategy of our study. The specific parameter range and values selected for each classifier are listed in Table 2. Previous studies have reported that several classification algorithms implemented in conjunction with AdaBoost achieved higher classification accuracy than did individual base classifiers [36–39]. In the present study, Adaptive Boost (or AdaBoost in short), a prominent classifier ensemble, was employed to further enhance the predictive power of the classifiers [40].

Tenfold cross-validation was applied to evaluate the predictive performance of the classifiers. Tenfold cross-validation is a practical statistical method in which sample data are divided into smaller subsets. The idea is to randomly divide the sample into 10 nonoverlapping subsamples, with the categories in each subsample similar to the original sample. Nine subsamples were used for training to establish the models, and the remaining (one) subsample was used for testing. The same procedure was performed 10 times (for all 10 combinations). For performance evaluation, the final accuracy was obtained by comparing the number of incorrect results with the original number of entries. The predictive performance of each classifier was measured by evaluating the accuracy, sensitivity, specificity, and AUC.

Table 2: Parameter setting in Weka.

Method	Parameters	Range	Value setting
C4.5	Confidence factor	0.1–0.5	0.25
	Minimum number of instances per leaf	2–20	20
Random forest	Number of trees	5–250	10
	Number of attributes to be used in random selection	2–8	4
Support vector machines	Kernel		PolyKernel
Multilayer perceptron	Number of hidden nodes	5–10	7
	Learning rate	0.1–0.6	0.3
	Momentum factor	0–0.9	0.2
	Maximum number of epochs	300–900	500

The most used empirical measure, accuracy, does not distinguish between the number of correct labels of different classes:

TP, true positives: number of examples predicted positive that are actually positive

FP, false positives: number of examples predicted positive that are actually negative

TN, true negatives: number of examples predicted negative that are actually negative

FN, false negatives: number of examples predicted negative that are actually positive

Accuracy: it refers to the total number of records that are correctly classified by the classifier. The accuracy of a classifier is defined as the percentage of test set tuples that are correctly classified by the model [31].

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}. \quad (1)$$

Sensitivity: refers to the true positive rate that means the proportion of positive tuples that were correctly identified [31].

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}}. \quad (2)$$

Specificity: refers to the rate at which a test or diagnostic method sets a correct (i.e., negative) diagnosis for a patient who is not ill [31].

$$\text{Specificity} = \frac{\text{TN}}{\text{FP} + \text{TN}}. \quad (3)$$

In general, if the predictive accuracy of the proposed model is perfect, its AUC is nearly 1. If the AUC is between 0.8 and 0.9, then the model has high predictive accuracy. If its AUC is between 0.7 and 0.8, then the proposed model is acceptable. We compared the pros and cons of each prediction model according to accuracy, sensitivity, specificity, and AUC and then selected the most appropriate model for predicting the course of disease possible delay diagnosis of bladder cancer in patients with hematuria.

4. Results and Discussion

4.1. Patient Selection. Table 2 presents the variables and descriptive statistics of the delayed and nondelayed diagnosis groups. The delayed diagnosis group comprised 210 patients with bladder cancer, and the nondelayed diagnosis group had 325 patients, of whom 67%-68% were men. The median age at enrollment was approximately 67 years in both groups (interquartile ranges: 31–93 and 16–98 years). Physician shopping behaviors in both groups were for surgery, gynecology, Chinese medicine, gastroenterology, and nephrology. In the delayed diagnosis group, over 97% of the patients underwent cystoscopy after hematuria was detected after a delay of 6 months, which means that delayed cystoscopy caused significantly delayed diagnosis. The demographics of patients and physicians and other variables are presented in Table 3.

4.2. Experimental Results for Different Models. Next, we combined the two groups (i.e., patients and physicians) to establish prediction models of delayed diagnosed in order to assist physicians in identifying patients at a high risk of bladder cancer. According to the law of large numbers, some useful instances in the study data may not be chosen by the classifier; therefore, we applied the models 30 times to construct datasets (by seed = 1–30) and averaged the evaluation results. For each generated dataset, tenfold cross-validation was applied in all the experimental evaluations. To evaluate the performance of our model, parameters such as accuracy, sensitivity, specificity, and AUC were considered.

The evaluation results for the six classifiers (i.e., C4.5, RT, random forest, SVM, logistic regression, and MLP) in the prediction models are presented in Table 4. For ease of explanation, this table presents only the mean and standard deviation of the 30 generated datasets. Summaries of other statistics are available upon request from the authors.

First, the average predictive accuracy rates of the C4.5 and RT classifiers were 0.859 and 0.879. These classifiers outperformed the SVM (0.746), LGR (0.788), and MLP (0.742) classifiers. The average sensitivity levels of the C4.5, RT, SVM, LGR, and MLP classifiers were 0.843, 0.875, 0.752, 0.799, and 0.720, respectively; their specificity levels were 0.858, 0.872, 0.769, 0.802, and 0.709, respectively; and their average AUC values were 0.871, 0.942, 0.705, 0.854, and 0.775, respectively. Apart from the SVM classifier, all the classifiers exhibited excellent predictive performance because the AUC values were >0.7. The tree-based classifiers (i.e., C4.5 and RF) outperformed the functional-based classifiers (SVM and LGR) and MLP in terms of prediction accuracy. Moreover, the average sensitivity and specificity levels of the C4.5 and RF classifiers were higher than those of

Table 3: Characteristics of bladder cancer patients with delay diagnosed and control subjects.

Variables	Value	Delay diagnosed group ($n = 210$)	Nondelayed diagnosed group ($n = 325$)
Age	Patient	67.3 (31–93)	67.4 (16–98)
	Physician	54.7 (47–63)	56 (43–63)
Seniority (physician)		11.7 (4–20)	13 (5–20
Gender (patient)*	Male	142 (67.6)	227 (68.8)
	Female	68 (32.4)	98 (30.2)
Gender (physician)*	Male	201 (95.7)	313 (96.3)
	Female	9(4)	12 (3.7)
Hospital level*	Medical center	60 (28.6)	113 (34.8)
	Regional hospital	81 (38.6)	135 (41.5)
	District hospital	36 (17.1)	43 (13.2)
	Clinic	33 (15.7)	32 (10.5)
Visit behavior*	Surgery	52 (24.7)	19 (5.8)
	Gynecology	38 (18.1)	13 (4)
	Chinese medicine	48 (22.3)	3 (0.9)
	Gastroenterology	61 (29.0)	6 (1.8)
	Nephrology	53 (25.2)	30 (9.2)
Cystoscopy after hematuria record for half year		205 (97.6)	120 (36.9)
Visit times	Surgery	1.69	0.03
	Gynecology	1.69	0.05
	Chinese medicine	1.73	0.01
	Gastroenterology	0	0.02
	Nephrology	0	0.18
Location (level of urbanization)*			
	City	82 (39.0)	114 (44.3)
	Commuting zone	54 (25.7)	80 (24.6)
	Towns and semidense areas	10 (4.8)	12 (3.7)
	Rural areas	64 (30.5)	89 (27.4)

*n (%), the others are $\mu(\sigma)$.

Table 4: Performance results of classifiers.

Classifier	Algorithms	Accuracy, μ/σ	Sensitivity, μ/σ	Specificity, μ/σ	AUC, μ/σ
Without AdaBoost	C4.5	0.859/0.014	0.843/0.003	0.858/0.016	0.871/0.042
	RF	0.879/0.071	0.875/0.045	0.8720.081	0.942/0.048
	SVM	0.746/0.007	0.752/0.005	0.769/0.040	0.705/0.008
	LGR	0.788/0.011	0.799/0.013	0.802/0.021	0.854/0.011
	MLP	0.742/0.079	0.720/0.048	0.709/0.127	0.775/0.104
With AdaBoost	C4.5	0.856/0.088	0.825/0.109	0.856/0.088	0.915/0.064
	RF	0.881/0.066	0.857/0.062	0.881/0.066	0.943/0.045
	SVM	0.743/0.011	0.722/0.008	0.743/0.001	0.762/0.010
	LGR	0.791/0.002	0.802/0.025	0.791/0.003	0.828/0.008
	MLP	0.751/0.088	0.791/0.029	0.752/0.087	0.786/0.111

the SVM, LGR, and MLP classifiers. As expected, the DT-based classifiers outperformed the other classifiers in terms of the five selected performance indicators. Therefore, we can conclude that the predictive performance of the DT-based classifier was superior to that of the functional-based and MLP classifiers; in particular, the RF classifier yielded superior accuracy compared with the other classifiers, and the difference was statistically significant ($P < 0.05$).

Second, our comparative results revealed an improvement in prediction ability—measured using all performance indicators in Table 4—when the classifiers were supplemented by AdaBoost. For example, the average accuracy, sensitivity, specificity, and AUC of the C4.5, RF, and MLP classifiers supplemented with AdaBoost were higher than those observed when these classifiers were implemented without AdaBoost; nevertheless, the difference not statistically significant. However, the SVM classifier supplemented with AdaBoost exhibited relatively low performance in one of the indicators than it did when implemented without AdaBoost; this means that the predictive performance was not stable. The evaluation results for the LGR classifier revealed a similar trend.

4.3. Variable Important Evaluation. In addition to comparing the performance of the classifiers in determining delayed diagnosis of bladder cancer, we determined their predictive performance. We further evaluated the importance ranking of each selected variable of patients' physician shopping by using ML technology. The GainRatioAttributeEval model in WEKA

was used to evaluate the importance level of all variables selected in this study. In the model, the gain ratio index was computed for each input variable; this index helped us in determining the relative importance of the variables.

As given in Table 5, we adjusted the input variables for our models (e.g., DT-based classifiers) for improving the model performance continually by calculating the gain ratios. We observed that hematuria was the most crucial variable influencing delayed diagnosis. However, undergoing cystoscopy 6 months after symptom appearance was determined to be the main reason for delayed diagnosis of bladder cancer, followed by variables related to patients' physician shopping. Many patients visit different departments to obtain more information about hematuria treatment strategies, contributing to a delayed diagnosis. Therefore, the mental stress associated with undergoing cystoscopy in patients with hematuria should be assessed, and patients should be encouraged to accept cystoscopy, which can help prevent the "physician shopping" behavior.

The ML classifiers' evaluation of the variables was compared with the experts' opinions about the variable ranking list in Table 1. The clinical experts believed that if a patient with some symptoms undergoes cystoscopy, the physician should ask the patient to stay in the hospital for further treatment. However, because patients in such situations increasingly pursue a second opinion, physicians should endeavor to reduce the possibility of delayed diagnosis. The characteristics of the patient and their physician shopping behavior are crucial variables, and follow-up action and related strategies should be emphasized.

Objective ML classifiers were used in this study to determine data-driven criteria for delayed diagnosis in order to reduce subjective bias in humans. The results revealed that clinical staff should pay more attention to patients' physician shopping behavior. Delayed diagnosis usually occurs when clinical staff passively wait for patients to visit clinics. Therefore, clinical staff should actively contact patients to improve their awareness about treatment strategies for their diseases.

Table 5: Variable importance ranking by GainRatioAttributeEval.

Rank	Variable	GainRatio
1	Cystoscopy after hematuria record for half year	0.3030443
2	Gastroenterology visiting times	0.122365
3	Gastroenterology visiting	0.122365
4	Chinese medicine visiting	0.103643
5	Chinese medicine visiting times	0.0636328
6	Surgery visiting times	0.073374
7	Gynecology visiting times	0.063343
8	Nephrology visiting times	0.035427
9	Hospital level	0.007911
10	Location	0.004625

5. Discussion

This study developed ML models for predicting delayed diagnosis of bladder cancer following hematuria symptoms. The anamnesis, full urine, and cystoscopy examinations of 591 patients who visited a urology clinic in Taiwan were collected from the NHIRD. Five AI analysis classifiers, namely, C4.5, RF, SVM, LGR, and MLP algorithms, which are frequently utilized in medical diagnosis systems, were used to create classification models based on the dataset. All models were tested using tenfold cross-validation, and their classification performance levels were compared and evaluated.

The results reveal that delayed diagnosis was related to sex, patients' physician shopping (whether patients had visited the gastroenterology department and the number of patients' visits to gynecology and gastroenterology departments), physician seniority, and whether cystoscopy was performed. These results supported the study hypotheses. However, information on the patients' physiological factors was not available in the data. Thus, although we identified a relationship between the patients' physician shopping and delayed diagnosis, we could not exclude the possibility of bias or presence of additional factors causing delayed diagnosis.

For any disease, the optimal strategy for reducing the chance of a poor patient prognosis is early diagnosis. A critical objective of preventive health care is to promote early diagnosis based on the standard procedure of checking the medical history and symptoms. However, malignant neoplasms of the bladder are in a rather sensitive area of the body, and examination is uncomfortable. In addition, the medical field currently contains numerous specialized fields. Consequently, patients may visit the wrong department when seeking medical treatment. Delayed diagnosis may be the result of a combination of all these factors. Therefore, to avoid delayed diagnosis and unnecessary medical expenditures, physicians from departments other than the urology department should consider the possibility of malignant neoplasms of the bladder when examining patients with hematuria.

To reduce delayed diagnosis, coordination and communication across departments of the healthcare system are essential. When first-line medical service providers doubt the existence of other possibilities in their patients' conditions, they must think from a comprehensive perspective to reduce misjudgments, which is difficult. However, because disease treatment strategies and comorbidities have become increasingly complex, relying only on physicians' decisions is insufficient.

In recent years, with the rise of AI, scholars and practitioners in the medical field have increasingly used big data to improve diagnostic accuracy. Numerous factors might be associated with a malignant neoplasm of the bladder. However, because of the sensitivity of the location of the neoplasm, the necessity of adopting an invasive procedure, and the pain and pressure a patient experiences during the procedure, patients commonly avoid active treatment. Therefore, this study analyzed patients' behaviors associated with consulting physicians to determine the risk of delayed diagnosis. The findings can help in improving the quality of diagnosis of malignant neoplasms of the bladder. In the future, psychological data could be introduced into hybrid AI algorithms to improve prediction accuracy.

6. Conclusions

The main symptom of malignant neoplasms of the bladder is hematuria. However, hematuria is caused by various factors. Patients with hematuria often visit a department other than the urology department, leading to delayed diagnosis. Psychological factors such as fear are also common causes of delayed examination and thus delayed diagnosis. Therefore, incorporating technology to identify factors related to the diagnosis of patients with malignant neoplasms of the bladder would be valuable. In this study, supervised ML classifiers were applied to establish prediction models for determining the behavioral characteristics of patients that could lead to delayed diagnosis in order to reduce the chance of delayed diagnosis.

This study has several limitations. The prediction models of delayed diagnosis were established using medical data from the NHIRD. However, because of the limitation related to the value-added data analysis of NHI expenditure application data, further analysis of patients' psychological factors was not possible. Moreover, because of the lack of socioeconomic data, the patients could not be grouped to analyze behavioral differences among socioeconomic clusters to identify possible indirect effects. Therefore, we could only assume that causal relationships existed between the factors analyzed and delayed diagnosis. Finally, the behavioral patterns of patients cannot be the main reference for making a diagnosis.

In summary, the problem of delayed diagnosis is sensitive, and the lack of discussion in previous studies is probably due to the pressure to avoid medical disputes. However, value-based payment is an increasingly common trend in healthcare insurance policy. The accuracy of medical diagnosis must be actively improved to maintain high medical treatment quality. Therefore, future studies can consider including more factors to establish models for predicting delayed diagnosis or consider integrating prediction algorithms into a computerized physician order entry system to create a practical clinical decision support system with warning functions.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

This research was supported in part by the Ministry of Science and Technology of the Taiwan (MOST-106-2410-H-242-001, MOST 109-2637-H-242-001, and MOST 110-2636-H-008-003-).

References

[1] C. Angst, R. Agarwal, G. Gao, J. Khuntia, and J. S. McCullough, "Information technology and voluntary quality disclosure by hospitals," *Decision Support Systems*, vol. 57, pp. 367–375, 2014.

[2] L. L. Leape, T. A. Brennan, N. Laird et al., "Incidence of adverse events and negligence in hospitalized patients. results of the harvard medical practice Study I," *New England Journal of Medicine*, vol. 324, pp. 370–376, 1991.

[3] M. O. Aaland and K. Smith, "Delayed diagnosis in a rural trauma center," *Surgery*, vol. 120, no. 4, pp. 774–779, 1996.

[4] R. A. Furnival, G. A. Woodward, and J. E. Schunk, "Delayed diagnosis of injury in pediatric trauma," *Pediatrics*, vol. 98, pp. 56–62, 1996.

[5] A. Brooks, B. Holroyd, and B. Riley, "Missed injury in major trauma patients," *Injury*, vol. 35, no. 4, pp. 407–410, 2004.

[6] C. B. Thomson and I. Greaves, "Missed injury and the tertiary trauma survey," *Injury*, vol. 39, no. 1, pp. 107–114, 2008.

[7] J. Ferlay, I. Soerjomataram, R. Dikshit et al., "Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012," *International Journal of Cancer*, vol. 136, no. 5, pp. E359–E386, 2015.

[8] A. B. Alexandroff, A. M. Jackson, M. A. O'Donnell, and K. James, "BCG immunotherapy of bladder cancer: 20 years on," *The Lancet*, vol. 353, no. 9165, pp. 1689–1694, 1999.

[9] E. L. Wynder and R. Goldsmith, "The epidemiology of bladder cancer: a second look," *Cancer*, vol. 40, no. 3, pp. 1246–1268, 1977.

[10] R. Davis, J. S. Jones, D. A Barocas et al., "Diagnosis, evaluation and follow-up of asymptomatic microhematuria (AMH) in adults: AUA guideline," *The Journal of Urology*, vol. 188, no. s6, pp. 2473–2481, 2012.

[11] A. M. Nieder, Y. Lotan, G. R. Nuss et al., "Are patients with hematuria appropriately referred to urology? a multi-institutional questionnaire based survey," *Urologic Oncology: Seminars and Original Investigations*, vol. 28, no. 5, pp. 500–503, 2010.

[12] S. Tarkan, C. Plaisant, B. Shneiderman, and A. Z. Hettinger, "Reducing missed laboratory results: defining temporal responsibility, generating user interfaces for test process tracking, and retrospective analyses to identify problems," *AMIA AnnuSymp Proc*, vol. 20111391 pages, 2011.

[13] E. H. Chen and T. Bodenheimer, "Improving population health through team-based panel management," *Archives of Internal Medicine*, vol. 171, no. 17, pp. 1558-1559, 2011.

[14] D. R. Murphy, B. Reis, D. F. Sittig, and H. Singh, "Notifications received by primary care practitioners in electronic health records: a taxonomy and time analysis," *The American Journal of Medicine*, vol. 125, no. 2, pp. 209–217, 2012.

[15] M. Smith, D. Murphy, A Laxmisan et al., "Developing software to "track and catch" missed follow-up of abnormal test results in a complex sociotechnical environment," *Applied Clinical Informatics*, vol. 4, no. 3, pp. 359–375, 2013.

[16] M. Schmidt-Hansen, S. Berendse, and W. Hamilton, "The association between symptoms and bladder or renal tract cancer in primary care: a systematic review," *British Journal of General Practice*, vol. 65, no. 640, pp. e769–e775, 2015.

[17] D. R. Murphy, A. N. D. Meyer, E. Russo, D. F. Sittig, L. Wei, and H. Singh, "The burden of inbox notifications in commercial electronic health records," *JAMA Internal Medicine*, vol. 176, no. 4, pp. 559-560, 2016.

[18] M. Nielsen and A. Qaseem, "Hematuria as a marker of occult urinary tract cancer: advice for high-value care from the American college of physicians," *Annals of Internal Medicine*, vol. 164, no. 7, pp. 488–497, 2016.

[19] M. Demissie, B. Lindtjorn, and Y. Berhane, "Patient and health service delay in the diagnosis of pulmonary tuberculosis in Ethiopia," *BMC Public Health*, vol. 2, no. 1, p. 23, 2002.

[20] W. G. Baxt, "Application of artificial neural networks to clinical medicine," *The Lancet*, vol. 346, no. 8983, pp. 1135–1138, 1995.

[21] M. F. Abbod, D. G. Von Keyserlingk, D. A. Linkens, and M. Mahfouf, "Survey of utilisation of fuzzy technology in medicine and healthcare," *Fuzzy Sets and Systems*, vol. 120, no. 2, pp. 331–349, 2001.

[22] A. A. N. Ramesh, J. Kambhampati, P. Monson, and P. J. Drew, "Artificial intelligence in medicine," *Annals of the Royal College of Surgeons of England*, vol. 86, no. 5, pp. 334–338, 2004.

[23] D. Delen and R. Sharda, "Artificial neural networks in decision support systems," in *Handbook on Decision Support Systems 1*, pp. 557–580, Springer, Berlin, Heidelberg, 2008.

[24] I. Saritas, I. A. Ozkan, and I. U. Sert, "Prognosis of prostate cancer by artificial neural networks," *Expert Systems with Applications*, vol. 37, no. 9, pp. 6646–6650, 2010.

[25] M. Dagli and I. Saritas, "Using artificial neural network for the prediction of anemia seen in behcet disease," *Energy Education Science and Technology Part A-Energy Science and Research*, vol. 28, pp. 1085–1092, 2012.

[26] I. Saritas, "Prediction of breast cancer using artificial neural networks," *Journal of Medical Systems*, vol. 36, no. 5, pp. 2901–2907, 2012.

[27] F. Amato, A. López, E. M. Peña-Méndez, P. Vaňhara, A. Hampl, and J. Havel, "Artificial neural networks in medical diagnosis," *Journal of Applied Biomedicine*, vol. 11, no. 2, pp. 47–58, 2013.

[28] A. Arvin, "Introduction to medical applications of artificial intelligence," in *Medical Applications of Artificial Intelligence*, pp. 1–8, CRC Press, Boca Raton, Florida, 2013.

[29] E. Kaya, M. E. Aktan, A. T. Koru, and E. Akdoan, "Diagnosis of anemia in children via artificial neural network," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 3, no. 1, pp. 24–27, 2015.

[30] T. L. Saaty, "Decision making - the analytic hierarchy and network processes (AHP/ANP)," *Journal of Systems Science and Systems Engineering*, vol. 13, no. 1, pp. 1–35, 2004.

[31] V. Vapnik, *Statistical Learning Theory*, Wiley, New York, 1998.

[32] A. Ozkanl, M. Koklu, and I. U. Sert, "Diagnosis of urinary tract infection based on artificial intelligence methods," *Computer Methods and Programs in Biomedicine*, vol. 166, pp. 51–59, 2008.

[33] V. Vapnik, *The Nature of Statistical Learning Theory*, Springer-Verlag, Berlin, Germany, 1995.

[34] Y.-H. Hu, C.-L. Lo, and S.-P. Shih, "Developing early warning systems to predict students online learning performance," *Computers in Human Behavior*, vol. 36, pp. 469–478, 2014.

[35] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, no. 6088, pp. 533–536, 1986.

[36] K. M. Iftekharuddin, S. Ahmed, and J. Hossen, "Multiresolution texture models for brain tumor segmentation in MRI," *Conference Proceedings IEEE Engineering in Medicine and Biology Society*, vol. 2011, Article ID 6091766, 6988 pages, 2011.

[37] F. Scalzo, R. Hamilton, S. Asgari, S. Kim, and X. Hu, "Intracranial hypertension prediction using extremely randomized decision trees," *Medical Engineering & Physics*, vol. 34, no. 8, pp. 1058–1065, 2012.

[38] R. Stollhoff, W. Sauerbrei, and M. Schumacher, "An experimental evaluation of boosting methods for classification," *Methods of Information in Medicine*, vol. 49, no. 3, pp. 219–229, 2010.

[39] D. B. Parker, "Learning-logic," MIT Technical Report TR-47, Center for Computational Research in Economics and Management Science, Cambridge, MA, USA, 1985.

[40] M. Galar, A. Fernandez, E. Barrenechea, H. Bustince, and F. Herrera, "A review on ensembles for the class imbalance problem: bagging-, boosting-, and hybrid-based approaches," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 42, no. 4, pp. 463–484, 2012.

Hindawi
Journal of Healthcare Engineering
Volume 2021, Article ID 7799100, 10 pages
https://doi.org/10.1155/2021/7799100

Research Article

Micro Expression Recognition via Dual-Stream Spatiotemporal Attention Network

Yan Wang,[1] Yikun Huang,[2] Can Liu,[3] Xiaoying Gu,[1] Dandan Yang,[1] Shuopeng Wang,[1] and Bo Zhang[1]

[1]*College of Information Engineering, Tianjin University of Commerce, Tianjin 300134, China*
[2]*Concord University College of Fujian Normal University, Fuzhou, Fujian 350117, China*
[3]*School of Artificial Intelligence, Hebei University of Technology, Tianjin 300401, China*

Correspondence should be addressed to Bo Zhang; 18002058162@163.com

Received 6 May 2021; Accepted 8 August 2021; Published 18 August 2021

Academic Editor: Chi-Hua Chen

Microexpression can manifest the real mood of humans, which has been widely concerned in clinical diagnosis and depression analysis. To solve the problem of missing discriminative spatiotemporal features in a small data set caused by the short duration and subtle movement changes of microexpression, we present a dual-stream spatiotemporal attention network (DSTAN) that integrates dual-stream spatiotemporal network and attention mechanism to capture the deformation features and spatiotemporal features of microexpression in the case of small samples. The Spatiotemporal networks in DSTAN are based on two lightweight networks, namely, the spatiotemporal appearance network (STAN) learning the appearance features from the microexpression sequences and the spatiotemporal motion network (STMN) learning the motion features from optical flow sequences. To focus on the discriminative motion areas of microexpression, we construct a novel attention mechanism for the spatial model of STAN and STMN, including a multiscale kernel spatial attention mechanism and global dual-pool channel attention mechanism. To obtain the importance of each frame in the microexpression sequence, we design a temporal attention mechanism for the temporal model of STAN and STMN to form spatiotemporal appearance network-attention (STAN-A) and spatiotemporal motion network-attention (STMN-A), which can adaptively perform dynamic feature refinement. Finally, the feature concatenate-SVM method is used to integrate STAN-A and STMN-A to a novel network, DSTAN. The extensive experiments on three small spontaneous microexpression data sets of SMIC, CASME, and CASME II demonstrate the proposed DSTAN can effectively cope with the recognition of microexpressions.

1. Introduction

Microexpression is a kind of spontaneous facial expression that can reveal the real emotion that people try to hide. The duration of microexpression is short, only lasting 1/25 s~1/5 s [1]. And the muscle movement caused by microexpression only appears in a small area of the face, which limits the performance of recognizing microexpression to a certain extent. In recent years, a large number of automatic recognition methods have emerged, which greatly improve the application feasibility of microexpression. At present, microexpression recognition has a wide application prospect in the police interrogation, clinical diagnosis, depression analysis, and other fields [2–5].

In the microexpression recognition procedures, feature extraction is the critical step and researchers strive to seek the reprehensive methods. LBP-TOP (local binary pattern with three orthogonal planes) [6] is a typical texture feature-based method for microexpression recognition and taken as the baseline of handcraft methods. Due to its shortcomings of sensitivity and sparse sampling, there are many improved methods, such as LBP-SIP (local binary pattern with six intersection points) [7], STLBP-IP (spatial-temporal local binary pattern with integral projection) [8], STCLQP

(spatial-temporal completed local quantized patterns) [9], and LCBP (local cube binary pattern) [10], to enhance the robustness of the features.

Furthermore, the optical flow-based methods exploit the changes of the pixels in the time domain and the correlations between the sequence frames to mine the correspondence in adjacent frames. The classical approaches include MDMO (main directional mean optical flow feature) [11], which can identify microexpression by locating the main direction of the face block and calculating the average optical flow of the face block. MDMO is characterized by simple calculation and small feature dimensions, but it is easy to miss the low-level manifold structure. Based on MDMO, researchers have proposed various improved methods, such as FDM (Facial Dynamics Map) [12], Sparse MDMO [13], and Bi-WOOF (Bi-Weighted Oriented Optical Flow) [14], to extract the motion features of microexpression more effectively.

Although the texture-based and motion-based methods solved the recognition problems to a certain extent, the features got by these handcraft methods need artificial design and it is difficult to automatically extract discriminative information using these methods. Latterly, deep learning approaches, such as CNNs (convolutional neural networks) and LSTM (long short-term memory), have shown their powerful abilities in many fields. These methods avoid the tedious handcraft feature design and can automatically capture the subtle changes of microexpressions. Kim et al. [15] adopted the two-step model CNN-LSTM to recognize microexpressions. They utilize CNNs to extract the spatial features of a frame and then put the features into LSTM to learn the temporal information of the microexpression sequence. Li et al. [16] proposed 3D-FCNN to extract the deep spatiotemporal features to identify the microexpression. Khor et al. [17] presented the method of ELRCN-TE.

They fused the original sequence, optical flow sequence, and optical strain sequence of microexpression and adopted VGG-LSTM to extract the spatiotemporal features of microexpression. Xia et al. [18] designed STRCN by fusing the appearance and geometry features to extend the connectivity of convolutional networks in the time domain. However, these methods do not consider the complementarity of high-level, low-level networks and the contribution of various image pixels to recognize microexpressions.

Since the attention mechanism has been successfully applied to many tasks, Chu et al. [19] combined CNN and multicontext attention to form an end-to-end framework for human pose estimation. Zhang et al. [20] used progressive attention to guide RNN for detecting salient objects. Due to the subtlety and short duration of microexpression, Yang et al. [21] proposed MERTA by utilizing three attention mechanisms to construct feature maps. Nevertheless, these attention-based methods mostly handle multilevel features without distinction, ignoring the differences between high-level and low-level network features and the intensity difference between the microexpression frames.

Inspired by these works, this paper constructs a dual-stream spatiotemporal network, DSTAN, by using STAN-A (STAN with attention mechanism) to extract appearance features and STMN-A (STMN with attention mechanism) to get motion features of microexpression sequences, respectively. Considering the small size of the microexpression data set and the low motion intensity of microexpression, this paper designs two lightweight networks, STAN and STMN, to extract subtle microexpression features. Besides, to focus on the key regions of microexpression, we introduce two attention mechanisms to the spatial model of STAN and STMN: the multiscale kernel spatial attention mechanism is applied to get the detailed low-level features, and the global dual-pool channel attention mechanism is applied to obtain the high-level features. Given the importance of different frames, the temporal attention mechanism is employed in the temporal model of STAN and STMN, so that the model can learn more representative features. Finally, the feature concatenate-SVM method is used to integrate the dual-stream networks STAN-A and STMN-A, which integrate the two spatiotemporal networks STAN and STMN and attention mechanisms to realize the task of microexpression recognition.

2. Proposed Method

The overall framework of DSTAN is shown in Figure 1. STAN-A extracts spatiotemporal appearance features from the original microexpression sequence, and STMN-A extracts the spatiotemporal motion features from the optical flow sequence to describe the subtle motion changes of the microexpression. STAN and STMN are two networks that extract appearance features and motion features of microexpression, respectively. The multiscale kernel spatial attention and global dual-pool channel attention are introduced into the spatial model of STAN and STMN to extract the refined spatial features of the microexpression. Then, the spatial features are input into the temporal model of STAN and STMN to get the spatiotemporal features of the microexpression. Finally, the STAN with attention (STAN-A) and the STMN with attention (STMN-A) are integrated by feature concatenate-SVM to obtain the predicted category of the microexpression.

2.1. Image Preprocessing. First, we carry on face detection for each frame and locate the feature points. Based on these key points, the face region is blocked. To remove the impacts of head movements on recognition, we conduct facial alignment on the images to eliminate the differences of faces and sequences in the expressionless state. Furthermore, the aligned frames are normalized in the spatial domain, that is, to maintain the size of the frame uniformity. Generally, the length of the microexpression sequence is different, but the deep learning network usually needs a fixed length of the input dimension in the training stage. Therefore, it is necessary to normalize the microexpression sequence in the time domain. We use the TIM (temporal interpolation model) [22] to handle the original sequence to a fixed number of frames, and the sequence with a fixed length is taken as the input of STAN-A. The optical flow information between two adjacent frames is calculated from the original sequence, and the

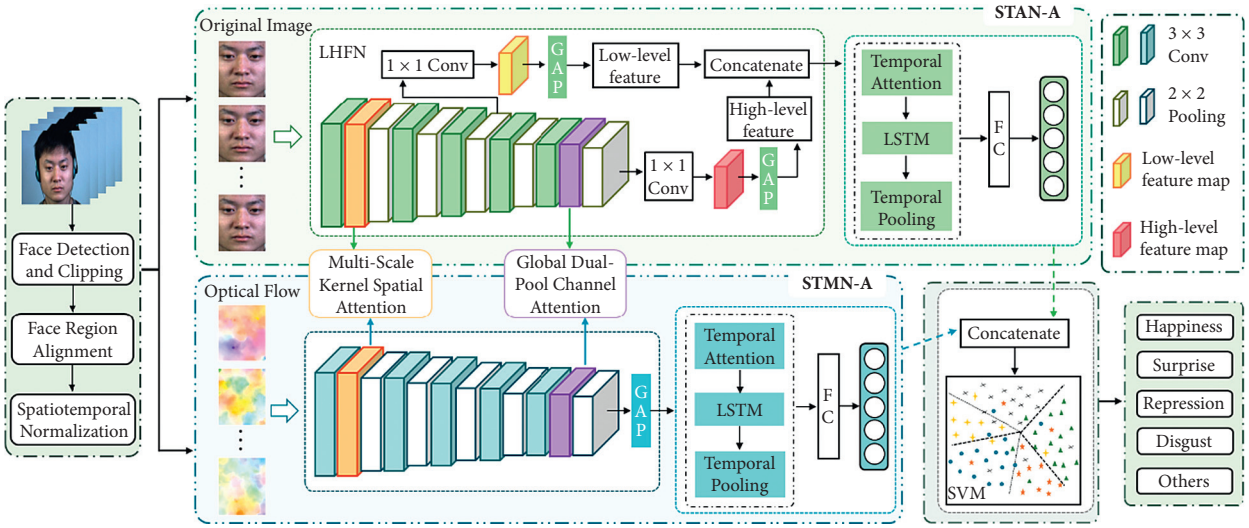


Figure 1: An illustration of the proposed DSTAN.

obtained fixed-length optical flow sequence is set as the input of STMN-A.

2.2. STMN. The spatial model of STMN is designed as a shallow network. Firstly, we use a kernel of 3×3 to conduct the convolution operations on the input optical flow sequence to extract local features. To avoid losing the edge information, zero padding is performed before convolution operation, and batch normalization [23] is used after convolution to accelerate the training speed of the model. We utilize ReLU as an activation function to enhance the nonlinear expression ability of the network. Each convolution layer is connected with a maximum pooling layer, and downsampling is performed under the condition of a 2×2 neighborhood and 2-step size. The local microexpression features are obtained after five times of convolution and pooling operation. Then, we adopt the GAP (global average pooling) to integrate these features and obtain the spatial features.

The temporal model of STMN is to obtain dynamic information between frames. The spatial feature vector describing the motion information of the microexpression obtained by the spatial model is input into the single-layer LSTM to learn the correlation between frames and obtain the feature vector v_i of each microexpression sequence. Then, the feature vectors are aggregated through a temporal average pooling operation to obtain the spatiotemporal feature f of the whole sequence:

$$f = \frac{1}{t} \sum_{i=1}^{t} v_i, \tag{1}$$

where $i = 1,2, \ldots, t$, and t represents the number of frames. Finally, the fully connected layer is applied to map the feature space to the label space through linear transformation, and softmax is used to map the output to (0, 1) to obtain the category of microexpression.

2.3. STAN. Considering that the features in different levels are complementary, we design the spatial model for STAN fusing high-level and low-level features, and the model can learn both deep semantic and low texture features, as the LHFN (low high feature fusion network) module in Figure 1 shows. In CNNs, different convolution layers learn different features; the third convolution layers can learn the low-level texture features [24], which play an important role in recognizing microexpression, so we fuse it with the last layer, which can learn the high-level semantic features to realize the high-level and low-level networks. The implementation of LHFN is based on the convolution calculation with a convolution kernel of 1×1 on the high-level feature map and low-level feature map to introduce more nonlinear relations. Then, we apply the GAP layer to obtain global low-level texture features and high-level semantic features. Finally, the high-level and low-level features are fused by a feature concatenate mode to obtain the spatial features describing the appearance information of each frame. The temporal model of STAN is the same as the temporal structure of STMN.

2.4. Attention Mechanism for Spatial Model. The existing microexpression recognition approaches handle the contribution of each pixel in the image or frame equally. However, the microexpression mainly appears in specific parts of the face, such as eyes, eyebrows, and mouth. According to the feature differences of the low-level and high-level networks, we introduce a novel attention mechanism to the spatial domain model, which is composed of a multiscale kernel spatial attention mechanism and a global dual-pool channel attention mechanism.

We introduce the multiscale kernel spatial attention to the low-level network and the global dual-pool channel attention to the high-level network to make the network focus on these significant motion areas.

2.4.1. Multiscale Kernel Spatial Attention Mechanism. The low-level network extracts the texture, edge, contour, and other low-level visual features of microexpression, and this information has almost no difference in different

channels. Therefore, we apply the multiscale kernel spatial attention to the low-level network to effectively distinguish each pixel in the spatial domain. The implementation process is shown in Figure 2, and the calculating processes are as follows:

Given the low-level feature map $F^l \in \mathbb{R}^{C\times H\times W}$, C is the number of feature channels, and H and W are the height and width of the feature map, respectively. The first step is to conduct convolution operations $\mathrm{Conv}_m^{n\times n}$ on feature map F^l by the convolution kernel n of 1×1, 3×3, and 5×5 to extract multiscale feature, and the spatial feature matrix of different scales $S_1 \in \mathbb{R}^{1\times H\times W}$, $S_2 \in \mathbb{R}^{1\times H\times W}$, and $S_3 \in \mathbb{R}^{1\times H\times W}$ are obtained:

$$\begin{aligned} S_1 &= \mathrm{Conv}_1^{1\times 1}\left(F^l\right), \\ S_2 &= \mathrm{Conv}_2^{3\times 3}\left(F^l\right), \\ S_3 &= \mathrm{Conv}_3^{5\times 5}\left(F^l\right). \end{aligned} \tag{2}$$

Then, we fuse S_1, S_2, and S_3 by concatenate mode and conduct convolution operation by the convolution kernel of 1×1 to obtain spatial features. Afterward, we obtain the weight $SA \in \mathbb{R}^{1\times H\times W}$ of spatial attention by normalizing as below:

$$SA = \sigma\left(\mathrm{Conv}_4^{1\times 1}\left(S_1, S_2, S_3\right)\right). \tag{3}$$

Finally, we multiply SA with F^l and get the refined spatial attention feature map $F_S \in \mathbb{R}^{C\times H\times W}$:

$$F_S = SA \otimes F^l, \tag{4}$$

where $\otimes$ denotes the matrix multiplied by elements, and σ is the sigmoid function.

2.4.2. Global Dual-Pooling Channel Attention Mechanism. A high-level network extracts high-level semantic feature information, and different feature channels have different responses to different semantic features [25]. Max-pooling can preserve more texture information, average pooling can retain more local information, and utilizing maximum pooling and average pooling at the same time can greatly improve the network's presentation capabilities [26]. Consequently, we present the global dual-pooling channel attention mechanism to the high-level network, which combines the max-pooling operation with average pooling operation effectively. This kind of attention mechanism automatically obtains the contribution of each feature channel. Through this attention mechanism, the effective features are enhanced while the features of little matter are suppressed. The global dual-pooling channel attention mechanism is shown in Figure 3.

The calculation process is as follows: given the input high-level feature graph $F^h \in \mathbb{R}^{C\times H\times W}$, C is the number of feature channels, and H and W are the height and width of the feature map, respectively. Firstly, we conduct GAP and GMP (global max-pooling) operation on F^h to aggregate the spatial information of the feature map and obtain the

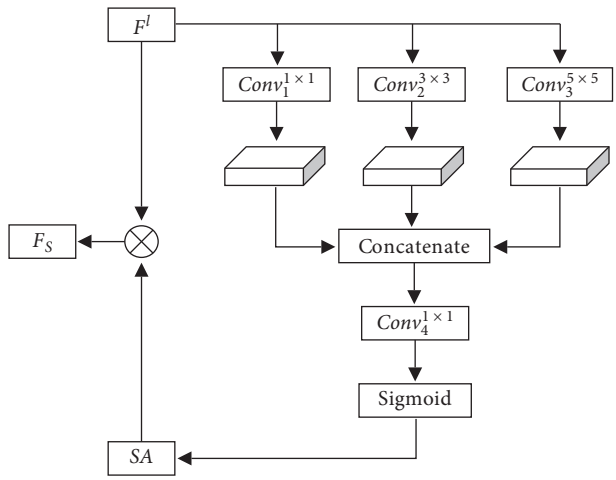


Figure 2: Multiscale kernel spatial attention mechanism.

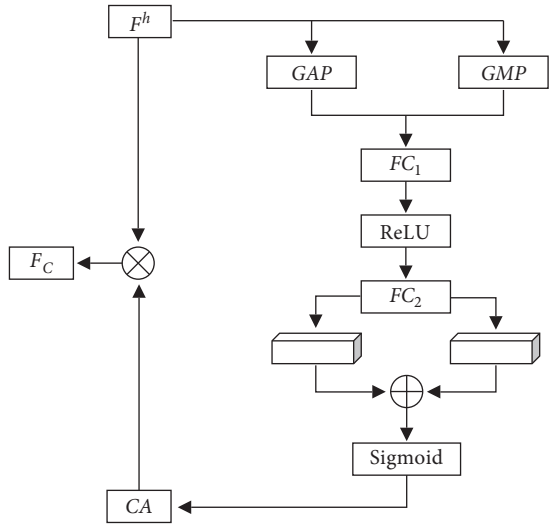


Figure 3: Global dual-pool channel attention mechanism.

global average pooling feature vector $F_C^{GMP} \in \mathbb{R}^{C\times 1\times 1}$ and global max-pooling feature vector $F_C^{GAP} \in \mathbb{R}^{C\times 1\times 1}$. Then, we use two consecutive full-connection layers FC_1 and FC_2 to fine-tune the parameters adaptively to learn the dependence and correlation of different channels. To reduce the model complexity, we set the number of units in FC_1 as C/r, where r is the compression ratio, and the number of units in FC_2 as C. Through a full-connection layer, we can get two-channel feature vectors $C_1 \in\in \mathbb{R}^{C\times 1\times 1}$ and $C_2 \in\in \mathbb{R}^{C\times 1\times 1}$:

$$\begin{aligned} C_1 &= FC_2\left(\delta\left(FC_1\left(GAP\left(F^h\right)\right)\right)\right) \\ &= w_2\left(\delta\left(w_1\left(F_C^{GAP}\right) + b_1\right)\right) + b_2, \\ C_2 &= FC_2\left(\delta\left(FC_1\left(GMP\left(F^h\right)\right)\right)\right) \\ &= w_2\left(\delta\left(w_1\left(F_C^{GMP}\right) + b_1\right)\right) + b_2. \end{aligned} \tag{5}$$

Next, we merge C_1 and C_2 through element summation. The weight $CA \in\in \mathbb{R}^{C\times 1\times 1}$ of channel attention can be got by normalizing as below:

$$CA = \sigma(C_1 \oplus C_2). \tag{6}$$

Finally, we obtain the refined feature map $F_C \in \mathbb{R}^{C\times H\times W}$ of channel attention by multiplying CA and F^h:

$$F_C = CA \otimes F^h, \tag{7}$$

where δ denotes the ReLU activation function, σ denotes the sigmoid function, $\otimes$ indicates that vectors are added, w_1 and w_2 are the weight of FC_1 and FC_2, and b_1 and b_2 are the offsets, respectively.

2.5. Attention Mechanism for Temporal Model. This paper introduces another attention mechanism into the temporal model so that the model can learn automatically and distinguish the important frames in the microexpression sequence. The attention mechanism for the temporal model is shown in Figure 4. The feature vector of each frame obtained by the spatial model is input into the model, and an attention weight representing the importance of the frame is calculated. Specifically, for the spatial feature vector S_i' corresponding to the *i-th* frame, we use the sigmoid function to obtain the attention weight r_i for each frame and then perform a weighted operation on the obtained attention weight to get the feature vector S_i' of each frame.

$$S_i' = S_i r_i, \tag{8}$$

where $i = 1, 2, \ldots, t$, and t represents the number of frames. The weighted spatial feature vector of each frame S_i' is input into the temporal models of STAN and STMN, respectively, to obtain the refined spatiotemporal appearance features and spatiotemporal motion features of a sequence.

2.6. Model Integration. We integrate STAN-A and STMN-A by the feature concatenate-SVM method. Firstly, the SVC (support vector classification) is initialized with a linear kernel function to define the classifier. Then, the linear multivariate classifier is trained by the microexpression data in the training set, as shown in Equation (9):

$$\left\| X:\ f(p_i, q_i), Y \right\| \longrightarrow SVM, \tag{9}$$

where p_i and q_i are the outputs of STAN-A and STMN-A, respectively, and $f(p_i, q_i)$ is the cascaded results, X denotes the features of the classifier, and Y represents the feature label.

3. Results and Discussion

3.1. Data Sets. To evaluate the performance of the proposed framework, we conduct experiments on three spontaneous microexpression data sets: SMIC (Spontaneous Micro-Expression Database) [27] CASME(Chinese Academy of Sciences Micro-Expression) [28], and CASME II [29]. SMIC contains three categories of microexpressions: positive (51), negative (70), and surprise (43), and a total of 164 samples from 16 subjects. In CASME, 172 samples of 19 subjects' microexpression sequences are collected and divided into 4 categories, that is, tense (70), expression (38), distinct (44), and surprise (20). There are 246 samples of 26 subjects on CASME II. There are divided into 5 categories: happiness (32), surprise (25), expression (27), distinct (63), and others (99).

3.2. Parameter Setting and Evaluation Criterion. We use the TIM model to normalize the length of the sequence to 9 frames, and the size of each frame is set to 224×224. In the global dual-pooling channel attention, the compression ratio r is set as 16. We adopt the cross-entropy loss function and Adam optimizer to train the model and set the batch size as 32.

To get a stable and reliable model, we conduct experiments on three microexpression data sets to evaluate the performance of the algorithm by using the LOSOCV, that is, all samples of a subject are taken as testing sets, and the rest are used as training sets.

We utilize accuracy, *F1-score*, precision, and recall as the evaluation criterion to evaluate the proposed model. Accuracy is the ratio of the correct predicting sample number to the total sample number:

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN}. \tag{10}$$

F1-Score is the harmonic average of accuracy. *F1-Score, Precision*, and *Recall* can be calculated as follows [30]:

$$\begin{aligned} F1 - \text{Score} &= 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}, \\ \text{Precision} &= \frac{TP}{TP + FP}, \\ \text{Recall} &= \frac{TP}{TP + FN}, \end{aligned} \tag{11}$$

where TP (true positive) indicates the number of samples that positive is predicted as positive; FP (false positive) indicates the number of samples that negative is predicted as positive; TN (true negative) is the number of samples that negative is predicted as negative; and FN (false negative) is the number of samples that positive is predicted as negative.

3.3. Experimental Analysis. In this section, ablation experiments and performance verification are performed on the proposed DSTAN framework, and comparative experiments are conducted with state-of-the-arts.

3.3.1. Comparison of Single Network with Dual-Stream Network. To verify the effectiveness of the dual-stream network, we compare the single-stream networks STAN-A and STMN-A with the dual-stream network DSTAN. Table 1 shows the comparison results on three data sets. It can be seen that the performance of the dual-stream network DSTAN is better than that of STAN-A and STMN-A on three data sets. Specifically, compared with STAN-A and STMN-A, the accuracy of DSTAN is increased by 9.15% and 12.2% and *F1-score* is increased by 10.64% and 12.84% on SMIC. On CASME, the accuracy of DSTAN is increased by

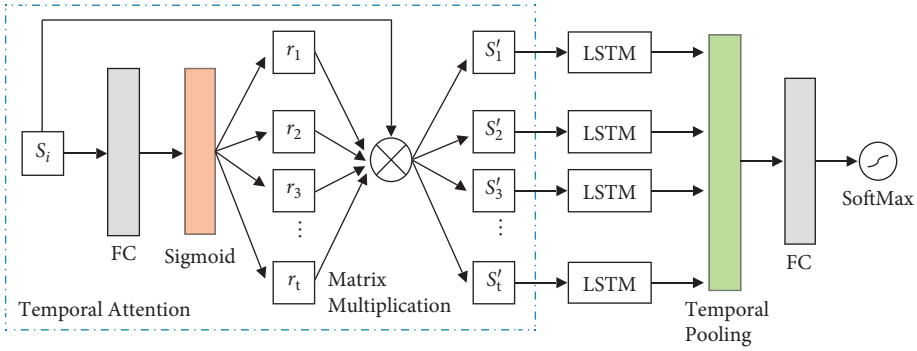


Figure 4: Attention mechanism for the temporal model.

Table 1: Comparison of single-stream with dual-stream network (F1 : *F1-score*).

Method	SMIC		CASME		CASME II	
	Accuracy	F1	Accuracy	F1	Accuracy	F1
STAN-A	0.683	0.672	0.669	0.627	0.602	0.559
STMN-A	0.652	0.649	0.651	0.612	0.638	0.607
DSTAN	0.774	0.778	0.779	0.752	0.752	0.728

11.05% and 12.79% and *F1-score* is increased by 12.47% and 13.97%. On CASME II, the accuracy of DSTAN is increased by 15.04% and 11.38% and *F1-score* is increased by 16.91% and 12.18%. The results show that the dual-stream network DSTAN outperforms the single-stream network, which can verify that DSTAN can make the model learn more discriminative features and improve the overall recognition performance.

Furthermore, we compare the recognition performance of each emotion on three data sets, as shown in Figure 5. On SMIC, as shown in Figure 5(a), STAN-A gets a higher recognition rate for "positive" but a poor result for "surprise".

However, STMN-A has a good performance for "surprise" and low accuracy for "positive". On CASME, as shown in Figure 5(b), STAN-A has a good performance on "disgust", but it behaves poorly on "repression" and "surprise". STMN-A has a good performance on these two emotions, but the "disgust" recognition result is lower. On CASME II, as can be seen from Figure 5(c), STAN-A and STMN-A are also complementary. Especially, the performance of STMN-A is lower than that of STAN-A in recognizing "disgust", but DSTAN gets an ideal recognition result. Overall, STAN-A and STMN-A promote and complement each other in the recognition of emotions and DSTAN can get the best performance.

3.3.2. Performance Verification of Different Modules. The proposed DSTAN combines high-level and low-level feature fusion modules (LHFN), spatial attention modules, and temporal attention modules based on the dual-stream network. To verify the effectiveness of different modules, ablation experiments are performed on the CASME II data set.

The basic model only contains the network, that is, the DSTAN removes the LHFN module, two spatial attention modules, and the temporal attention module. We compare the basic model with the models that are added LHFN module (basic model + LHFN), spatial attention modules (basic model + LHFN + SA), and temporal attention module (basic model + LHFN + SA + TA). Table 2 shows the comparison result. It can be seen that by adding three modules to the basic model, the recognition result has been further improved. By adding the LHFN module, accuracy is increased by 1.62% and *F1-score* is increased by 1.67%. By adding the spatial attention modules, accuracy is increased by 4.07% and *F1-score* is increased by 3.64%. After adding the temporal attention module, accuracy is increased by 4.06% and *F1-score* is increased by 3.81%. The basic model + LHFN + SA + TA model (DSTAN) obtains the best recognition result and robustness. Therefore, these modules can improve the performance of recognizing microexpressions, which verifies the effectiveness of the modules. The LHFN module enables the model to learn discriminative semantic information of the microexpression sequence. The spatial attention module and temporal attention module can make the model learn more detailed and effective features.

3.3.3. DSTAN Performance Analysis. We evaluate the DSTAN by using each subject as a testing set on three data sets. The experimental results are shown in Figure 6. The abscissa is the coding number of the subject, and the ordinate is the recognition accuracy of the subject. On SMIC, as shown in Figure 6(a), the DSTAN has good recognition results for most subjects, but the accuracy of the 3*rd* and the 4*th* subject is poor due to the action units of "negative" are similar to "surprise", which makes them easy to be confused. On CASME, as shown in Figure 6(b), the accuracy rates on all of the 9 subjects are 100%, but the recognition result of the 1*st* subject is lower, it is because that "repression" is easily confused with "disgust" and "surprise". As shown in Figure 6(c) on CASME II, the accuracy rate of the 16*th* subject is lower because there is a small number of this subject, only 4 sequences.

We calculated the confusion matrix of DSTAN on SMIC, CASME, and CASME II, as shown in Figure 7. On SMIC, as shown in Figure 7(a), the DSTAN performs well on identifying negative, positive, and superior emotions because the

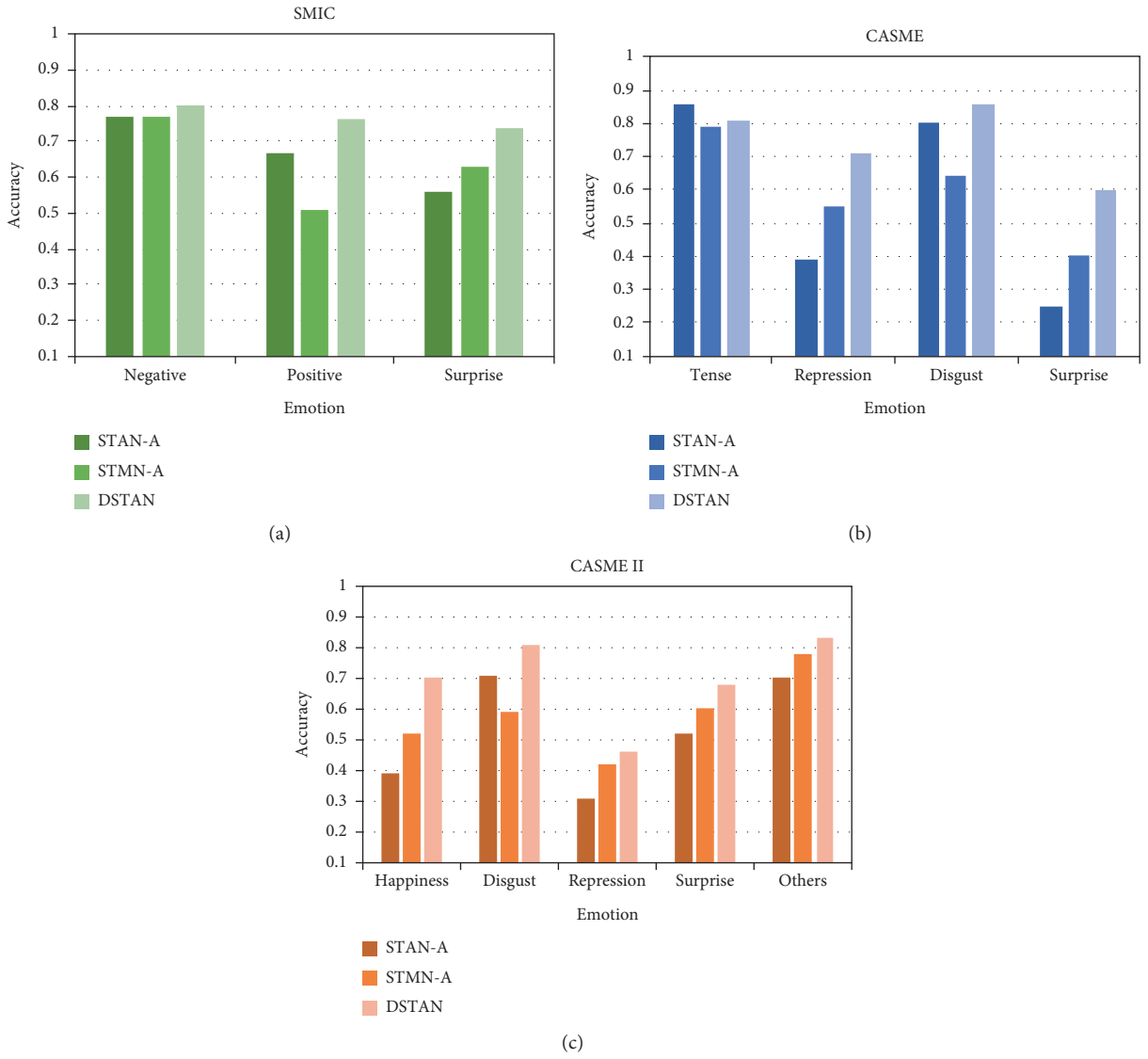


Figure 5: Comparison of a single network with dual-stream network: (a) SMIE, (b) CASME, and (c) CASME II.

Table 2: The influence of different modules on the network.

Method	CASME II	
	Accuracy	*F1-score*
Basic model	0.655	0.637
Basic model + LHFN	0.671	0.654
Basic model + LHFN + SA	0.711	0.690
Basic model + LHFN + SA + TA (DSTAN)	0.752	0.728

distribution of samples on this data set is relatively uniform. On CASME, as shown in Figure 7(b), the DSTAN gets a higher result for "tense", "repression", and "disgust", but it is not good at recognizing "surprise" due to its small range of muscle motion. On CASME II, as shown in Figure 7(c), the DSTAN performs poorly in emotions "surprise" and "repression". It is because "surprise," "repression", and "others" are easy to be confused as a result of the number of "others" has the largest data and the data set is unbalanced. The experimental results show that for the microexpression recognition task, the total number of samples of each emotion, the difference number of emotions, and the motion amplitude of microexpression are the important factors.

3.3.4. Integration Mode Validation. Since most of the approaches adopt the weighted sum model to integrate, we compare it with the feature concatenate-SVM method.

The DSTAN with weighted sum integration mode is labeled as DSTAN-Average, and the DSTAN integrated by feature concatenate-SVM is marked as DSTAN-SVM. Table 3 shows the comparison results on three microexpression data sets. The evaluation index *Precision* represents the discrimination ability of the model for negative samples, and *Recall* represents the recognition ability of the model for positive samples. As can be seen from Table 3, the performance of DSTAN-SVM is better than that of DSTAN-Average to a certain extent. On SMIC, CASME, and CASME II, the *Precision* of DSTAN-SVM is 3.29%, 2.94%, and 2.69%

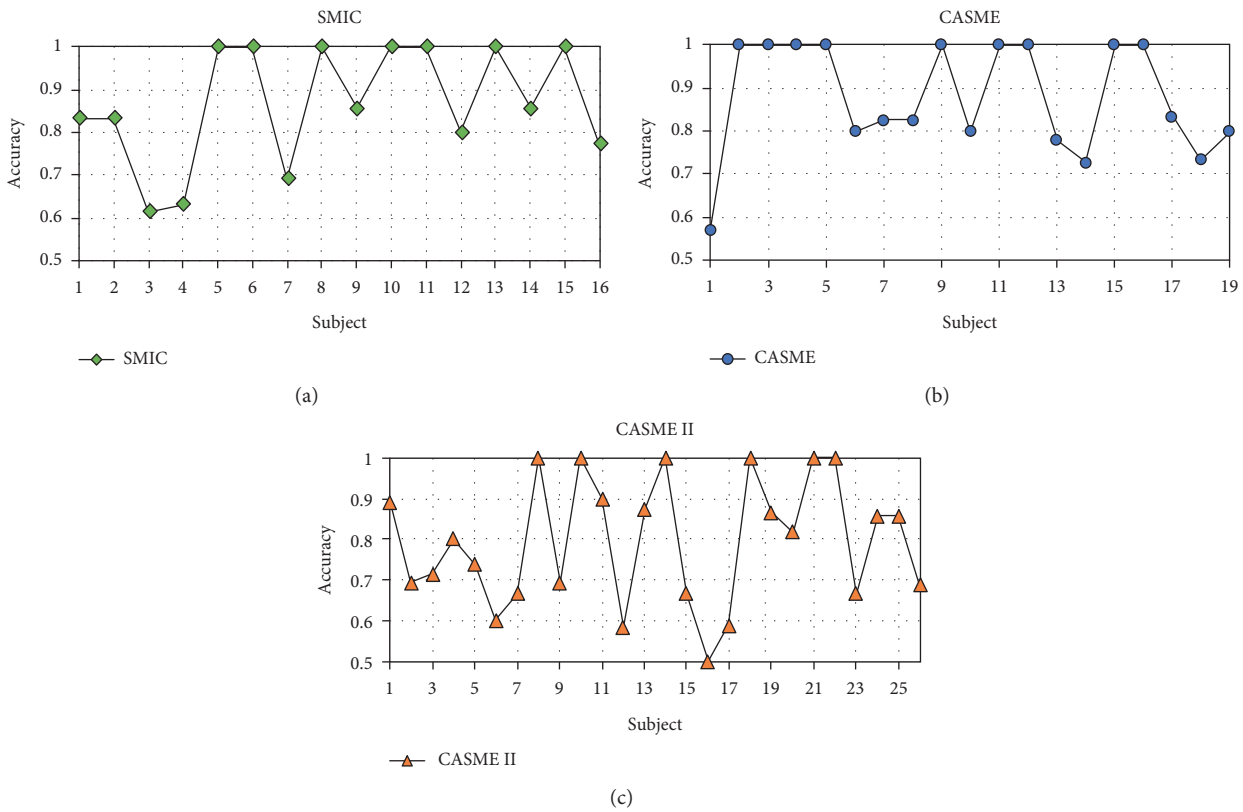


Figure 6: The performance of DSTAN for each subject: (a) SMIE, (b) CASME, and (c) CASME II.

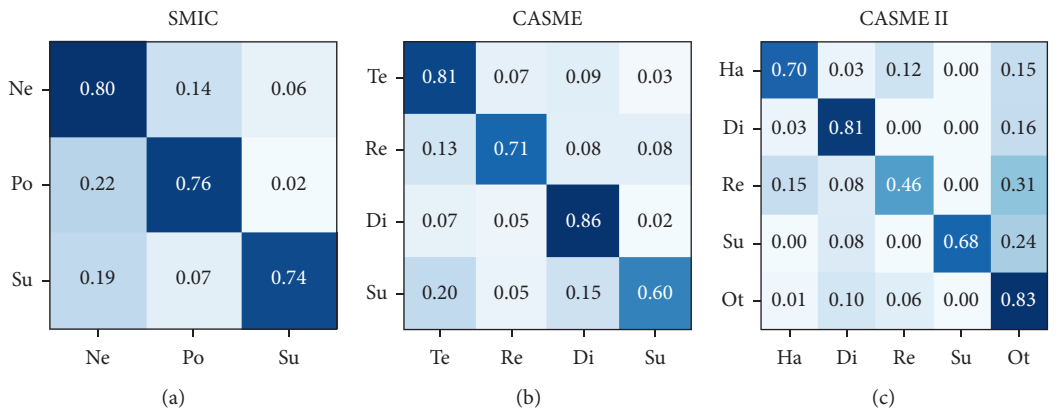


Figure 7: Confusion matrix of DSTAN on three data sets: (a) SMIE, (b) CASME, and (c) CASME II.

higher than that of DSTAN-Average, and *Recall* is increased by 4.46%, 2.39%, and 6.21%, respectively, which indicates that DSTAN-SVM has the discrimination ability for positive and negative samples. The accuracy of DSTAN-SVM is improved by 4.27%, 3.49%, and 4.47% compared with that of DSTAN-Average, and the *F1-score* is improved by 3.89%, 2.67%, and 4.68%, respectively, which indicates that DSTAN-SVM has better recognition performance and robustness.

3.3.5. Comparison with State-of-the-Arts. The recognition performance of DSTAN is compared with some state-of-the-arts. The experimental results of the three data sets are shown in Table 4. LBP-TOP, LBP-SIP, STLBP-IP, STCLQP, and LCBP are texture feature-based methods. FDM, MDMO, Sparse MDMO, and Bi-WOOF are optical flow-based methods. 3D-FCNN, ELRCN-TE, STRCN, and MERTA are depth learning-based methods.

As can be seen from Table 4, on SMIC, the accuracy of DSTAN is 77.44%, which is 6.14% higher than the best method STRCN, and *F1-Score* is increased to 0.7783, which is 7.42% higher than Sparse MDMO. On CASME, the accuracy of DSTAN reaches 77.91%, which is 3.08% higher than Sparse MDMO, and *F1-Score* is 0.7516, which is 0.18% higher than Sparse MDMO. On CASME II, the Accuracy of DSTAN reaches 75.20%, which is 5.1% higher than the

Table 3: Comparison results of two model integration methods.

Metrics	DSTAN-Average			DSTAN-SVM		
	SMIC	CASME	CASME II	SMIC	CASME	CASME II
Accuracy	0.73	0.74	0.71	0.77	0.78	0.75
Precision	0.75	0.73	0.74	0.79	0.76	0.77
Recall	0.73	0.72	0.63	0.77	0.75	0.70
F1-score	0.74	0.73	0.68	0.78	0.75	0.73

Table 4: Comparison of DSTAN with other approaches (F1 : *F1-Score*).

Method	SMIC		CASME		CASME II	
	Accuracy	F1	Accuracy	F1	Accuracy	F1
LBP-TOP [6]	0.44	0.41	0.37	0.33	0.47	0.35
LBP-SIP [7]	0.45	0.45	0.37	0.33	0.47	0.45
STLBP-IP [8]	0.58	—	—	—	0.56	—
STCLQP [9]	0.64	0.64	0.57	0.50	0.58	0.58
LCBP [10]	0.71	0.68	—	—	0.70	0.70
MDMO [11]	0.59	0.59	0.56	0.56	0.52	0.50
FDM [12]	0.55	0.54	0.56	0.49	0.46	0.41
Sparse MDMO [13]	0.71	0.70	0.75	0.75	0.67	0.69
Bi-WOOF [14]	0.62	0.62	—	—	0.59	0.61
CNN-LSTM [15]	—	—	—	—	0.61	—
3D-FCNN [16]	0.56	—	0.54	—	0.59	—
ELRCN-TE [17]	—	—	—	—	0.52	0.50
STRCN [18]	0.72	0.69	—	—	—	—
MERTA [21]	—	—	—	—	0.61	—
DSTAN	0.77	0.78	0.78	0.75	0.75	0.73

baseline method LCBP, and *F1-Score* is increased to 0.7283, which is 2.83% higher than LCBP. Experimental results show that the proposed DSTAN has better recognition performance than state-of-the-arts.

4. Conclusion

In this paper, we have presented a novel architecture for dynamic facial microexpression recognition combining deep and handcraft features, which can recognize the micro-expressions with higher accuracy. Both the deep learning method and the handcraft method are fused to identify the microexpressions by learning features not only the tiny skin change but also the semantic properties from sequences. The approach successfully exploits spatial and temporal features of microexpression simultaneously. Particularly, the feature framework has been established to identify the dynamic microexpressions successfully by extracting robust features from data. In the end, we conduct extensive validation experiments to demonstrate the proposed method. The excessive experimental results showed that with an accuracy of 75.51% on SMIC, an accuracy of 81.26% on CASME_B, and an accuracy of 76.14% on CASME 2 in terms of the 5-class microexpression recognition, our framework can surpass other methods.

In the future, we aim to evaluate our approach on additional microexpression data sets. We also consider training our approach on cross-data-set experiments and explore the effective method to improve the recognition performance of microexpressions on action units.

Data Availability

All data included in this study are available upon request by contact with the corresponding author.

Disclosure

Yan Wang and Yikun Huang are co-first authors.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Yan Wang and Yikun Huang contributed equally to this work

Acknowledgments

This work was supported in part by the Research Innovation Team of Concord University College Fujian Normal University in 2020 (No. 2020-TD-001), in part by the third batch of key lifelong education projects in Fujian province (No. ZS20033), and in part by Fujian Province 13th Five-Year Plan Teaching Reform Project in 2019 (No. FBJG20190156).

References

[1] W.-J. Yan, Q. Wu, J. Liang, Y.-H. Chen, and X. Fu, "How fast are the leaked facial expressions: the duration of micro-expressions," *Journal of Nonverbal Behavior*, vol. 37, no. 4, pp. 217–230, 2013.

[2] M. A. Takalkar, M. Xu, Q. Wu, and Z. Chaczko, "A survey: facial micro-expression recognition," *Multimedia Tools and Applications*, vol. 77, no. 15, pp. 19301–19325, 2018.

[3] X. Duan, Q. Dai, X. Wang, Y. Wang, and Z. Huo, "Recognizing spontaneous micro-expression from eye region," *Neurocomputing*, vol. 217, pp. 27–36, 2016.

[4] X. Xue, J. Chen, and X. Yao, "Efficient user involvement in semi-automatic ontology matching," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 5, no. 2, pp. 214–224, 2021.

[5] G. Liu, X. Chen, R. Zhou, S. Xu, Y.-C. Chen, and G. Chen, "Social learning discrete particle swarm optimization based two-stage X-routing for IC design under intelligent edge computing architecture," *Applied Soft Computing*, vol. 104, no. 6, Article ID 107215, 2021.

[6] T. Pfister, X. Li, G. Zhao, and M. Pietikäinen, "Recognizing spontaneous facial micro-expressions," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 1449–1456, Barcelona, Spain, January 2011.

[7] Y. Wang, J. See, R. C. W. Phan, and Y.-H. Oh, "LBP with six intersection points: reducing redundant information in LBP-TOP for micro-expression recognition," in *Proceedings of the Asian Conference on Computer Vision*, pp. 525–537, Singapore, November 2014.

[8] X. Huang, S. J. Wang, G. Zhao, and M. Pietikäinen, "Facial micro-expression recognition using spatiotemporal local binary pattern with integral projection," in *Proceedings of the IEEE International Conference on Computer Vision Workshops*, pp. 1–9, Santiago, CH, USA, December 2015.

[9] X. Huang, G. Zhao, X. Hong, W. Zheng, and M. Pietikäinen, "Spontaneous facial micro-expression analysis using Spatiotemporal completed local quantized patterns," *Neurocomputing*, vol. 175, pp. 564–578, 2016.

[10] M. Yu, Z. Q. Guo, Y. Yu, Y. Wang, and S. Cen, "Spatiotemporal feature descriptor for micro-expression recognition using local cube binary pattern," *IEEE Access*, vol. 7, p. 1, 2019.

[11] Y. Liu, J. Zhang, W. Yan, S. Wang, G. Zhao, and X. Fu, "A main directional mean optical flow feature for spontaneous micro-expression recognition," *IEEE Transactions on Affective Computing*, vol. 7, no. 4, pp. 299–310, 2016.

[12] F. Xu, J. Zhang, and J. Z Wang, "Micro expression identification and categorization using a facial dynamics map," *IEEE Transactions on Affective Computing*, vol. 8, no. 2, pp. 254–267, 2017.

[13] Y. Liu, B. Li, and Y. Lai, "Sparse MDMO: learning a discriminative feature for spontaneous micro-expression recognition," *IEEE Transactions on Affective Computing*, vol. 12, no. 1, pp. 254–261, 2018.

[14] S. Liong, J. See, K. Wong, and R. C. W. Phan, "Less is more: micro-expression recognition from video using apex frame," *Signal Processing: Image Communication*, vol. 62, pp. 82–92, 2018.

[15] D. H. Kim, W. J. Baddar, and Y. M. Ro, "Micro-expression recognition with expression-state constrained Spatial-temporal feature representations," in *Proceedings of the 24th ACM International Conference on Multimedia*, pp. 382–386, Amsterdam, Netherlands, October 2016.

[16] J. Li, Y. Wang, J. See et al., "Micro-expression recognition based on 3d flow convolutional neural network," *Pattern Analysis & Applications*, vol. 22, no. 4, pp. 1331–1339, 2019.

[17] H. Q. Khor, J. See, R. C. W. Phan, and W. Lin, "Enriched long-term recurrent convolutional network for facial micro-expression recognition," in *Proceedings of the 13th IEEE International Conference on Automatic Face & Gesture Recognition*, pp. 667–674, Xian's, China, May 2018.

[18] Z. Xia, X. Hong, X. Gao, X. Feng, and G. Zhao, "Spatiotemporal recurrent convolutional networks for recognizing spontaneous micro-expressions," *IEEE Transactions on Multimedia*, vol. 22, no. 3, pp. 626–640, 2020.

[19] X. Chu, W. Yang, W. Ouyang, C. Ma, A. L. Yuille, and X. Wang, "Multi-context attention for human pose estimation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5669–5678, Honolulu, HW, USA, July 2017.

[20] X. Zhang, T. Wang, J. Qi, H. Lu, and G. Wang, "Progressive attention guided recurrent network for salient object detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 714–722, Salt Lake, UT, USA, July 2018.

[21] B. Yang, J. Cheng, Y. Yang, B. Zhang, and J. Li, "MERTA: micro-expression recognition with ternary attentions," *Multimedia Tools and Applications*, pp. 1–16, 2019.

[22] Z. Zhou, G. Zhao, Y. Guo, and M. Pietikainen, "An image-based visual speech animation system," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 10, pp. 1420–1432, 2012.

[23] S. Ioffe and C. Szegedy, "Batch normalization: accelerating deep network training by reducing internal covariate shift," in *Proceedings of the International Conference on Machine Learning*, pp. 448–456, Lille, France, July 2015.

[24] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," in *Proceedings of the European Conference on Computer Vision*, pp. 818–833, Springer, Zurich, Switzerland, July 2014.

[25] J. Hu, L. Shen, G. Sun, and E. Wu, "Squeeze-and-Excitation networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 7132–7141, Salt Lake City, UT, USA, June 2018.

[26] S. Woo, J. Park, J. Lee, and S. Kweon, "CBAM: convolutional block attention module," in *Proceedings of the European Conference on Computer Vision*, pp. 3–19, Munich, Germany, September 2018.

[27] X Li, T Pfister, X Huang, J. Zhao, and M. Pietikainen, "A spontaneous micro-expression database: inducement, collection, and baseline," in *Proceedings of the 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition*, pp. 1–6, Shanghai, China, April 2013.

[28] W. J. Yan, Q. Wu, Y. J. Liu, S. W. Wang, and X. Fu, "CASME database: a dataset of spontaneous micro-expressions collected from neutralized faces," in *Proceedings of the 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition*, pp. 1–7, Shanghai, China, April 2013.

[29] W. J. Yan, X. Li, S. J. Wang et al., "Casme II: an improved spontaneous micro-expression database and the baseline evaluation," *PLoS One*, vol. 9, no. 1, Article ID e86041, 2014.

[30] A. C. Ngo, R. C. Phan, and J. See, "Spontaneous subtle expression recognition: imbalanced databases and solutions," in *Proceedings of the Asian Conference on Computer Vision*, pp. 33–48, Singapore, November 2014.

Hindawi
Journal of Healthcare Engineering
Volume 2021, Article ID 9957828, 9 pages
https://doi.org/10.1155/2021/9957828

Research Article

Study on the Effect of Driving Time on Fatigue of Grassland Road Based on EEG

Yule Zhang and Shoulin Zhu

College of Energy and Transportation Engineering, Inner Mongolia Agricultural University, Hohhot 010000, China

Correspondence should be addressed to Shoulin Zhu; zylscience@emails.imau.edu.cn

Received 23 March 2021; Revised 8 May 2021; Accepted 17 June 2021; Published 9 July 2021

Academic Editor: Hao Chun Lu

In order to study the change law of the fatigue degree of grassland expressway drivers over time, this paper takes the semidesert grassland landscape of Xilinhot city as the experimental environment and takes the provincial highway S101 (K278–K424) as an example to design an actual driving test. Taking Urumqi, Inner Mongolia Autonomous Region, as the experimental section, combined with the Biopac MP150 multichannel physiological instrument and its auxiliary knowledge software and mathematical statistics methods, the relationship between EEG and time was studied. The test results show that the primary fatigue factor F_1 and the secondary fatigue factor F_2 can summarize the fatigue law characterized by 96.42% of EEG information. During 130 minutes of driving on the prairie highway, the periods of high fatigue were 105 minutes and 125 minutes, respectively. Driving fatigue can be divided into three stages over time: 5–65 min fatigue-free stage, 70–85 min fatigue transition stage, and 90–130 min fatigue stage. Fatigue changes over time. The law follows the Gaussian function and the sine function.

1. Introduction

Drivers need to keep their concentration in the road environment to ensure the safety of driving tasks [1]. Studies have shown that low load conditions such as monotony can lead to fatigue symptoms and impaired driving performance, although drivers are neither tired nor lack sleep before driving tasks [2, 3]. This kind of fatigue is caused by driving itself, which is called passive fatigue. Passive fatigue is caused by nerve habituation, such as driving on similar and repeated routes [4], which is different from the driver's active fatigue caused by physiological rhythm and lack of sleep [5]. Fell and Black conducted a driver fatigue accident survey in Sydney, which covered all accidents caused by fatigue. It was found that 45% of crash drivers were not tired at all before the accident, and the road geometry was very monotonous [6]. It can be seen that the accidents caused by driving fatigue in the monotonous environment are potential. The research on this kind of fatigue can effectively reduce the accident rate. Relevant research shows that the most reliable way to quantify driver mental fatigue in the monotonous environment is to use electroencephalogram (EEG), which is a very promising indicator [7].

The influencing factors of driving fatigue include not only the monotonous degree of the road environment, the physiological rhythm of drivers, and the time of driving but also the duration of work. Driving for a long time will increase the fatigue degree. Research shows that 30 minutes of monotonous driving can cause alertness disorder [8], 11% of accidents are related to fatigue, and 62% of accidents occur in driving less than 100 miles, that is, 87.6 minutes [4]. Larue et al. proposed that the effect of quantitative time on fatigue needs to be set beyond 40 minutes [9]. Therefore, in the monotone road environment, the time task test of more than one hour is the basis of quantifying the effect of time on fatigue.

The mileage of roads in the Inner Mongolia Autonomous Region is increasing constantly. By the end of 2018, the total mileage of class roads was 192,200 kilometers, of which grassland roads accounted for 73.3%. Grassland road has the characteristics of monotonous roadside landscape and road alignment. When driving on prairie roads, drivers will

encounter the problem of insufficient stimulus information, which will easily lead to the decrease of driver's attention and alertness, premature driving fatigue, and traffic accidents [10].

The innovative content of this paper is as follows:

(1) Combining the Biopac MP150 multichannel physiological instrument and its auxiliary knowledge software and mathematical statistics methods, the relationship between EEG and time has been studied

(2) The EEG model is used to analyze and research related fatigue data

(3) The EEG model and EEG are used to study the relationship between driving fatigue and fatigue data

2. Materials and Methods

2.1. Test Personnel. In the experimental study, the number and representativeness of the subjects will directly affect the reliability of the experimental results. In formula (1), n is the required sample size [11], Z is the statistics with a certain confidence level, and s is the standard deviation of each test index of driving characteristics [12]. Based on the research results at home and abroad, $s = 0.094$ and $d = 0.04$; in this paper, $z = 1.96$ at 95% confidence level is taken; the minimum number of subjects in this experimental study is 22. Taking into account the characteristics of significant individual differences in EEG signals and susceptibility to environmental factors, the number of subjects was determined as 30 in this experiment.

$$n \geq \frac{z^2 \times s^2}{d^2}. \quad (1)$$

According to the statistics that the proportion of male and female drivers with legal driving license is 7/3 and the analysis of driving age and age of drivers in traffic accidents [13], the test subjects are determined, as shown in Table 1.

In order to ensure the effectiveness of data collection, the requirements for the test driver are as follows: good health, sufficient sleep before the test, no activities that stimulate the heart within 24 hours, no smoking, and vigorous exercise 1 hour before the test.

2.2. Road Sections and Test Vehicles. The test section is S101 provincial road in Xilinhot city, Inner Mongolia. The relevant parameters of the section are shown in Table 2: 84.9% of the road lines are long straight lines, with large curve radius and 61% of large radius curves. The longitudinal slope is gentle, and the undulation is not obvious. In addition, the vegetation coverage is low, and the desertification is serious. The test section has the characteristics of monotonous landscape and line type.

By the end of 2017, China's car ownership was 209.0667 million, including 180.3869 million small passenger vehicles, accounting for 86.28% (China National Bureau of statistics, http://data.stats.gov.cn/index.htm). The research group investigated the traffic flow of the grassland road, and the proportion of small vehicles was 91.5%. Therefore, the Honda Civic car (automatic transmission) is used as the test vehicle. According to the time task test, the time of the vehicle to be tested must be more than 40 min. The test road section is selected to be 150 km long in one way, 300 km back and forth, and 150 min under the speed limit of 80 km/h.

2.3. Test Process. The choice of test time and the control of the environment will seriously affect the test results. In order to reduce the influence of factors other than time on the results, sunny days with similar illuminance and postlunch dip were selected for testing [14], which was the peak period of daytime fatigue accidents. The subjects arrived at the test site at 12:30. Familiar with the test equipment, understand the test requirements, complete the wiring work; 12:50, drive for 5 minutes, familiar with the environment and equipment; 13:00, officially start the test.

2.4. Test Instruments and Requirements. The MP150 multichannel physiological recorder (maximum sampling rate: 400 kHz, including the EEG amplifier) and AcqKnowledge 4.1 data analysis software are used to obtain the EEG data. The research related to the fatigue of monotonous road drivers is to characterize the driver's mental fatigue by the average power of three waveforms of α (8–13 Hz), β (14–30 Hz), and θ (4–7 Hz). The sampling frequency is adjusted to 250 Hz based on kurtosis and probability criteria, and the band-pass filter frequency is set to 1 Hz (high-pass filter) to 30 Hz (low-pass filter) [15]. Two lead 110 shield wires are, respectively, connected to the disposable electrode piece on the driver's left forehead and right mastoid, and the other end is connected to the EEG100C EEG amplifier (the amplifier gain is set at 5000). AcqKnowledge 4.1 software is used to extract one EEG power every 30 s as absolute power. The 5-minute absolute power at the beginning and end of driving is deleted, SPSS 20.0 software and the method of Pauta criterion are used to remove the abnormal value of EEG power [16], and finally, the analysis of EEG power within 130 minutes is determined.

3. Results

3.1. Determination of the Fatigue Factor. The absolute power magnitude of the EEG signal varies greatly (as shown in Figure 1). The EEG power of each subject is normalized according to the formula [17], and an average value is taken every five minutes to average the EEG power of 22 drivers. The results are shown in Figure 2. Among them, 30 minutes ago, the change rule was the same, and there were differences among the three wave patterns in 30–130 minutes. The related research shows that the power of α and θ increases with the increase of fatigue, the increase of β-power indicates the increase of alertness, and the larger the ratio of fast wave to slow wave is, the deeper the fatigue is [18, 19]. Therefore, the accumulation rule of fatigue with time shows strong time characteristics. In order to obtain the fatigue factors with obvious time characteristics and strong representativeness, dimension reduction is carried out [20].

Table 1: Subject information.

Gender	Number	Age	Driving age	BIM
Female	6	32.51 ± 5.23	3.12 ± 1.33	19.53 ± 5.42
Male	24	35.46 ± 7.47	3.97 ± 2.29	21.59 ± 2.77

Table 2: Statistics of relevant parameters for the S101 section line.

Speed limit	Line type						Landscape		
80 km/h	Route length (km)	The longest straight line (m)	Linear density (km/km)	Maximum curve radius (m)	Curve radius distribution (m) and proportion (%)	Maximum longitudinal grade (%)	Desertification is serious (%)	Medium vegetation coverage (%)	High vegetation coverage (%)
	150	8741	0.849	40,000	1000–2000/61	−4.92	13	67	20

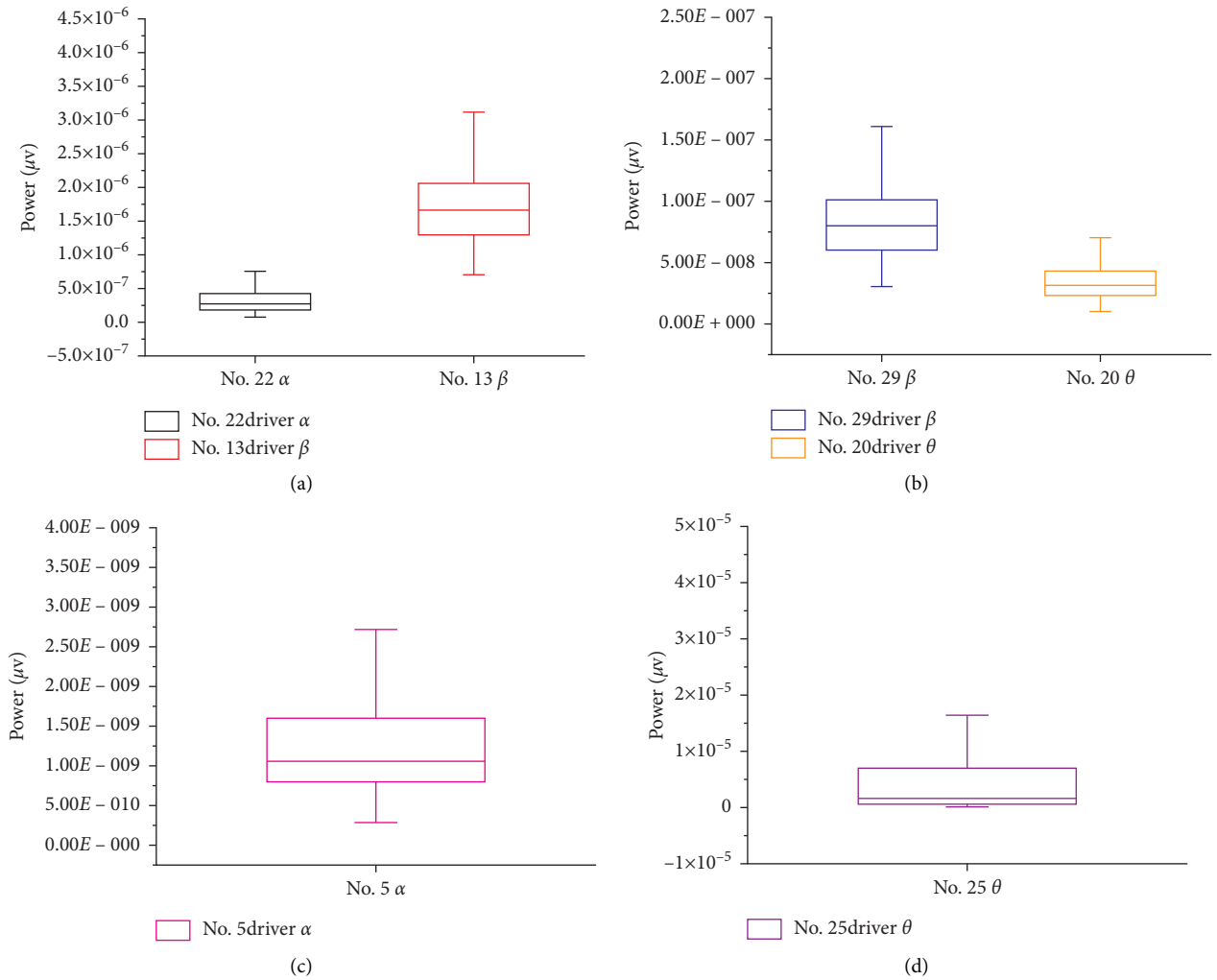


Figure 1: Absolute power (α, β, and θ) of 6 drivers.

SPSS 20.0 and OriginPro 2019 software are used to analyze the five indicators [21]. The results are shown in Figure 3 and Table 3. Autoencoder dimension reduction is shown in Figure 4.

As shown in Figure 3, the coordinate axis and the long and short axis of the confidence ellipse are parallel, indicating that the variable of the long axis describes the main change of the data, and the variable of the short axis

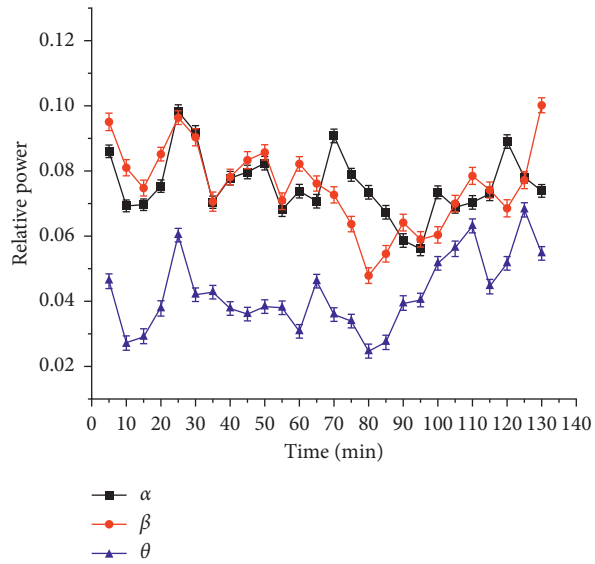


Figure 2: Changes in the EEG relative power of 22 drivers over time.

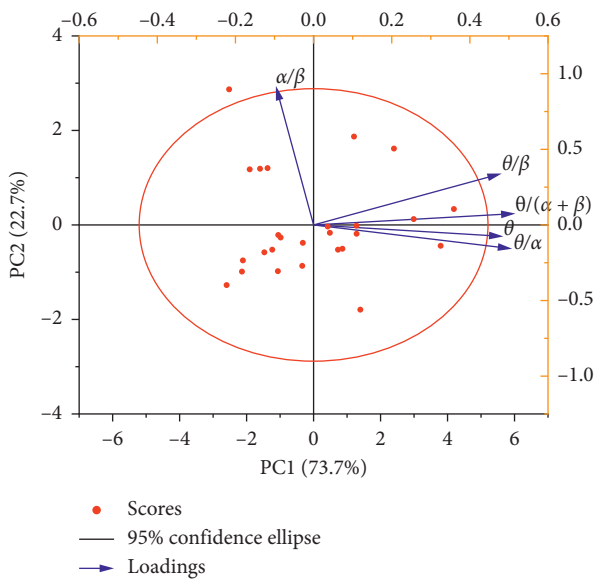


Figure 3: Principal component analysis confidence ellipses for 5 indicators.

describes the secondary change of the data [22]. The characteristic value shows that F_1 represents 3.686 times of the original data, F_2 represents 1.135 times of the original data, and the cumulative percentage of them is 96.42%, which can explain the fatigue rule represented by five indicators [23]. Similar to the principal component analysis (PCA), we also use the autoencoder to provide visual analysis on the two feature dimensions, and the relationship of the power of α, β, and θ is also given accordingly. Expressions (2) and (3) of the main fatigue factor F_1 and the secondary fatigue factor F_2 are derived from the linear relationship between the main components and variables. The change rule of fatigue factors F_1 and F_2 with time is shown in Figure 5 [24].

The larger the fatigue factor is, the deeper the fatigue degree is. The smaller the fatigue factor is, the shallower the fatigue degree is. There is no physiological significance for positive and negative factors [25].

$$F_1 = 0.931\theta + \frac{0.99\theta}{\alpha+\beta} + \frac{0.975\theta}{\alpha} + \frac{0.923\theta}{\beta} - 0.189\frac{\alpha}{\beta}, \quad (2)$$

$$F_2 = -0.079\theta + \frac{0.081\theta}{\alpha+\beta} - \frac{0.164\theta}{\alpha} + \frac{0.367\theta}{\beta} + 0.98\frac{\alpha}{\beta}. \quad (3)$$

3.2. Clustering of Fatigue Factors and Determination of High-Fatigue Time. In order to describe the stage characteristics of fatigue, *K*-means test was used to cluster F_1 and F_2. According to the trend analysis of Section 2.1, F_1 and F_2 were clustered into three categories (I and IV, II and V, and III and VI, in which the fatigue degree ranked I < II < III, IV < V < VI, but I, II, III and IV, V, VI had no statistical relationship) [26]. The results are shown in Figure 6.

In order to quantitatively describe the rule of fatigue factor and time, regression analysis is used to fit F_1 with logistic, and the results are shown in Figure 7 and Table 4. According to Figure 5, the fatigue state in 130 min can be divided into two stages: 5–85 min and 90 min–130 min (one EEG average value is taken every 5 min, which represents the EEG power in 5 min. In this paper, the time point at the end of every 5 min represents the EEG power in this period) [27]. In 5–85 min, class I F_1 is dominant, class II F_1 only accounts for 17.65%, and class II and class III F_1 alternate in 90 min–130 min, which account for the same proportion. It can be seen that the driver's fatigue accumulates and fluctuates greatly in 90–130 min. Further sine fitting is carried out for the fatigue factor $F_{1\text{-}1}$ (100–130 min) in this time, and the results are shown in Tables 3 and 5. Although 90 min and 95 min belong to class II, they are quite different and have no statistical significance, so they are not fitted. F_1 logistic fitting is shown in Figure 7. F_1 sine fitting is shown in Figure 8 [28].

In the same way, Gauss fitting is carried out for F_2, and the results are shown in Figure 9 and Table 5. The overall fatigue state is divided into three stages: 5–65 min, 70–85 min, and 90–130 min [29]. In 5–65 min, F_2 completely belongs to category IV, in 70–85 min, there are two categories V and VI, and the peak value appears in 80 min. In 90–130 min, category IV and category V F_2 alternately appear and share the same proportion. It can be seen that the driver fatigue accumulates and fluctuates greatly in 90–130 min. Furthermore, 70–85 min and 90–130 min were fitted. The results are shown in Figures 10 and 11 and Table 4 [30].

In addition, the time corresponding to the extreme point is the high-fatigue period. According to the application scope of each function, the high-fatigue period of F_1 tends to the end of the driving task; the high-fatigue period of $F_{1\text{-}1}$ is 106.47 min and 124.47 min; the high-fatigue period of $F_{2\text{-}1}$ is 79 min; the high-fatigue period of $F_{2\text{-}1}$ is 80 min; the high-fatigue period of $F_{2\text{-}2}$ is 104.13 min and 123.17 min. Among them, $F_{1\text{-}1}$ and $F_{2\text{-}2}$ were the same for 105 min and 125 min, while F_2 and $F_{2\text{-}2}$ indicate that the fatigue factor fluctuates greatly in 70–85 min

Table 3: Principal component analysis results.

	The total variance of the principal components		Component matrix				
	Characteristic value	Cumulative percentage (%)	θ	$\theta/(\alpha+\beta)$	θ/α	θ/β	α/β
F_1	3.686	73.728	0.931	0.990	0.975	0.923	−0.189
F_2	1.135	96.419	−0.079	0.081	−0.164	0.367	0.980

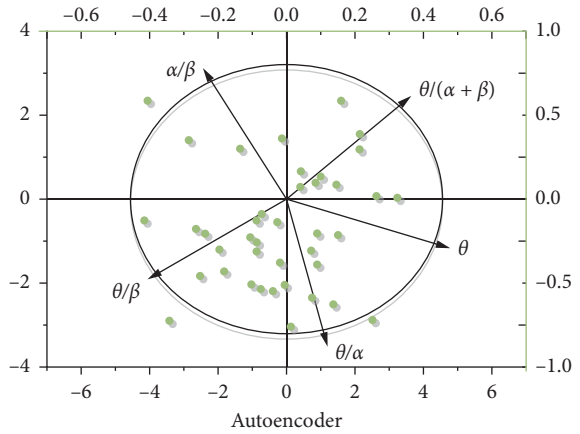


Figure 4: Autoencoder dimension reduction.

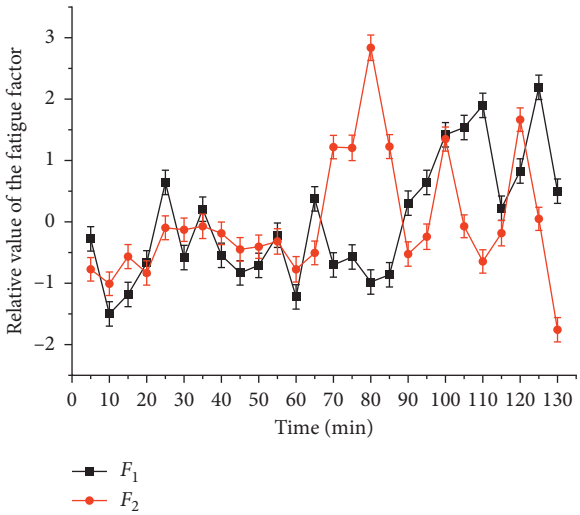


Figure 5: Driver fatigue factor changes with time.

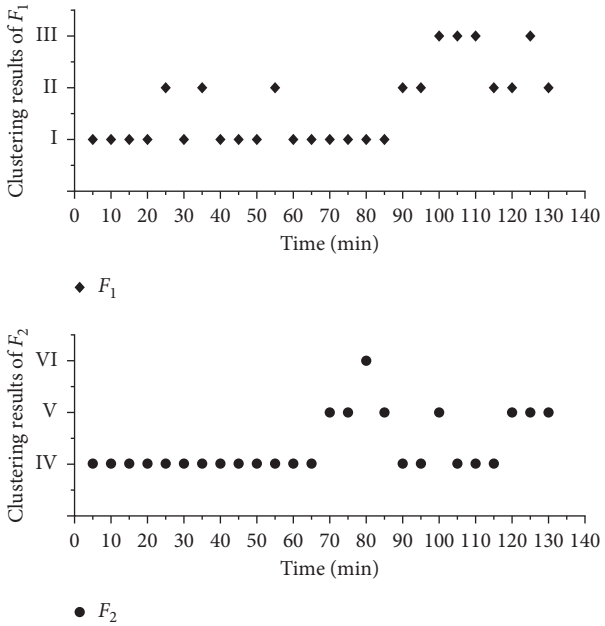


Figure 6: Cluster analysis results.

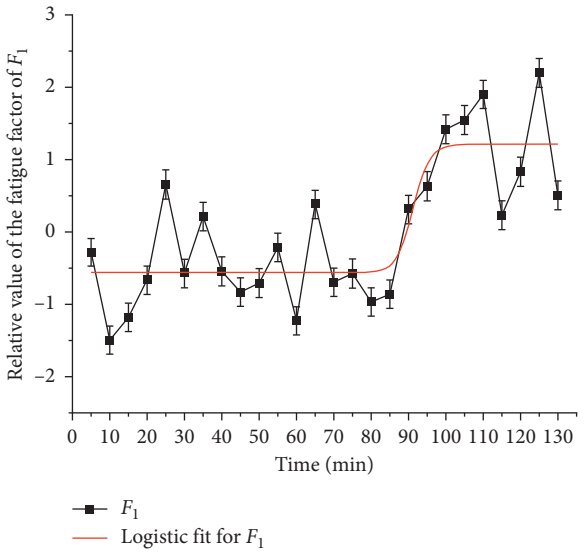


Figure 7: F_1 logistic fitting.

and peaks at 80 min. At this time, the F_1 fitting curve in Figure 7 shows a significant inflection point and the trend of sudden increase. It can be seen that 70–85 min is the transition stage from the driver to the fatigue state [31]. Therefore, fatigue can be divided into three stages: 5–65 min no fatigue stage, 70–85 min fatigue transition stage, and 90–130 min fatigue stage. $F_{2\text{-}1}$ Gauss fitting is shown in Figure 10.

3.3. Fatigue Variation with Time. Based on the above analysis, the driver fatigue shows significant time characteristics. There is no obvious fatigue accumulation within 5–65 minutes, and the fatigue factor fluctuates in a small range. Each increase and decrease is controlled within 10 minutes. The driver's fatigue regulation mechanism can completely prevent the accumulation of fatigue on the time axis. The specific rule is F_1 and F_2 decrease within 5–10 min, the driving task is just started, the driver needs to adapt to the driving task, fbe concentrated with full of freshness to the environment, and should not show a negative mental state of fatigue. Within 15–65 minutes, the driver's mental state is stable, even though F_1 fluctuates slightly,

Table 4: The accuracy of classification results.

Classification matrix	Model					
	ANN-BP (%)	ANN-PSO (%)	RBF-ROLS + D-opt (%)	[32] (%)	[33] (%)	Our model (%)
Sensitivity/TPR	76.67	71.81	80.00	92.36	82.36	93.26
Specificity/TNR	76.94	62.78	95.56	93.06	83.09	92.78
Accuracy (%)	76.81	67.29	87.78	86.98	85.72	90.07

Table 5: Fitting formula and fitting degree.

Fitting formula	Adj. R^2	t	$dF/dt = 0$
$F_1 = 1.21 - (1.77/1 + (0.01t)^{40.94})$	0.61 > 0.4	5–130	$t \longrightarrow +\infty$
$F_{1-1} = 1.2 + \sin[(t + 6.03/9)\pi]$	0.48 > 0.4	100–130	$t = -1.53 \pm 9k$, $k = 0, 1, 2, 3\cdots$
$F_2 = -0.29 + 2.83e^{-0.02(t-79)^2}$	0.45 > 0.4	5–130	79
$F_{2-1} = 1.22 + 1.62e^{-1.06(t-80)^2}$	1.00 > 0.4	70–85	80
$F_{2-2} = 0.3 + \sin[(t - 1.37/9.5)\pi]$	0.80 > 0.4	90–130	$t = 18.45 \pm 9.52k$, $k = 0, 1, 2, 3\cdots$

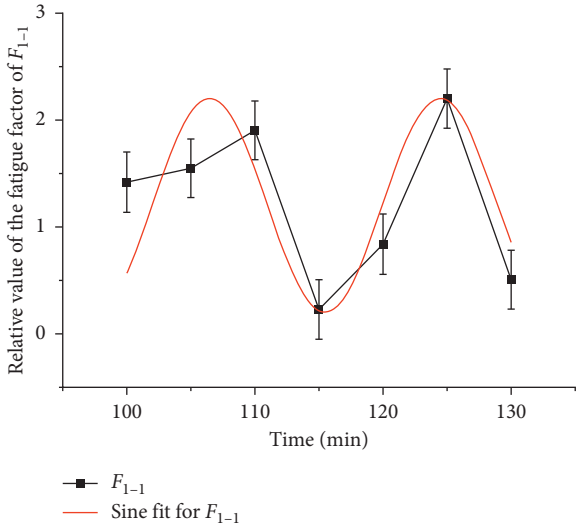


Figure 8: $F_{1\text{-}1}$ sine fitting.

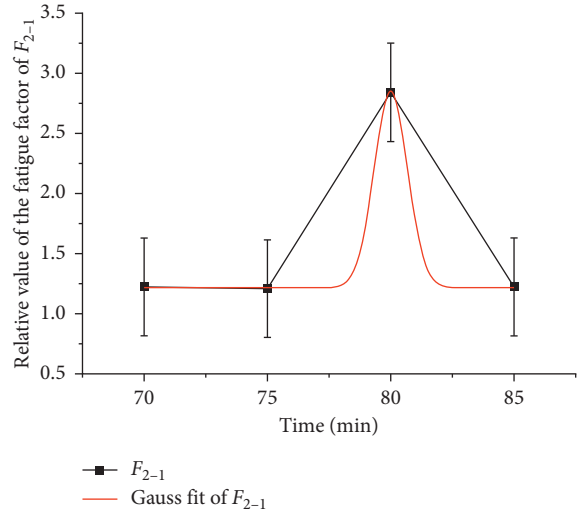


Figure 10: $F_{2\text{-}1}$ Gauss fitting.

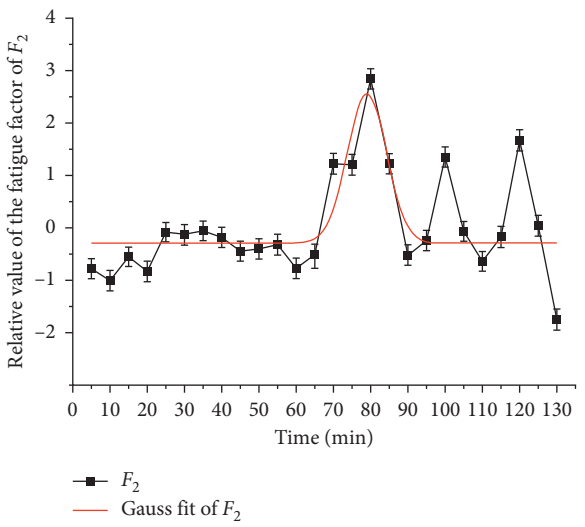


Figure 9: F_2 Gauss fitting.

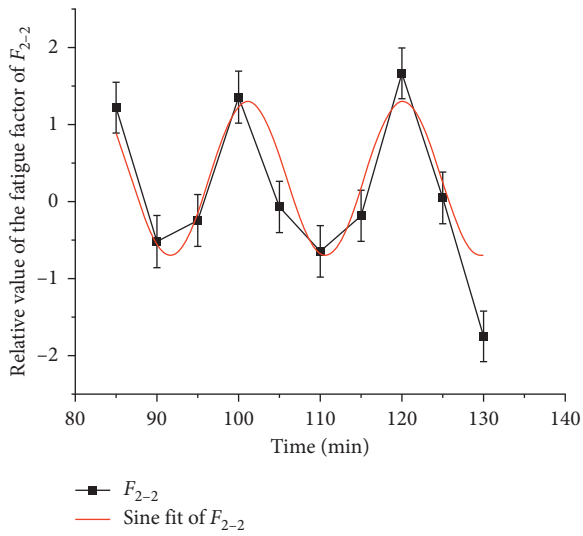


Figure 11: $F_{2\text{-}2}$ sine fitting.

but fatigue does not accumulate with time. $F_{2\text{-}2}$ sine fitting is shown in Figure 11.

In 70–85 min, F_2 showed a significant Gauss wave pattern. The significant peak appeared at 80 min, and the recovery period from the peak appeared to 15 min.

In 90–130 minutes, F_1 and F_2 showed a significant sine wave pattern. $F_{1\text{-}1}$ fluctuates periodically in 9 min and F_2 in 9.52 min. At the same time, the peak value of the two increases. It can be seen that although the driver's fatigue regulation ability can play a role, fatigue still increases with time. From the analysis of Section 2.2, it is known that the fit of F_1 and F_2 for the high-fatigue period is only 2.34 min, within 5 min, and the gap tends to decrease, so it is more reasonable to use the secondary fatigue factor F_2 to describe the rule of the fatigue transition stage and fatigue stage. The fatigue model of the grassland road driver is summarized as the following formula:

$$F = \begin{cases} 1.22 + 1.62e^{-1.06(t-80)^2}, 70 \le x \le 85, \\ \\ 0.3 + \sin\left[\left(\dfrac{t-1.37}{9.5}\right)\pi\right], 90 \le x \le 130. \end{cases} \quad (4)$$

3.4. The Classification Accuracy in Driving Fatigue Detection. Table 4 summarizes the classification accuracy in driving fatigue detection as achieved by the six classification models for each subject. The results show that our model classifier achieves the highest accuracy for all the subjects in classifying the fatigue vs. alert states. Our model achieves higher accuracy and yields lower variance than the RBF-ROLS + D-opt network. This shows that among these two RBF-based classifiers, the EEG driving fatigue detection proposed in this paper is a more accurate and robust classifier.

Compared to the other models, our model exhibits the best performance regardless of the specificity, sensitivity, and accuracy. In addition, the RBF-TLLH model significantly outperforms the RBF-ROLS + D-opt model in sensitivity, demonstrating the superiority of the proposed approach to detect driving fatigue. Compared to RBF-ROLS + D-opt, the proposed RBF-TLLH model achieves a slightly lower specificity, but a much higher accuracy and sensitivity.

4. Discussion

(1) The conclusion shows that the driver enters the fatigue stage obviously after 90 minutes. Although the fatigue factor increases and fluctuates greatly, the peak increase of the sine wave is small. Up to now, there is no clear division of light, medium, and heavy fatigue in the world. Because the driving task in this paper is only 130 minutes, there is no comparison between other types of fatigue, and it can only be defined as fatigue. As for the fatigue level, it is true that it is necessary to arrange long-term driving tasks for comparative study.

(2) 70–85 min is the transition period from the driver to fatigue. According to F_2 interpretation, the rule changes with the sine function after 80 min peak, and each cycle of fatigue and recovery is 9.52 min; according to F_1 interpretation, 70–85 min does not show fatigue, and it suddenly increases within 85–100 min, from which fatigue begins to accumulate. The derivative of F_1 in 85–110 min is analyzed, as shown in Figure 11; the derivative is 37 times of the derivative of 85 min, which is similar to the mutation rule characterized by F_2 in 70–85 min. The mutation period is 15 min at the same time, and F_2 is mutated before F_1. It is confirmed that the transition from the fatigue-free stage to fatigue stage needs to be completed by mutation and recovery of the fatigue factor, but it is impossible to determine that F_1 mutation lags behind F_2. The analysis of the change rule of the fatigue transition period is of guiding significance for the measures to prevent the formation of fatigue. Therefore, a large number of experimental samples are needed to study the change rule of the fatigue transition period.

(3) For the fatigue prediction after 100 min, the sine wave function of F_1 and F_2 has a high fitting degree, with a period difference of 0.52 min (see Figure 11). At first, the peak value corresponds to a time difference of 1/4 sine cycles, F_2 is earlier than F_1, but because of the long period of F_2, F_1 will reach the peak value at $t = 142.47$ min and F_2 will reach the peak value at $t = 142.21$ min with the increase of time; at this time, F_1 and F_2 change rule coincide. Later, if F_1 and $F2$ still meet the current sine function law, they will fluctuate according to the superposition sine wave law. The peak value of F_1 is 0.76 min earlier than that of F_2. According to the superposition principle of wave, at this time, fatigue will accumulate to a higher peak value in the time-axis direction, indicating the deepening of fatigue. Due to the limited time of driving tasks in this paper, the driving fatigue after 100 minutes cannot be determined, only the effect can be predicted, and large samples and simulation experiments are needed to arrange long-term driving tasks for further explanation.

5. Conclusions

In this paper, we propose an ant colony optimization algorithm based on mobile sink data collection in industrial wireless sensor networks. The fatigue factors are determined as $F_1 = 0.931\theta + (0.99\theta/\alpha + \beta) + (0.975\theta/\alpha) + (0.923\theta/\beta) - (0.189\alpha/\beta)$ and $F_2 = -0.079\theta + (0.081\theta/\alpha + \beta) - (0.164\theta/\alpha) + (0.367\theta/\beta) + (0.98\alpha/\beta)$.

F_1 and F_2 contain 96.42% EEG power information of fatigue changing with time during 130 min driving. Through regression analysis, F_1 and F_2 together show that the high incidence time of driving fatigue is 105 min and 125 min.

The change of driving fatigue with time can be divided into three stages: 5–65 min is the nonfatigue stage, 70–85 min is the fatigue transition stage, and 90–130 min is the fatigue stage. The variation of fatigue with time obeys function expression (4).

Data Availability

All the data used to support the findings of this study are available from the corresponding author upon request.

Ethical Approval

The authors confirm that all methods were carried out in accordance with the provisions of the Chinese Academy of Sciences on human-machine engineering experiments and strictly abide by the relevant provisions of the Chinese Traffic Law. All experimental protocols were approved by the Inner Mongolia Agricultural University Committee.

Consent

The authors confirm that informed consent was obtained from all subjects.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Y. L. Z. analyzed the data and wrote the manuscript. Y. L. Z. and S. L. Z. performed the experiments and participated in data analysis. S. L. Z. conceived and designed the whole project and provided funding support. Both authors read and approved the final manuscript.

Acknowledgments

This work was supported by the Natural Science Foundation of China (no. 51768057).

References

[1] P. Thiffault and J. Bergeron, "Fatigue and individual differences in monotonous simulated driving," *Personality and Individual Differences*, vol. 34, no. 1, pp. 159–176, 2003.

[2] P. Thiffault and J. Bergeron, "Monotony of road environment and driver fatigue: a simulator study," *Accident Analysis & Prevention*, vol. 35, no. 3, pp. 381–391, 2003.

[3] T. Oron-Gilad, A. Ronen, and D. Shinar, "Alertness maintaining tasks (AMTs) while driving," *Accident Analysis & Prevention*, vol. 40, no. 3, pp. 851–860, 2008.

[4] S. K. L. Lal and A. Craig, "A critical review of psychophysiological of driver fatigue," *Biological Psychology*, vol. 55, no. 3, pp. 173–194, 2001.

[5] G. S. Larue, A. Rakotonirainy, and A. N. Pettitt, "Driving performance impairments due to hypovigilance on monotonous roads," *Accident Analysis & Prevention*, vol. 46, 2011.

[6] D. L. Fell and B. Black, "Driver fatigue in the city," *Accident Analysis & Prevention*, vol. 29, pp. 463–469, 1997.

[7] V. E. Pollock, L. S. Schneider, and S. A. Lyness, "Reliability of topographic quantitative EEG amplitude in healthy late-middle-aged and elderly subjects," *Electroencephalography and Clinical Neurophysiology*, vol. 79, no. 1, pp. 20–26, 1991.

[8] L. Fletcher, L. Petersson, and A. Zelinsky, "Road scene monotony detection in a fatigue management driver assistance system," in *Proceeding of the IEEE Intelligent Vehicles Symposium*, pp. 484–489, Piscataway, NJ, USA, September 2005.

[9] G. S. Larue, A. Rakotonirainy, and A. N. Pettitt, "Driving performance impairments due to hypovigilance on monotonous roads," *Accident Analysis & Prevention*, vol. 43, no. 6, pp. 2037–2046, 2011.

[10] W. Li, Q. C. He, X. M. Fan, and Z. M. Fe, "Study on driving fatigue of grassland highway based on EEG signal," *Neuroscience Letters*, vol. 506, 2011.

[11] M. Abojaradeh, B. Jrew, and H. Ababsah, "Study on driver's driving characteristics and road traffic safety countermeasures," *Transportation and Engineering*, vol. 6, 2014.

[12] P. Phillip, J. Taillard, E. Klein et al., "Experimental Study on fatigue accumulation of long-distance bus drivers," *Journal of Psychosomatic Research*, vol. 55, 2003.

[13] H. Huang, F. Zhang, L. Yu et al., "Overview of non-contact pantograph-catenary arc detection based on image processing," in *Proceedings of the International Symposium for Intelligent Transportation and Smart City (ITASC)*, pp. 279–289, Shanghai, China, 2017.

[14] M. Gillberg, G. Kecklund, and T. Åkerstedt, "Sleepiness and performance of professional drivers in a truck simulator - comparisons between day and night driving," *Journal of Sleep Research*, vol. 5, no. 1, pp. 12–15, 1996.

[15] E. Wascher, S. Arnau, I. Gutberlet, M. Karthaus, and S. Getzmann, "Evaluating Pro- and Re-active driving behavior by means of the EEG," *Frontiers in Human Neuroscience*, vol. 12, no. May, pp. 1–8, 2018.

[16] X. Zhao, Z. Wei, Z. Li, Y. Zhang, and X. Feng, "Threshold research on highway length under typical landscape patterns based on drivers' physiological performance," *Discrete Dynamics in Nature and Society*, vol. 2015, Article ID 753732, 15 pages, 2015.

[17] S. Ahn, T. Nguyen, H. Jang, J. G. Kim, and S. C. Jun, "Exploring neuro-physiological correlates of drivers' mental fatigue caused by sleep deprivation using simultaneous EEG, ECG, and fNIRS data," *Frontiers in Human Neuroscience*, vol. 10, pp. 1–14, 2016.

[18] C. H. Bastien, C. Ladouceur, and K. B. Campbell, "EEG characteristics prior to and following the evoked K-Complex," *Canadian Journal of Experimental Psychology*, vol. 54, no. 4, pp. 255–265, 2000.

[19] S. K. L. Lal, A. Craig, P. Boord, L. Kirkup, and H. Nguyen, "Development of an algorithm for an EEG-based driver fatigue countermeasure," *Journal of Safety Research*, vol. 34, no. 3, pp. 321–328, 2003.

[20] T. Zhao, C. Qi, S. Zhu et al., "Study on multi index division of short-term driving fatigue of grassland highway," *Chinese Journal of safety Sciences*, vol. 26, no. 8, pp. 13–18, 2016.

[21] C. H. Chen, F. Song, F. J Hwang, and L. Wu, "A probability density function generator based on neural networks," *Physica A: Statistical Mechanics and its Applications*, vol. 541, Article ID 123344, 2020.

[22] G. D. J. Farias and J. V. Savian, "Integrated crop-livestock system with system fertilization approach improves food production and resource-use efficiency in agricultural lands," *Agronomy for Sustainable Development*, vol. 40, no. 39, 2020.

[23] A. Kumar and S. Bawa, "DAIS: dynamic access and integration services framework for cloud-oriented storage systems," *Cluster Computing*, vol. 23, pp. 1–20, 2020.

[24] M. L. Nandi, S. Nandi, H. Moya, and H. Kaynak, "Blockchain technology-enabled supply chain systems and supply chain performance: a resource-based view," *Supply Chain Management: International Journal*, vol. 25, no. 6, pp. 841–862, 2020.

[25] S. S. Sahoo, T. Nguyen, B. Veeravalli et al., "Multi-objective design space exploration for system partitioning of FPGA-based Dynamic Partially Reconfigurable Systems," *Integration*, vol. 67, pp. 95–107, 2019.

[26] T. Sharma and P. Balachandra, "Model based approach for planning dynamic integration of renewable energy in a transitioning electricity system," *International Journal of Electrical Power & Energy Systems*, vol. 105, no. FEB, pp. 642–659, 2019.

[27] H. Roekel and M. Steen, "Integration as unrealised ideal of ERP systems: an exploration of complexity resulting from multiple variations of integration," *Qualitative Research in Accounting and Management*, vol. 16, no. 4, 2019.

[28] L. Zhang and J. Pu, "An improved back propagation neural network in objects recognition," in *Proceedings of the 2011 IEEE International Conference on Automation and Logistics (ICAL) (IEEE)*, pp. 507–511, Chongqing, China, June 2020.

[29] C. Li and X. Liu, "An improved PSO-BP neural network and its application to earthquake prediction," in *Proceedings of the 2016 Chinese Control and Decision Conference (CCDC)*, pp. 3434–3438, Yinchuan: IEEE), Yinchuan China, May 2016.

[30] S. Chen, X. Hong, and C. Harris, "Sparse kernel regression modeling using combined locally regularized orthogonal least squares and D-optimality experimental design," *IEEE Transactions on Automatic Control*, vol. 48, pp. 1029–1036, 2003.

[31] Z. Ren, R. Li, B. Chen et al., "EEG-based driving fatigue detection using a two-level learning hierarchy radial basis function," *Frontiers in Neurorobotics*, vol. 15, Article ID 618408, 2021.

[32] F. Wang, S. Wu, J. Ping, Z. Xu, and H. Chu, "EEG driving fatigue detection with PDC-based brain functional network," *IEEE Sensors Journal*, vol. 21, no. 9, pp. 10811–10823, 2021.

[33] E. Karakose, M. T. Gencoglu, M. Karakose, O. Yaman et al., "A new arc detection method based on fuzzy logic using S-transform for pantograph-catenary systems," *Journal of Intelligent Manufacturing*, vol. 29, no. 4, pp. 839–856, 2018.

Hindawi
Journal of Healthcare Engineering
Volume 2021, Article ID 5709513, 16 pages
https://doi.org/10.1155/2021/5709513

Research Article

Privacy-Preserving Reversible Data Hiding for Medical Images Employing Local Rotation

Guo-Dong Su,[1,2,3] **Chia-Chen Lin**,[4] **and Chin-Chen Chang**[2]

[1]*School of Big Data and Artificial Intelligence, Fujian Polytechnic Normal University, Fuzhou 350300, China*
[2]*Department of Information Engineering and Computer Science, Feng Chia University, Taichung 40724, Taiwan*
[3]*Engineering Research Center for ICH Digitalization and Multi-Source Information Fusion (Fujian Polytechnic Normal University), Fujian Province University, Fuzhou 350300, China*
[4]*Department of Computer Science and Information Engineering, National Chin-Yi University of Technology, Taichung 41170, Taiwan*

Correspondence should be addressed to Chia-Chen Lin; ally.cclin@ncut.edu.tw

Received 4 May 2021; Accepted 14 June 2021; Published 2 July 2021

Academic Editor: Chi-Hua Chen

As digitalization becomes more common, patients' concerns about the leakage of private information, such as electronic medical record, are increasing, and those concerns motivated this case study of secure covert communication. Therefore, in this paper, a novel reversible data hiding method based on pixel rotation is proposed for medical images. Using pixel rotation, a state mapping model is constructed to represent the payload. More specifically, many intermediate states are derived from an image block, and each of them is used to form a one-to-one mapping relationship with a specific sequence of payload bits. In addition, to ensure the visual quality of stego-medical-images, the payload bits are only concealed in the regular blocks and the other blocks are unchanged. Moreover, the smoother regular image block will be priority to be used to embed the payload to enhance the visual quality of stego-medical-image. The experimental results showed that the stego-medical-images generated by the proposed reversible data hiding method have better visual quality with an average PSNR of 47.0307 dB, which is higher than that provided by some state-of-the-art methods.

1. Introduction

With the rapid development of telecommunication and computer sciences, informatization and digitization have been in many fields, such as banking and healthcare, and the result is e-banking and e-healthcare. For example, doctors use e-healthcare images for diagnosis and to study how to cure diseases. In addition, the e-healthcare platform was developed to provide a secure, covert channel of communication for the exchange of information among patients, doctors, and other relevant practitioners with the aim of ensuring that patients can access better quality medical service. However, several challenges will be encountered when electronic medical records (EMRs) are transmitted and shared in e-healthcare systems. Thus, one of the key challenges is how to ensure the security of the EMRs. Usually, the cryptography-based information security technique [1, 2] is considered to be an effective solution for this concern, but the encrypted information can easily attract the attention of attackers. Thus, the researchers in this area have focused their attention on irreversible data hiding [3–5] and reversible data hiding (RDH) [6–25] in medical images, which are techniques that can provide security and avoid attracting the attention of attackers.

RDH is a technique for processing multimedia signals that aims to hide secret data, such as EMR, into the original images, resulting in stego-images, in which the data are embedded. After removing the confidential data from the stego-images, the original images can be recovered completely. Due to its reversibility, RDH has been explored and

used extensively in many applications, and detailed discussions of some of these applications are presented in Section 2.

In recent years, many RDH methods have been presented for medical images to address patients' concerns about the leakage of their private data. In 2013, Huang et al. [26] expanded the histogram shifting (HS) framework-based RDH technique to medical images that have high bit depths. For each medical image block, a difference-histogram is constructed by using different values, and it is used to carry secret data. This method satisfies 6 criteria, three of which are the inclusion of a free location map and the ability to adjust both embedding capacity (EC) and peak signal-to-noise ratio (PSNR). To improve EC, Kelkar et al. [27] presented two innovative variations of the traditional HS technique. The first technique uses HS to embed the secret data into nonoverlapping image blocks, and the second technique separates images into the region of interest (ROI) and the region of noninterest (NROI), and it only embeds the secret data in NROI. Their method leaves the medical information intact and achieves a high PSNR (above 45 dB). In the method of [28], the pixel-to-block conversion technique is used to predict the interpolated pixels in the cover image. Subsequently, the EMR related data are embedded into the LSBs of the pixels in the 4×4 medical image block. Instead of using the pixel-to-block conversion technique, the method in [29] uses another technique, i.e., the rhombus mean interpolation technique, as an effective alternative to interpolation for the generation of the cover image to ensure the reversibility of the medical images. Both techniques obtain a high EC and acceptable quality of the stego-medical-image. As to the method in [30], Gao et al. proposed an automatic contrast enhancement algorithm to achieve larger EC and better quality of the stego-medical-images. First, given the characteristics of the medical image, ROI and NROI are separated by using an adaptable threshold detector. Then, the ROI's histogram is stretched to enlarge the EC of ROI. For NROI, the replacement of the LSBs is applied directly, if necessary, to conceal the secret data. Gao et al.'s scheme does well in EC, but there is a significant visual difference between the stego-medical-image and its original version.

To better balance the EC and the quality of the stego-medical-image, in this paper, we propose a novel RDH method based on a local pixel rotation. The main contributions of this paper are as follows:

(1) Propose an RDH based on pixel rotation. Using the rotation mechanism, many intermediate states are produced and can be used to represent the payload bits. This is quite different from other RDH methods.

(2) Achieve a good balance between the EC and the quality of the stego-medical-image. Compared to some recent works, the proposed RDH method has an excellent performance in both EC and PSNR.

(3) Provide secure covert communication.

The rest of this paper is organized as follows. Section 2 provides a review of RDH, and the proposed RDH method is presented in Section 3. The experimental results and analyses are presented in Section 4, and Section 5 provides our conclusions.

2. Related Work

The focal points of RDH are (1) how to improve the payload capability of stego-images and (2) how to enhance the quality of the stego-images. However, there is a trade-off problem between the EC and the quality of the stego-images. Most common RDH methods essentially are classified into four types, i.e., (1) difference expansion-based (DE-based) techniques [6–10], (2) prediction error expansion-based (PEE-based) techniques [11–15], (3) HS-based techniques [16–20], and (4) other techniques [21–25].

It is our understanding that the first DE framework-based RDH method was put forward by Tian [6] in 2002. In his scheme, every two adjacent pixels are compared and computed to detect the difference errors, and some of them are selected for difference expansion along with secret data embedding. Alattar [7] extended difference errors into a vector that consists of k distinct pixels to generate more available difference values to be used for embedding data. To improve EC, Hu et al. [8] proposed a novel DE embedding algorithm based on a dynamically expandable difference search and selection mechanism, which uses both the vertical and horizontal difference errors for embedding data. To enhance the quality of stego-images, a novel, multiple-based, RDH method based on edge prediction and lossless joint photographic experts group (JPEG-LS) pixel value prediction was presented by Wu et al. [9]. This method conceals more secret data into smooth areas rather than complex areas, thereby achieving better quality stego-images. In addition, Wang et al. [10] studied a robust RDH method using significance-bit-difference expansion. This method can be used against unintentional attacks, such as image compression and sometimes unavoidable addition of random noise, which is below a certain level and does not change the content of an image.

The PEE framework-based RDH method was developed to enhance the EC. In 2004, Thodi and Rodriguez [11] explored an algorithm that uses the correlation inherent among neighboring pixels in a local image area using a predictor. Then, the secret data are inserted by expanding the prediction error values. In 2009, Tseng and Hsieh [12] presented an RDH method based on various predictors that is capable of providing a great EC without creating any noticeable distortion. Ou et al. [13] considered the correlations among the prediction errors and used them to design a pairwise prediction error expansion for embedding secret data, and their method provided a high EC [13]. In the method of [14], the joint use of the pixel-value-ordering and PEE was developed to present a high-fidelity RDH method. Earlier, Li et al. [14] also conducted block selection before embedding the payload. Hence, this method guarantees that the PSNR of the stego-image will be greater than 51.14 dB. In 2020, Chang trained the generative adversarial networks to predict bit-planes that have been used to hide secret data.

This work achieved a good performance, and it clarified how deep learning can breathe new life into older RDH methods.

The aim of the HS framework-based RDH method is to embed secret data into the peak bin that has the maximum frequency in its histogram distribution. Ni et al.'s method [16] is one of the key works in HS-based RDH methods. In [16], the zero or the minimum peaks of the image's histogram are used to embed more secret data with only slight modifications of the pixels. Next, Lee et al. [17] explored the spatial correlation of natural images to generate a more appropriate histogram, i.e., the difference-histogram. Since the difference-histogram has a much higher peak-bin, it is more suitable for RDH. Hong et al.'s HS-based RDH method [18] focuses on the modification of the prediction-error histogram to create more vacant positions for embedding secret data. Their scheme achieves a good EC and guarantees the PSNR to be above 48 dB. Jia et al. [19] considered that most of the existing RDH methods do not fully take into account that the texture of natural images influences the embedding distortion. Thus, they proposed the HS-based RDH method to reduce the invalid shifting of pixels. In this way, their method provides better-quality stego images. Also, Peng et al. [20] proposed a novel HS framework-based RDH method by a specific secret data coding strategy and a multisegment left and right histogram-shifting mechanism. Experiments showed that it achieved good performance in EC and PSNR.

In this paper, we propose an RDH method using pixel rotation to improve the EC and enhance the visual quality of medical images. More specifically, we constructed the mapping relationship between the pixel sequences and the to-be-embedded bits of the payload. This is quite different from conventional RDH methods, and the details are described in Section 3.

3. Proposed Scheme

In this paper, we present a novel RDH method based on pixel rotation to further enhance the EC and visual quality of stego-medical-image for medical images. Our method consists mainly of two procedures, i.e., (1) the payload embedding procedure and (2) the payload extraction procedure. However, before that, the definition of state mapping and the sequence selection are introduced.

3.1. Definition of State Mapping. Considering that the pixel values from a local block in a medical image are either the same or similar to each other, a rotation-based state mapping mechanism is defined and later used for embedding the payload. Let us assume that a sequence derived from a local medical image block is

$$P = (p_1, p_2, \ldots, p_k, \ldots, p_K), \tag{1}$$

where K represents the total number of elements in the current sequence, and p_k is its k^{th} element, and all ranged in $[0, V_{\max}]$. Next, we transfer P into $P\prime$ by

$$\begin{bmatrix} p_1' \\ p_2' \\ \ldots \\ p_k' \\ \ldots \\ p_K' \end{bmatrix} = \begin{bmatrix} p_1 \\ p_2 \\ \ldots \\ p_k \\ \ldots \\ p_K \end{bmatrix} + \varepsilon \begin{bmatrix} e_1 \\ e_2 \\ \ldots \\ e_k \\ \ldots \\ e_K \end{bmatrix}, \tag{2}$$

where p_k' is the k^{th} element of $P\prime$ and ε is a positive integer. The value of e_k is determined by

$$e_k = \begin{cases} 1, & \text{if } k = \text{pos}, \\ 0, & \text{others}, \end{cases} \tag{3}$$

where pos is the highlighted position in P, and it is in the range of $[1, K]$.

After that, in most cases, the p'_{pos} is expected to be the unique maximum value in P'. When P' is prepared, K distinct states Θ^k $(1 \le k \le K)$ can be derived using the left-oriented rotation. An illustration of the state mapping definition is demonstrated in Table 1. For each row, a given state, Θ^k, has a unique maximum value, p'_{pos}. In addition, for each Θ^k, the p'_{pos} always is located at a distinct position (i.e., in the 3rd column), which can be used to carry a $\log_2 K$ bit payload s_1. Given a pos, the value of a can be computed by

$$a = \begin{cases} 1, & \text{if } k > \text{pos}, \\ 0, & \text{others}, \end{cases} \tag{4}$$

where k varies from 1 to K.

To demonstrate the definition of proposed state mapping better, Table 2 gives an example of a specific case. Here, the pos is set to 4, and the initial sequence $[p_1', p_2', p_3', p_4']$ is [20, 20, 20, 21]. First, we rotate the initial sequence, i.e., 0, 1, 2, and 3 times separately, to generate four states, i.e., Θ^1, Θ^2, Θ^3 and Θ^4. We will find, for each state, that the position, where the maximum value is located, is different from that of the other states. Therefore, each state can be used to represent a $\log_2 4$ bit payload.

3.2. Sequence Selection. For a given sequence, it is not always suitable to carry the payload. Thus, we classified the sequences into four types, i.e., the regular sequence, the singular sequence, the unusable sequence, and the overflow sequence. The definitions of those four types can be described as follows:

(1) Regular sequence: the sequence has a unique maximum value, and the position of the maximum value is equal to pos

(2) Singular sequence: the sequence has a unique maximum value, but the position of this maximum value is not equal to pos

(3) Unusable sequence: there are two or more values that are maximum in the sequence

(4) Overflow sequence: the value of the pos$^{\text{th}}$ element in sequence exceeds $V_{\max}$

Table 1: Illustration of the state mapping definition (rotation left).

	States Θ^k $(1 \le k \le K)$	Position of maximum value	Payload s_1
Θ^1	$[p_1', p_2', \ldots, p_{\text{pos}}', \ldots, p_K']$	$\text{pos} + K\ a$	0
Θ^2	$[p_2', \ldots, p_{\text{pos}}', \ldots, p_K', p_1']$	$\text{pos} - 1 + K\ a$	1
	...		
Θ^k	$[p_k', \ldots, p_{\text{pos}}', \ldots, p_K', p_1', \ldots, p_{K-1}']$	$\text{pos} - (k-1) + K\ a$	$k-1$
	...		
Θ^K	$[p_K', p_1', p_2' \ldots, p_{\text{pos}}', \ldots, p_{K-1}']$	$\text{pos} - (K-1) + K\ a$	$K-1$

Table 2: An example of constructing the state mapping.

	States Θ^k	Position of p_{pos}'	Payload s_1
Θ^1	[20, 20, 20, 21]	4	0
Θ^2	[20, 20, 21, 20]	3	1
Θ^3	[20, 21, 20, 20]	2	2
Θ^4	[21, 20, 20, 20]	1	3

To explain the different types of sequences clearly, examples are provided in Table 3, where the pos and $V_{\max}$ are set to 4 and 127, respectively.

3.3. Payload Embedding. For a given 8-depth $H \times W$ medical image I, all pixels are divided into N nonoverlapping image blocks with the sizes of $h \times w$, denoted as $I = \{I^t | 1 \le t \le N\}$, where $N = (H/h) \times (W/w)$. For simplicity, we represent $I^t = (I_1^t, I_2^t, \ldots, I_K^t)$, where $K = h \times w$. Before embedding the payload, we should separate all image blocks into four types according to the definition in Subsection 3.2, and the details can be described as follows:

Step 1: process I^t into $P = (p_1, p_2, \ldots, p_k, \ldots, p_K)$ using

$$p_k = \text{rounding}\left(\frac{I_k^t}{2}\right), \tag{5}$$

where $1 \le k \le K$ and the function rounding represents the rounding operation. p_k is ranged in [0, 127].

Step 2: given a ε and pos, process P into $P\prime = (p_1', p_2', \ldots p_k', \ldots, p_K')$ using (2) and (3).

Step 3: identify the sequence $P\prime$ as one of four types.

Step 4: determine I^t as one of four types. More specifically, if P' is identified as a regular sequence, the I^t is determined as a regular block; if P' is identified as a singular sequence, the I^t is determined as a singular block; if P' is identified as an unusable sequence, the I^t is determined as an unusable block; if P' is identified as an overflow sequence, the I^t is determined as an overflow block.

Examples of the different types of image blocks are given in Figure 1, where $\varepsilon = 1$ and pos = 4.

After the separation of the image blocks, all regular blocks are roughly classified into various priority levels according to the standard deviations of their corresponding P's. For a given $P\prime = (p_1', p_2', \ldots p_k', \ldots, p_K^v)$, its various priority levels, namely, L, are defined as

Table 3: Examples of different types of sequences (pos = 4).

Sequences	Situations	Block types
[20, 20, 20, 21]	$p_{\text{pos}}' = p_4'$ is unique maximum value	Regular
[20, 20, 21, 20]	$p_{\text{pos}}' = p_4'$ is not unique maximum value	Singular
[20, 19, 20, 21]	$p_{\text{pos}}' = p_4'$ is unique maximum value	Regular
[20, 20, 21, 21]	$p_{\text{pos}}' = p_4'$ is not unique maximum value	Unusable
[127, 127, 126, 128]	$p_{\text{pos}}' = p_4' = 128 > V_{\max} = 127$	Overflow

$$L = \text{rounding}\left(\sqrt{\frac{\sum_1^K (p_k' - \overline{p}')^2}{K}}\right), \tag{6}$$

where $\overline{p}'$ is the mean of elements in P'. The image block with the lower value of L will be the priority to be used to embed the payload for the aim of enhancing the high stego-medical-image quality. It is noted that the priority level of a regular image block will be reserved perfectly before and after payload embedding. This is because the proposed rotation based embedding strategy only changes the order of elements in P' rather than changing their values.

After the classification of the regular image blocks, the payload can be embedded. First, to ensure reversibility, a location map (LM) should be used to differentiate between a regular block and a singular block. This is because a regular block may be transferred into a singular block after the payload is embedded. More specifically, if an image block is identified as a regular block, we mark it with bit "1" in LM; if an image block is identified as a singular block or unusable block or overflow block, we mark it with bit "0" in LM. Then, the LM is processed further using the quadtree-based compression technique. Next, the parameters ε, pos, h, w, |LM| and the compressed LM are concatenated to form the auxiliary information, where |LM| is the length of the compressed LM. The detailed analysis of LM will be given in Subsection 4.2.2. For simplicity, we used 4 bits, 4 bits, 4 bits, 4 bits, and 16 bits in our experiments to store the values of ε, pos, h, w and |LM|, respectively. Finally, the auxiliary information was embedded into the LSB of the first 32 + |LM| pixels. Certainly, before embedding the auxiliary information, the original 32 + |LM| LSBs are recorded and concatenated with secret data to form the payload. The detailed

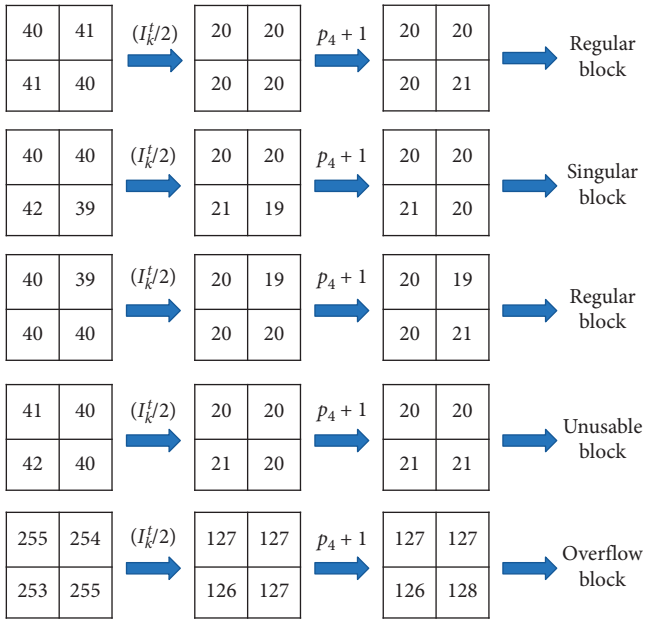


FIGURE 1: Examples of the different types of image blocks ($\varepsilon = 1$ and pos = 4).

procedure of embedding payload into regular blocks is described as follows:

Step 1: count the embedding capacity by multiplying the $\log_2 K$ and the number of regular blocks. If the length of the payload is equal to or less than the embedding capacity, the embedding procedure continues; otherwise, it is terminated.

Step 2: process all regular image blocks and process them into the corresponding sequence P'. Calculate the priority level for each sequence P' using equation [6].

Step 3: take $\log_2 K$ bits from the head of the payload and convert them into a decimal value, dv.

Step 4: read in an unused sequence P' with the highest priority level. Then, rotate it left dv times and derive the decided sequence P^* where s_k is equal to dv.

Step 5: transfer P^* to the stego-image block using the inverse process of (2).

Step 6: repeat Steps 3 to 5 until all regular blocks have been processed, or all bits in the payload have been embedded.

After that, the stego-medical-image IS embedded with payload is generated and sent to the recipient. Figure 2 demonstrates two examples of embedding payload bits into regular blocks, where $\varepsilon = 1$ and pos = 4. In Figure 2(a), the image block is [40, 41; 41, 40], and the to-be-embedded payload bits are "01." First, we can obtain the initial sequence [20, 20, 20, 21] (see Figure 1). Since the decimal value of payload bits is dv = 1, so this sequence is rotated left 1 time. The rotated sequence is decided as [20, 20, 20, 21] and put back to generate the stego-image block. Figure 2(b) illustrates the embedding of payload bits "11" into a regular image block [40, 39; 40, 40]. In the same way, the initial sequence [19, 20, 20, 21] is obtained and rotated left 3 times. The rotated result [19, 20, 20, 21] is locked and used to generate the stego-image block [42, 41; 38, 40]. Finally, the payload bits "11" are carried out.

Besides, we also can see from Figure 2 that the distortion in Figure 2(a) is small, and the distortion in Figure 2(b) is relatively large. Overall, the distortion caused by embedding payload bits into medical images using our rotation based RDH will be affected by two aspects: (1) the more complex the image block is, the larger the distortion is; on the contrary, the smoother the image block is, the smaller the distortion is; (2) for the image block, whose complexities are similar or the same, the distortion is relevant to the times of rotation. Thus, in this paper, the smoother regular image block will be priority to be used to embed the payload to ensure the considerable visual quality of stego-medical-image.

3.4. Payload Extraction and Image Recovery

3.4.1. State Mapping Definition in the Extraction Process. Assume that the initial sequence derived from the image block in a stego-medical-image is $PS' = (ps_1', ps_2', \ldots ps_k', \ldots, ps_K')$. For simplicity, we also suppose that the ps_k' has the maximum value among this sequence. Next, the definition of state mapping used in the extraction process can be described as Table 4. An example of constructing the state mapping is given in Table 5, where $PS' = [20, 20, 21, 20]$ and pos = 4. Obviously, the sequence, whose position of the maximum value equals to pos, is the original sequence, and the corresponding φ is the to-be-extracted payload.

3.4.2. Payload Extraction and Image Recovery. When the recipient holds the stego-medical-image *IS*, s/he can implement the extraction of payload and the recovery of the original medical image *I*. Firstly, s/he derives the auxiliary information from the every LSB of the front 32 + |LM| pixels and parses out the parameters ε, pos, h, w, |LM| and the compressed LM. Using quadtree-based decompression, the location map is obtained that indicates which block is the regular one. Next, all pixels in *IS* are divided into N non-overlapping image blocks with the sizes of $h \times w$, denoted as $IS = \{IS^t | 1 \le t \le N\}$, where $N = (H/h) \times (W/w)$. For simplicity, we represent $IS^t = (IS_1^t, IS_2^t, \ldots, IS_K^t)$, where $K = h \times w$. For the regular image blocks, the payload extraction and image recovery are conducted as follows:

Step 1: process every regular image block IS^t into $PS = (ps_1, ps_2, \ldots, ps_k, \ldots, ps_K)$ using

$$ps_k = \text{rounding}\left(\frac{IS_k^t}{2}\right), \tag{7}$$

where $1 \le k \le K$.

Step 2: according to the parameters ε and pos, process all *PS* into the corresponding sequence *PS′* using (2) and (3).

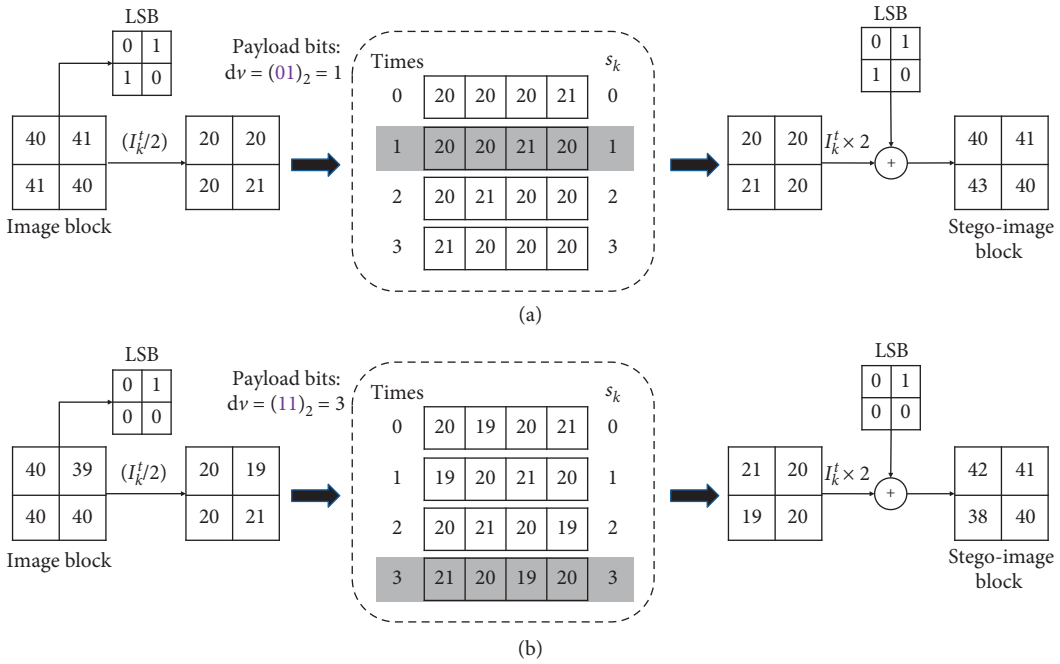


Figure 2: Examples of embedding payload bits into regular blocks ($\varepsilon = 1$ and pos = 4).

Table 4: Illustration of the state mapping definition used in extraction process (rotation right).

	States Θ^k $(1 \le k \le K)$	Position of maximum value	Payload s_1
Θ^1	$[ps'_1, ps'_2, \ldots, ps'_k, \ldots, ps'_K]$	k	0
Θ^2	$[ps'_K, ps'_1, ps'_2, \ldots, ps'_k, \ldots, ps'_{K-1}]$	$k+1$	1
	...		
Θ^k	$[p'_k, \ldots, p'_{\text{pos}}, \ldots, p'_K, p'_1, \ldots, p'_{k-1}]$	pos	φ
	...		
Θ^K	$[p'_K, p'_1, p'_2 \ldots, p'_{\text{pos}}, \ldots, p'_{K-1}]$	$k-1$	$K-1$

Table 5: An example of constructing the state mapping used in extraction process (rotation right).

	States Θ^k	Position of maximum value	Payload s_1
Θ^1	[20, 20, 21, 20]	Payload s_1	0
Θ^2	[20, 20, 20, 21]	4 (pos)	1 (φ)
Θ^3	[21, 20, 20, 20]	1	2
Θ^4	[20, 21, 20, 20]	2	3

Step 3: calculate the priority level of all sequences PS' using equation [6].

Step 4: read in an unused sequence PS' with highest priority level and rotate it right φ times, making the value of ps'_{pos} of a sequence its maximum value. Put this sequence back to generate the original medical image block, and convert φ into binary representation to form the payload bits.

Step 5: repeat Steps 1 to 4 until all regular blocks have been processed.

Finally, the payload is extracted completely, and the approximate original medical image is reconstructed. Next, the first 32 + |LM| bits of payload are cut out and used to replace the LSB of the front 32 + |LM| pixels in this approximate original medical image. By now, the secret data has been gained, and the original medical image has been recovered in a lossless way. For ease of understanding, two examples of extracting the payload bits and recovering the image blocks are illustrated in Figure 3.

4. Experimental Results

In this section, the results of extensive experiments are provided to evaluate the performance of the proposed RDH method. There are six 512×512 medical images, i.e., "Kindey_A," "Kindey_B," "Brain_A," "Brain_B," "Sketeton_A," and "Sketeton_B," used as test images, and they are shown in Figure 4. In our experiments, several statistical metrics, such as PSNR (Peak signal-to-noise ratio) [31] and EC, are measured for the performance evaluation.

PSNR is a metric that can evaluate the visual quality of the stego-medical-image, and it is defined as

$$\text{PSNR} = 10 \times \log_{10}\left(\frac{255 \times H \times W}{\sum_{r=1}^{H}\sum_{c=1}^{W}\left(I_{r,c} - IS_{r,c}\right)^2}\right), \quad (8)$$

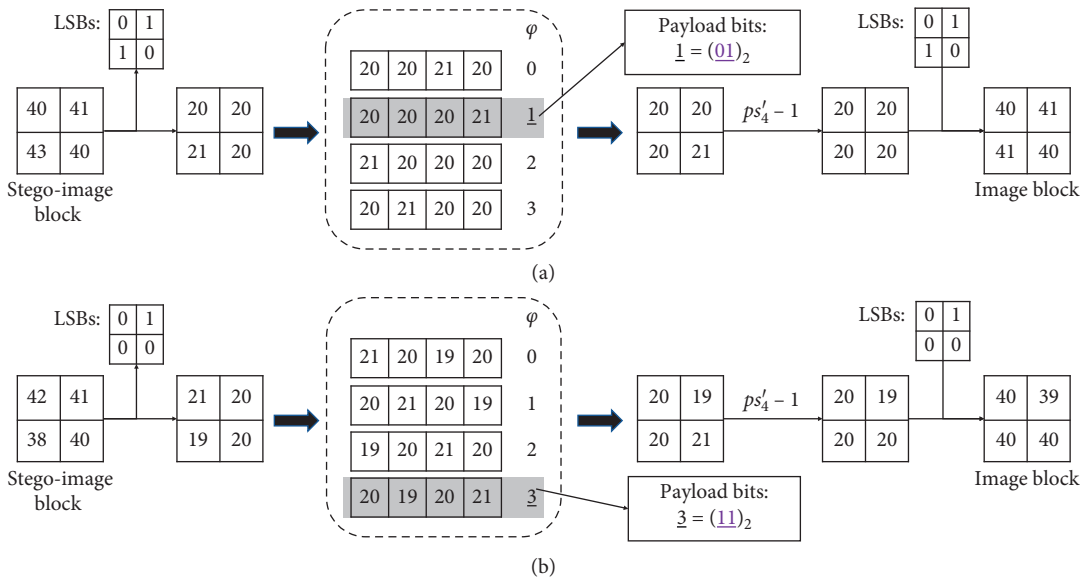


Figure 3: Examples of extracting the payload bits and recovering the image blocks ($\varepsilon = 1$ and pos = 4).

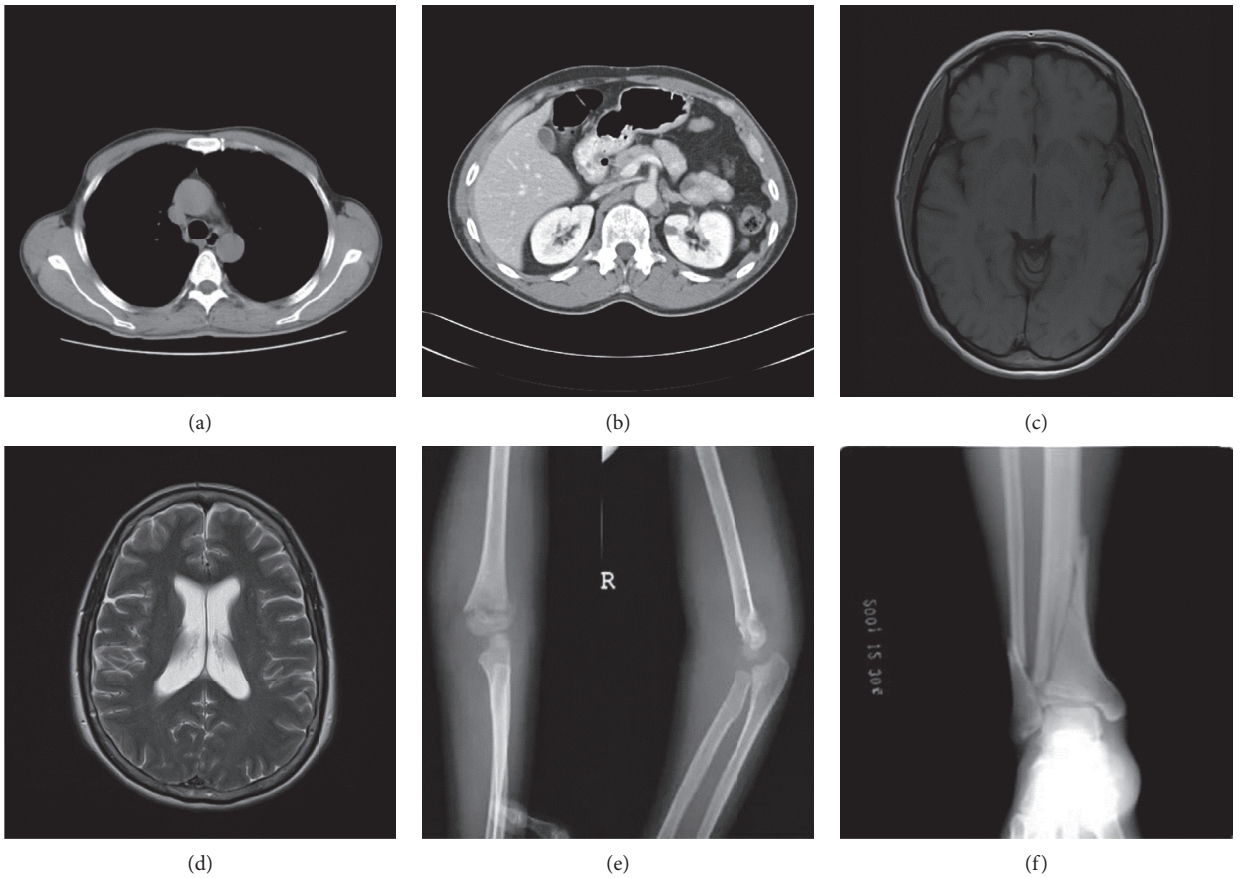


Figure 4: Six test medical images. (a) Kidney_A; (b) Kidney_B; (c) Brain_A; (d) Brain_B; (e) Sketeton_A; (f) Sketeton_B.

where $I_{r,c}$ and $IS_{r,c}$ are the pixel values that are located on the r^{th} row and c^{th} column in images I and IS, respectively. In general, the higher the value of PSNR is, the better the visual quality of the stego-medical-image is.

EC is a metric that is used to evaluate the ability of the stego-medical-image to carry secret data, and it is defined as follows:

$$\begin{aligned} \text{EC} &= \frac{H \times W}{h \times w} \times \log_2 (h \times w), \\ \text{pure EC} &= \text{EC} - (|\text{LM}| + 32). \end{aligned} \tag{9}$$

4.1. Security Analysis. To prove that the proposed RDH method can provide the imperceptibility of secret data and stego-medical-image, security analyses, including pixel value difference (PVD) histogram [32], Shannon entropy, the number of pixels change rate (NPCR) [33], and the unified average changing intensity (UCAI) [33], were used to evaluate the stego-medical-image with full payload under $\varepsilon = 1$, pos = 4, $h = 2$ and $w = 2$.

4.1.1. PVD Histogram. The PVD histogram is an indicator that can provide the degree of difference between every two adjacent pixels in an image. Generally speaking, the spatially anomalous distribution of the PVD histogram leaks the existence of secret data, and it even can be used to get a rough estimation of the amount of secret data. Figure 5 illustrates the variation tendency of the PVD histogram curves for six couples of images. As can be seen, the gaps of the PVD histogram curves between the original image, I, and its corresponding stego-image, IS, are close to each other, which implies that the proposed RDH method can resist the steganalysis of the PVD histogram.

4.1.2. Shannon Entropy. Shannon entropy is a metric that can be used to evaluate the divergence of a stego-image from its original version. Generally speaking, two images coincide if their Shannon entropies are close to each other, and the system is considered to be perfectly secure. Figure 6 shows the Shannon entropy curves of the image I and the stego-image IS, where image IS is embedded with different amounts of the payload. It is easy to observe that the gaps for the curves between the images I and IS are extremely close to each other. Thus, it was concluded that the proposed RDH method is extremely secure.

4.1.3. Differential Attack Analysis. In addition, two measurements, i.e., NPCR and UACI, are used to provide a quantitative analysis with respect to the changes from the original image to the stego-image. NPCR is used to determine the rate at which the pixels changed for a stego-image caused by payload embedding, and it has the maximum theoretical value of 1. UACI is used to indicate the average intensity of the change of pixel values, and it has a theoretical value of 0.3346. The smaller the NPCR and UACI are, the slighter the changes in the pixel are. The UACI and NPCR are defined as follows:

$$\begin{aligned} \text{UACI} &= \frac{1}{H \times W} \times \sum_{r=1}^{H} \sum_{c=1}^{W} \frac{\left| I_{r,c} - IS_{r,c} \right|}{255} \times 100\%, \\ \text{NPCR} &= \frac{1}{H \times W} \times \sum_{r=1}^{H} \sum_{c=1}^{W} D_{r,c} \times 100\%, \\ D_{r,c} &= \begin{cases} 0, & \text{if } I_{r,c} = IS_{r,c}, \\ 1, & \text{otherwise.} \end{cases} \end{aligned} \tag{10}$$

Figure 7 shows the variation tendency of curves of NPCR and UACI for the stego-images that are produced for six medical images under various ECs when ε is set to range from 1 to 4. Obviously, in most cases, the UACI and NPCR values of the six test images are far away from their theoretical maximum values. This indicates that the proposed RDH method can effectively against differential attacks.

4.2. Performance Analysis. In this section, the proposed RDH method is analyzed based on its performance in terms of PSNR and EC. In these experiments, the parameters $\varepsilon = 1$, pos = 4, $h = 2$ and $w = 2$ are the default values unless specified elsewhere.

4.2.1. Visual Quality of the Stego-Images. After embedding the payload into the original medical image, it is always expected that the stego-image is the same or similar to the original version. Figure 8 shows six stego-images when the EC of 60000 bits is achieved. It is obvious that the visual quality of six stego-medical-images is good, and it is difficult for the human eyes to distinguish the stego-medical-images from the original medical images.

4.2.2. Analysis of LM. As analyzed in Subsection 3.3, the LM is used to differentiate the type of each image block. Because most image blocks in an image will be identified as the regular block, thus, it is expected that the LM is a sparse matrix, which can be effectively compressed into the reduced version with fewer bits. Table 6 gives the ratio of the regular block and the size of |LM| with/without compression for six medical images. As can be seen, the ratio of the regular block reaches 76.0801% on average. What is more, the size of |LM| can be reduced from 65535 bits to 24240.67 bits. This positively contributes to achieve a high embedding capacity.

4.2.3. PSNR and EC. Commonly, it is considered that it is difficult for the human vision system to detect the distortion in images, as long as the PSNR is greater than 30 dB. Figure 9 demonstrates the graphs in terms of PSNR and EC for six images when ε varies from 1 to 4. It can be seen that the proposed RDH scheme can achieve a good EC that exceeds 10×10^4 bits. Also, it is not surprising that the PSNR decreases as EC increases since the more payload bits are embedded onto the stego-medical-image. In addition, since the more regular blocks can be used to carry payload bits, it

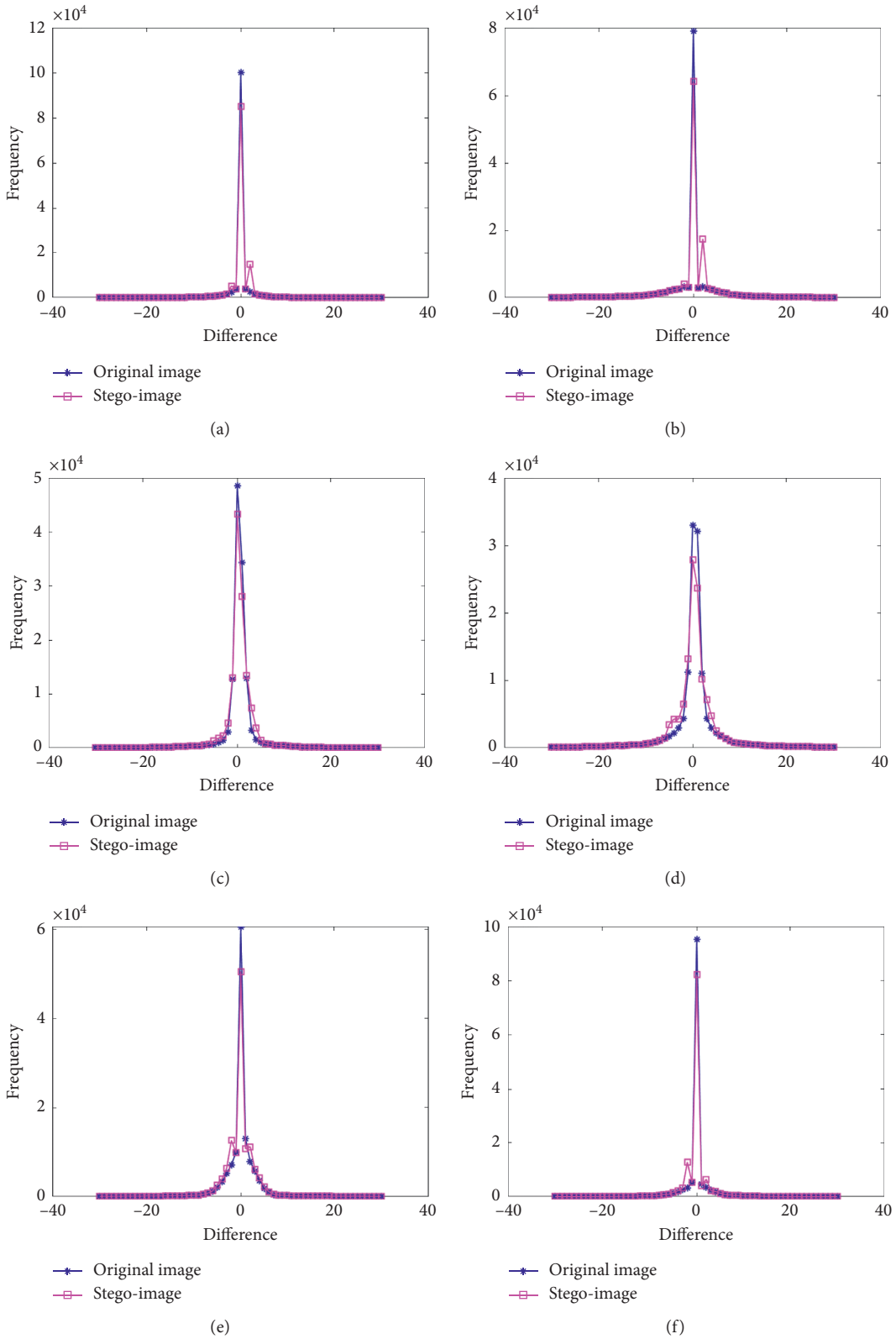


Figure 5: Six PVD histograms of images (I) and *IS*. (a) Kidney_A; (b) Kidney_A; (c) Brain_A; (d) Brain_B; (e) Sketeton_A; (f) Sketeton_B.

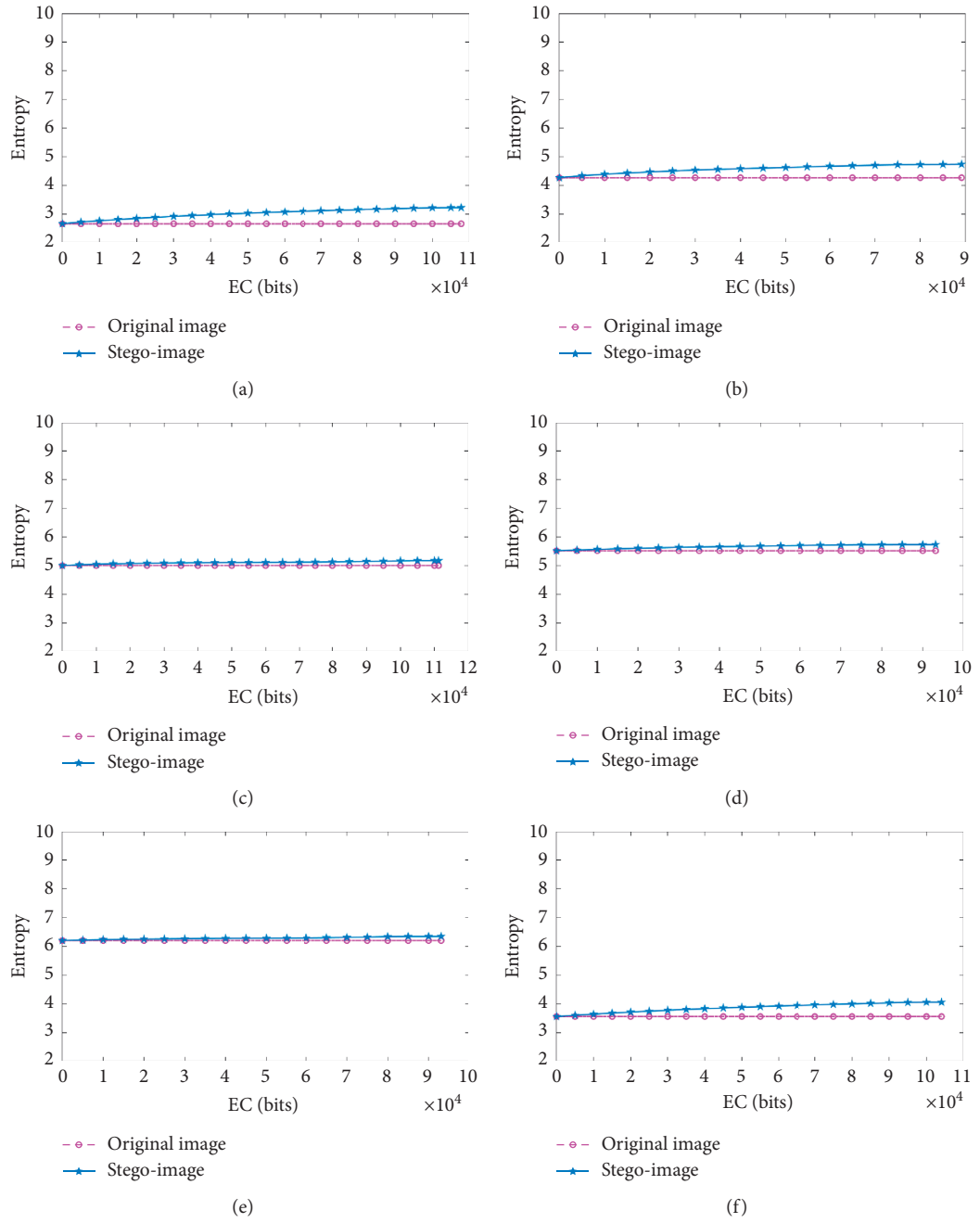


Figure 6: Shannon entropy curves of images (I) and *IS*. (a) Kidney_A; (b) Kidney_B; (c) Brain_A; (d) Brain_B; (e) Sketeton_A; (f) Sketeton_B.

can be seen that the EC increases as ε increases. However, for a given EC, the PSNR obviously decreases when the value of ε is large. This is mainly due to the fact that the larger value of ε that is selected will lead to significant changes in pixel values during the procedure of embedding the payload.

Also, we conducted experiments using six medical images to determine the trends of PSNR and EC when the value of pos was set from 1 to 4, and Figure 10 shows the results. Figure 10 clearly indicates that, in some cases, both the trend of the PSNR curves and EC are almost consistent with each other for different values of pos. The details in the difference of PSNR are demonstrated in the subgraph embedded inside Figure 10. For images "Brain_A" and "Brain_B," the PSNR values obtained under pos = 4 were a little lower than the values of the others, as shown in the front part of the curve. But, in general, the EC gained under pos = 4 is much higher than that of other cases. Thus, in most of our experiments, the pos = 4 is selected to achieve a good performance.

In addition, Table 7 lists the pure ECs and PSNRs obtained on six medical images for different sizes of blocks. In theory, the larger the size of the block is, the more bits the block carries. However, when the size of the block is 4×4,

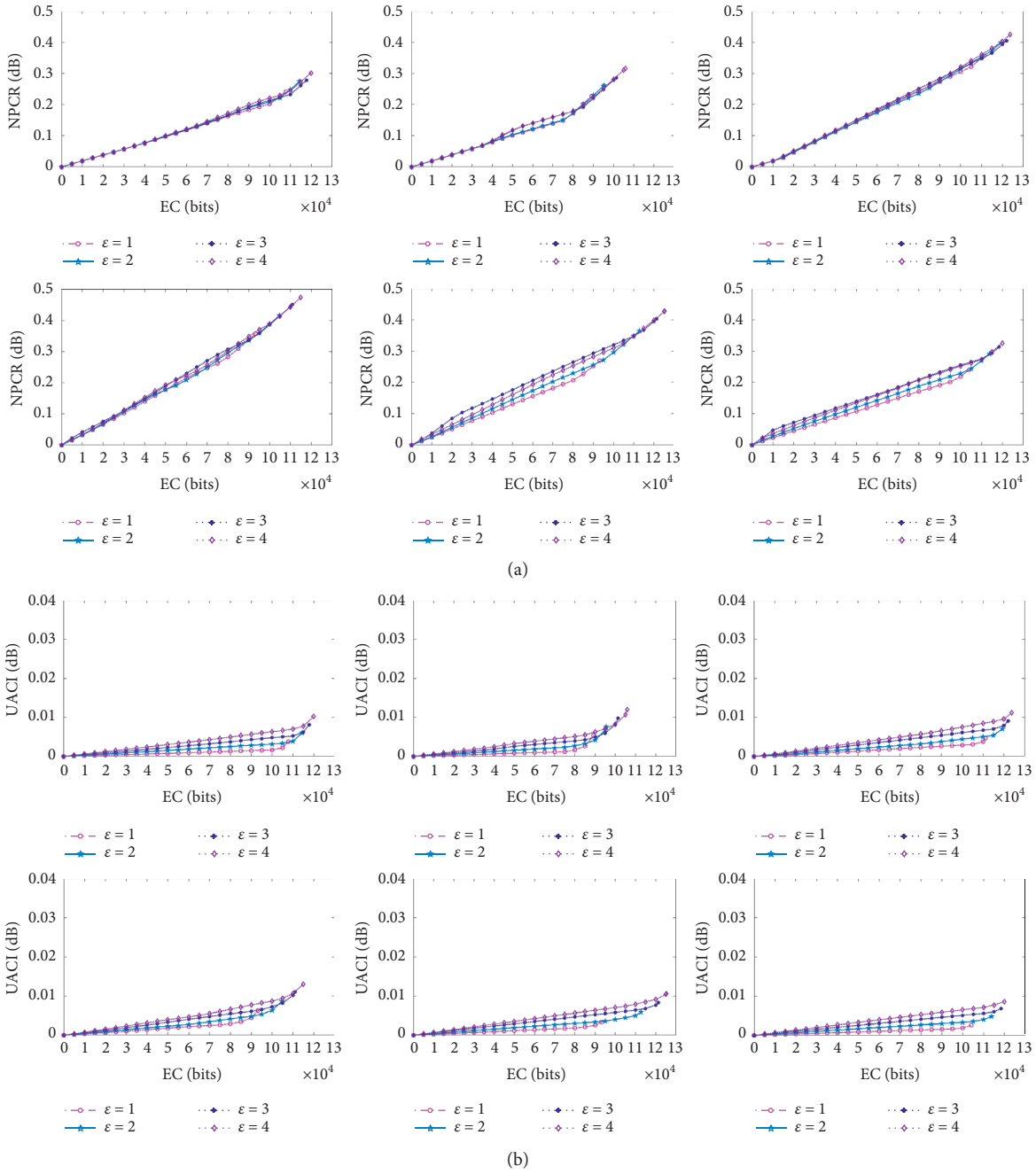
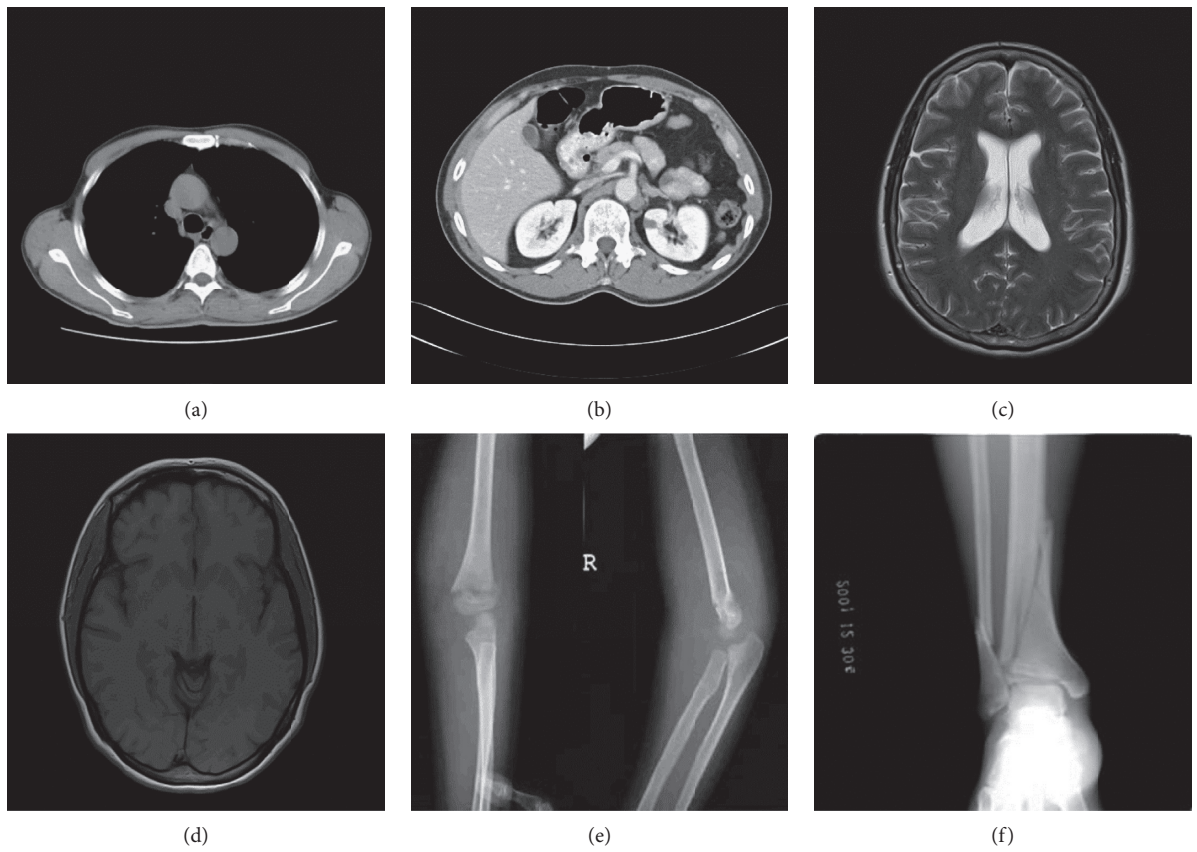


Figure 7: Curves of NPRC and UACI for the stego-images. (a) NPRC. (b) UACI.

the mean value of pure EC was around 29110 bits, which was much lower than that of other cases, in which the types of block size were 4×4, 4×1, and 1×4. This occurred because the elements within a sequence derived from an image block are not so quite similar to each other when the block size is larger, and this leads to a serious decrement of the number of the regular blocks. In addition, it is interesting to note that, for a similar block size, the pure EC obtained for the 4×1 block was about 13000 bits higher than the pure EC obtained for the 1×4 block. That is to say, the pure EC and PSNR will be affected by the different sizes of the blocks. For example, in our experiments, the means of pure EC and PSNR reached the highest values, i.e., 75479 bits and 36.4917 dB, respectively, when the size of the block was set to 2×2.

4.3. Performance Comparison. In this section, we compare the results provided by the proposed RDH method, Parah et al.'s method [28], Geetha et al.'s method [29], and Gao et al.'s method [30] to demonstrate the excellent performance of our approach.

First, a comparison in terms of PSNR between methods [28–30] and the proposed RDH method was conducted on the six medical images that are listed in Table 8. For the sake



(a) (b) (c)

(d) (e) (f)

Figure 8: Stego-medical-images: (a) Kidney_A; (b) Kidney_B; (c) Brain_A; (d) Brain_B; (e) Sketeton_A; (f) Sketeton_B.

Table 6: Size of the |LM| and ratio of the regular block for six medical images

	|LM| without compression	|LM| with compression	Ratio of regular block (%)
Kidney_A	65536 bits	20898 bits	82.2159
Kidney_B	65536 bits	38756 bits	67.9978
Brain_A	65536 bits	16429 bits	84.8206
Brain_B	65536 bits	33596 bits	71.0617
Sketeton_A	65536 bits	22077 bits	71.0022
Sketeton_B	65536 bits	13688 bits	79.3823
Average	**65536** bits	**24240.67**	**76.0801**

of fairness, Table 8 gives the PSNR provided by the above four methods under the EC of 50,000 bits. It can be observed that the proposed RDH method obtains a higher PSNR than that of Parah et al.'s method [28], Geetha et al.'s method [29], and Gao et al.'s method [30], and the differences were about 18.4055 dB, 12.1389 dB, and 10.4440 dB, respectively. The main reasons are that, in the front two methods [28, 29], the use of the pixel to block technique seriously degraded the quality of the images, and, in the last method [30], the large scale in shifting pixels in order to enhance the contrast in the image leads to a decrement in the quality of the image. Also, it can be seen that the algorithm designed by Gao et al. [30] was not effective for the four images, i.e., "Kindey_A," "Kindey_B," "Sketeton_A," and "Sketeton_B," so they are marked as "NA".

For this, we used the proposed RDH methods and methods [28, 29] on the "Brain01" and "Xray" images used in Gao et al. [30] to further evaluate their performances, and the results are presented in Figure 11. For the EC aspect, the proposed RDH method was slightly inferior compared to the other methods [28–30]. However, it was apparent that the proposed RDH scheme also achieved a considerable PSNR value, which was better than those of the other three methods [28–30] in most cases.

Secondly, comparisons of the various features of the different RDH methods are given in Table 9. In methods of [28, 29], the pixel to block conversion technique and the rhombus mean interpolation technique were used as effective means to interpolate an original medical image to a cover medical image, respectively. Then, the secret data was inserted into the LSBs of the pixels in this cover medical image to ensure its reversibility. Both of these methods can extract secret data without errors. In addition, they can reconstruct the cover medical image (but not the original

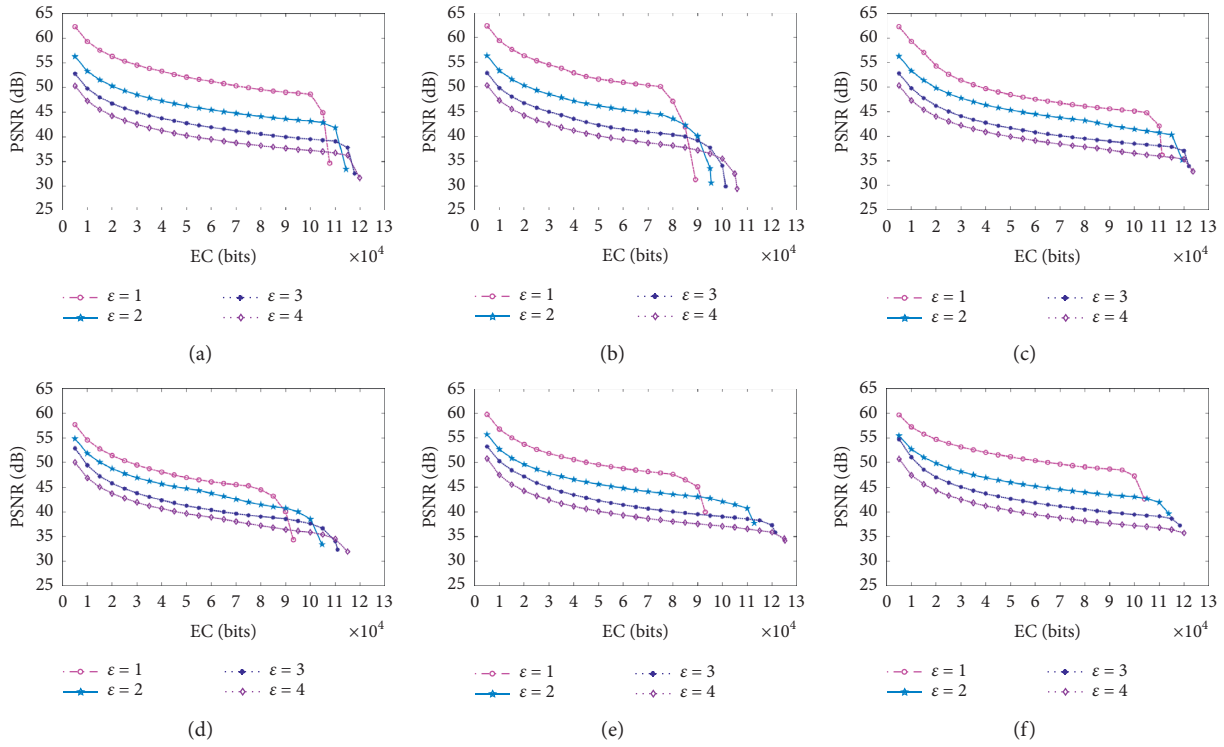


Figure 9: Graphs in terms of PSNR and EC for six medical images when ε varies from 1 to 4. (a) Kidney_A; (b) Kidney_B; (c) Brain_A; (d) Brain_B; (e) Sketeton_A; (f) Sketeton_B.

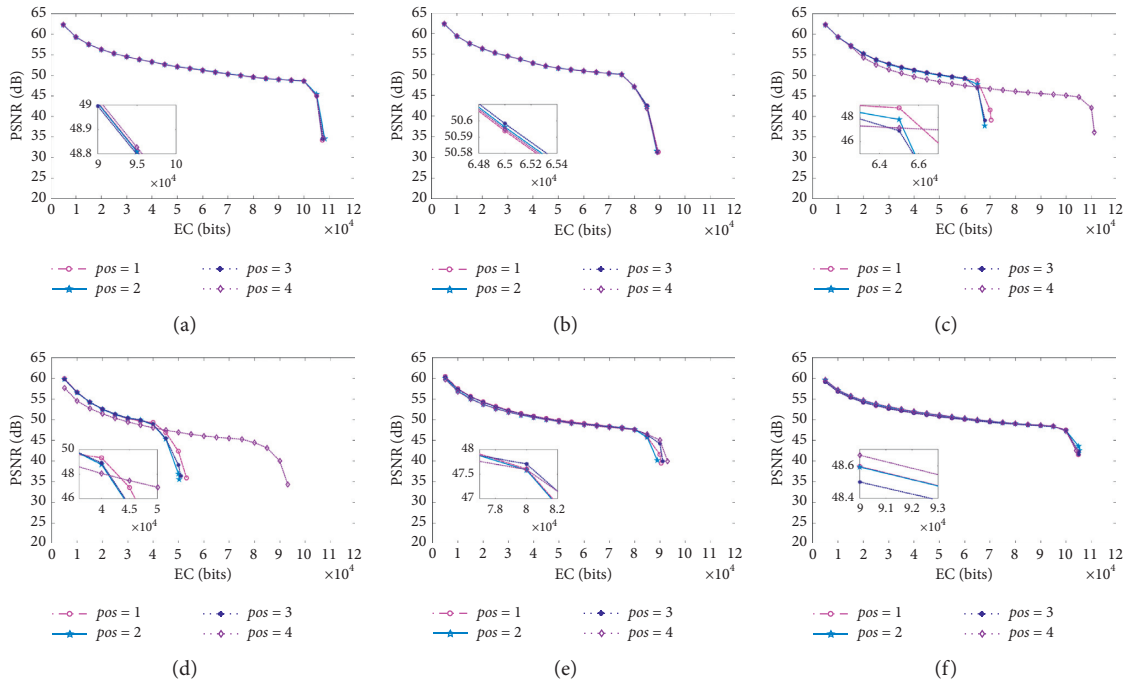


Figure 10: Graphs in terms of PSNR and EC for six medical images where pos varies from 1 to 4. (a) Kidney_A; (b) Kidney_B; (c) Brain_A; (d) Brain_B; (e) Sketeton_A; (f) Sketeton_B.

medical image) in a lossless way, which was quite different from method [30] and our RDH method. Not only that, the visual quality of stego-medical-image provided by methods [28, 29] was relatively lower than that of the proposed method. Concerning the method in [30], the traditional histogram shifting technique was used to hide secret data into pixels that have peak-bin. An automatic histogram stretching technique was designed to vacate a more

Table 7: Pure ECs (bits) and PSNRs (dB) obtained on six medical images with different block sizes.

Images	Block size (unit: pixels)							
	2 × 2		4 × 1		1 × 4		4 × 4	
	Pure EC	PSNR	Pure EC	PSNR	Pure EC	PSNR	Pure EC	PSNR
Kidney_A	86864	34.6457	86755	28.5421	79886	31.7735	42777	31.3795
Kidney_B	50370	31.2902	50756	26.3472	37386	28.2342	27379	30.1879
Brain_A	94747	36.1669	71924	30.9600	62540	32.7430	7151	33.2414
Brain_B	59546	34.3169	41513	30.7882	34905	29.9029	NA	NA
Sketeton_A	70987	39.9571	79830	39.73034	55944	32.9566	25352	36.7719
Sketeton_B	90360	42.5736	92830	41.0707	81843	37.1164	42889	38.2895
Average	**75479**	**36.4917**	**70601**	**32.9064**	**58751**	**32.1211**	**29110**	**33.9740**

NA: Not applicable.

Table 8: Comparisons of PSNRs for different RDH methods under the EC of 50,000 bits.

Images	PSNRs (dB)			
	Parah et al. [28]	Geetha et al. [29]	Gao et al. [30]	Proposed method
Kidney_A	28.6677	34.4016	NA	52.0954
Kidney_B	24.4430	28.5094	NA	51.5742
Brain_A	32.6299	39.3391	39.8084	48.4490
Brain_B	30.1521	36.1370	38.2170	46.9342
Sketeton_A	35.3404	42.6602	NA	49.5485
Sketeton_B	38.0766	45.8597	NA	51.1387
Average	**31.5512**	**37.8178**	**39.5127**	**49.9567**

NA: not applicable.

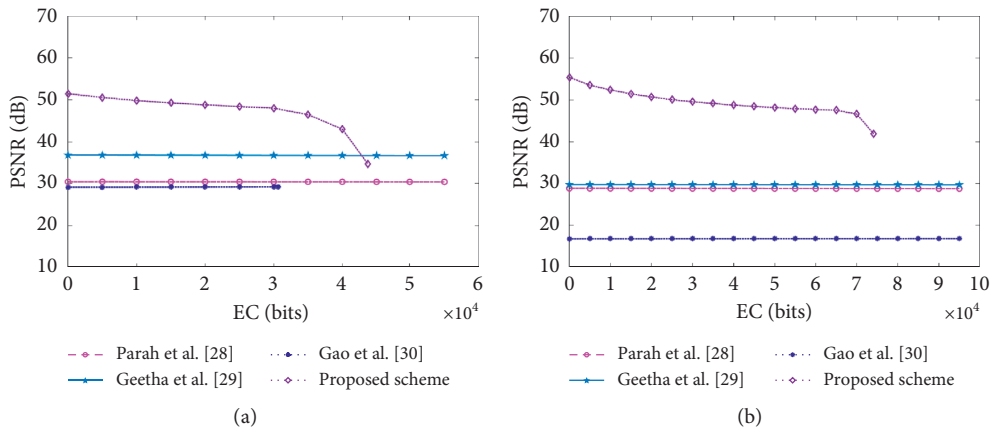


Figure 11: Graphs showing the PSNR and EC values for two medical images. (a) "Brain01"; (b) "Xray".

Table 9: Comparisons of features for different RDH methods.

Features	Parah et al. [28]	Geetha et al. [29]	Gao et al. [30]	Proposed method
Methodology	Pixel to block technique	Rhombus mean interpolation technique	Histogram shifting	Pixels rotation
Hidden component	LSBs	LSBs	Pixels	Pixels
Bit errors	No	No	No	No
PSNR (on average)	−(30.4335 dB)	(36.7489 dB)	−(29.1779 dB)	(47.0307 dB)
Reversibility	No	No	Yes	Yes

More "+" means better image quality when compared to the Refs. [28, 30]. The symbol "−" represents the baseline in terms of PSNR.

embeddable room. In other words, the method in [30] obtains more EC at the expense of an apparent decrease in the visual quality of stego-medical-images. It is important to note that this method has the feature of reversibility if the replacement of the LSBs is not conducted for its NROI. The proposed RDH method employs the idea of pixel rotation to insert the secret data into the medical image with reversibility. More specifically, the payload bits are carried by a specific sequence generated from a regular image block, instead of changes in the pixel values. Therefore, the

proposed RDH method has a considerable PSNR of stego-medical-image.

5. Conclusions

In this paper, a novel rotation based RDH method for medical images is presented.

According to the characteristic of pixel distribution of the medical image, the proposed RDH method separates all image blocks into the regular blocks, singular blocks, unusable blocks, and overflow blocks. The payload is only inserted into the regular blocks to reduce the invalid rotation on pixels. Additionally, different from the conventional RDH method, we define a state mapping model to construct a mapping between various states of an image block and the payload in a one-one manner. Then, the selected state is put back into the original medical image block to form the stego-medical-image, and the payload is carried. We implemented the proposed RDH method and evaluated it with extensive experiments. It was demonstrated that our rotation based RDH method can achieve excellent performance, exceeding the performance of some recent works in both EC and PSNR. In the future, we plan to investigate the improvement of our algorithm by further considering the ROI and NORI to optimize the visual quality of stego-medical-images.

Data Availability

All the data can be accessed in the public database.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper to the best of their knowledge.

Acknowledgments

This work was supported in part by the Open Fund of Engineering Research Center for ICH Digitalization and Multi-Source Information Fusion of Fujian Province under Grant No. FJ-ICH201901 and in part by the Education-Scientific Research Project for Middle-Aged and Young of Fujian Province under Grant Nos. JAT190488 and JAT201368. This work was also supported by Ministry of Science and Technology, Taiwan, under Grant No. MOST109-2410-H-167-014.

References

[1] C.-C. Chang, C.-T. Li, and Y.-Q. Shi, "Privacy-aware reversible watermarking in cloud computing environments," *IEEE Access*, vol. 6, pp. 70720–70733, 2018.

[2] C.-C. Chang, C.-T. Li, and K. Chen, "Privacy-preserving reversible information hiding based on arithmetic of quadratic residues," *IEEE Access*, vol. 7, pp. 54117–54132, 2019.

[3] C. C. Chang, Y. Liu, and T. S. Nguyen, "A novel turtle shell based scheme for data hiding," in *Proceedings of the 10th International Conference on Intelligent Information Hiding & Multimedia Signal Processing*, pp. 89–93, Kitakyushu, Japan, August 2014.

[4] D.-C. Lou, M.-C. Hu, and J.-L. Liu, "Multiple layer data hiding scheme for medical images," *Computer Standards & Interfaces*, vol. 31, no. 2, pp. 329–335, 2009.

[5] G. D. Su, C. C. Chang, C. C. Lin, and Z. Q. Yao, "Secure high capacity tetris-based scheme for data hiding," *IET Image Processing*, vol. 14, no. 17, pp. 4633–4645, 2021.

[6] J. Jun Tian, "Reversible data embedding using a difference expansion," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 8, pp. 890–896, 2003.

[7] A. M. Alattar, "Reversible watermark using the difference expansion of a generalized integer transform," *IEEE Transactions on Image Processing*, vol. 13, no. 8, pp. 1147–1156, 2004.

[8] Y. Hu, H.-K. Lee, K. Chen, and J. Li, "Difference expansion based reversible data hiding using two embedding directions," *IEEE Transactions on Multimedia*, vol. 10, no. 8, pp. 1500–1512, 2008.

[9] H.-C. Wu, C.-C. Lee, C.-S. Tsai, Y.-P. Chu, and H.-R. Chen, "A high capacity reversible data hiding scheme with edge prediction and difference expansion," *Journal of Systems and Software*, vol. 82, no. 12, pp. 1966–1973, 2009.

[10] W. Wang, J. Ye, T. Wang, and W. Wang, "Reversible data hiding scheme based on significant-bit-difference expansion," *IET Image Processing*, vol. 11, no. 11, pp. 1002–1014, 2017.

[11] D. M. Thodi and J. J. Rodriguez, "Prediction-error based reversible watermarking," vol. 3, pp. 1549–1552, in *Proceedings of the 2004 International Conference on Image Processing*, vol. 3, , IEEE, Singapore, October 2004.

[12] H.-W. Tseng and C.-P. Hsieh, "Prediction-based reversible data hiding," *Information Sciences*, vol. 179, no. 14, pp. 2460–2469, 2009.

[13] B. Ou, X. Li, Y. Zhao, R. Ni, and Y.-Q. Shi, "Pairwise prediction-error expansion for efficient reversible data hiding," *IEEE Transactions on Image Processing*, vol. 22, no. 12, pp. 5010–5021, 2013.

[14] X. Li, J. Li, B. Li, and B. Yang, "High-fidelity reversible data hiding scheme based on pixel-value-ordering and prediction-error expansion," *Signal Processing*, vol. 93, no. 1, pp. 198–205, 2013.

[15] C.-C. Chang, "Adversarial learning for invertible steganography," *IEEE Access*, vol. 8, pp. 198425–198435, 2020.

[16] Z. Zhicheng Ni, Y. Q. Yun-Qing Shi, N. Ansari, and W. Wei Su, "Reversible data hiding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 16, no. 3, pp. 354–362, 2006.

[17] S. K. Lee, Y. H. Suh, and Y. S. Ho, "Reversible image authentication based on watermarking," in *Proceedings of the 2006 IEEE International Conference on Multimedia and Expo*, pp. 1321–1324, IEEE, Toronto, Canada, July 2006.

[18] W. Hong, T.-S. Chen, and C.-W. Shiu, "Reversible data hiding for high quality images using modification of prediction errors," *Journal of Systems and Software*, vol. 82, no. 11, pp. 1833–1842, 2009.

[19] Y. Jia, Z. Yin, X. Zhang, and Y. Luo, "Reversible data hiding based on reducing invalid shifting of pixels in histogram shifting," *Signal Processing*, vol. 163, pp. 238–246, 2019.

[20] F. Peng, Y. Zhao, X. Zhang, M. Long, and W.-Q. Pan, "Reversible data hiding based on RSBEMD coding and adaptive multi-segment left and right histogram shifting," *Signal Processing: Image Communication*, vol. 81, Article ID 115715, 2020.

[21] W. He, J. Cai, K. Zhou, and G. Xiong, "Efficient PVO-based reversible data hiding using multistage blocking and

prediction accuracy matrix," *Journal of Visual Communication and Image Representation*, vol. 46, pp. 58–69, 2017.

[22] B. Jana, D. Giri, and S. K. Mondal, "Partial reversible data hiding scheme using (7, 4) hamming code," *Multimedia Tools and Applications*, vol. 76, no. 20, pp. 21691–21706, 2017.

[23] G.-D. Su, Y. Liu, and C.-C. Chang, "A square lattice oriented reversible information hiding scheme with reversibility and adaptivity for dual images," *Journal of Visual Communication and Image Representation*, vol. 64, Article ID 102618, 2019.

[24] C.-C. Chang, C. T. Li, and C.-T. Li, "Algebraic secret sharing using privacy homomorphisms for IoT-basedhealthcare systems," *Mathematical Biosciences and Engineering*, vol. 16, no. 5, pp. 3367–3381, 2019.

[25] C. C. Chang, "Neural reversible steganography with long short-term memory," *Security and Communication Networks*, vol. 2021, Article ID 5580272, 14 pages, 2021.

[26] L.-C. Huang, L.-Y. Tseng, and M.-S. Hwang, "A reversible data hiding method by histogram shifting in high quality medical images," *Journal of Systems and Software*, vol. 86, no. 3, pp. 716–727, 2013.

[27] V. Kelkar, K. Tuckley, and H. Nemade, "Novel variants of a histogram shift-based reversible watermarking technique for medical images to improve hiding capacity," *Journal of Healthcare Engineering*, vol. 2017, Article ID 3538979, 7 pages, 2017.

[28] S. A. Parah, F. Ahad, J. A. Sheikh, and G. M. Bhat, "Hiding clinical information in medical images: a new high capacity and reversible data hiding technique," *Journal of Biomedical Informatics*, vol. 66, pp. 214–230, 2017.

[29] R. Geetha and S. Geetha, "Embedding electronic patient information in clinical images: an improved and efficient reversible data hiding technique," *Multimedia Tools and Applications*, vol. 79, no. 19, pp. 12869–12890, 2020.

[30] G. Gao, S. Tong, Z. Xia, B. Wu, L. Xu, and Z. Zhao, "Reversible data hiding with automatic contrast enhancement for medical images," *Signal Processing*, vol. 178, Article ID 107817, 2021.

[31] C. C. Chang, "Cryptospace invertible steganography with conditional generative adversarial networks," *Security and Communication Networks*, vol. 2021, Article ID 5538720, 14 pages, 2021.

[32] Y. Liu, C. C. Chang, and T. S. Nguyen, "High capacity turtle shell-based data hiding," *IET Image Processing*, vol. 10, no. 2, pp. 130–137, 2016.

[33] F. Li, H. Zhu, J. Yu, and C. Qin, "Double linear regression prediction based reversible data hiding in encrypted images," *Multimedia Tools and Applications*, vol. 80, no. 2, pp. 2141–2159, 2021.

Hindawi
Journal of Healthcare Engineering
Volume 2021, Article ID 9958256, 14 pages
https://doi.org/10.1155/2021/9958256

Research Article

An IoT-Based Motion Tracking System for Next-Generation Foot-Related Sports Training and Talent Selection

Shanshan Lu,[1] Xiao Zhang,[1] Jiangqing Wang,[1] Yufan Wang,[2] Mengjiao Fan,[2] and Yu Zhou[3]

[1]*Department of Computer Science, South-Central University for Nationalities, Wuhan 430074, China*
[2]*Department of Industrial Engineering & Management, School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China*
[3]*College of Computer Science and Software Engineering, Shenzhen University, Shenzhen, China*

Correspondence should be addressed to Xiao Zhang; xiao.zhang@my.cityu.edu.hk and Yu Zhou; yu.zhou@szu.edu.cn

Received 4 March 2021; Revised 7 May 2021; Accepted 12 June 2021; Published 25 June 2021

Academic Editor: Chi-Hua Chen

Motion tracking in different fields (medical, military, film, etc.) based on microelectromechanical systems (MEMS) sensing technology has been attracted by world's leading researchers and engineers in recent years; however, there is still a lack of research covering the sports field. In this study, we propose a new AIoT (AI + IoT) paradigm for next-generation foot-driven sports (soccer, football, takraw, etc.) training and talent selection. The system built is cost-effective and easy-to-use and requires much fewer computational resources than traditional video-based analysis on monitoring motions of players during training. The system built includes a customized wireless wearable sensing device (WWSDs), a mobile application, and a data processing interface-based cloud with an ankle attitude angle analysis model. Eleven right-foot male participators wore the WWSD on their ankle while each performed 20 instances of different actions in a formal soccer field. The experimental outcome demonstrates the proposed motion tracking system based on AIoT and MEMS sensing technologies capable of recognizing different motions and assessing the players' skills. The talent selection function can partition the elite and amateur players at an accuracy of 93%. This intelligent system can be an emerging technology based on wearable sensors and attain the experience-driven to data-driven transition in the field of sports training and talent selection and can be easily extended to analyze other foot-related sports motions (e.g., taekwondo, tumble, and gymnastics) and skill levels.

1. Introduction

Recent trends in smart wearable technologies based on the Internet of Things (IoT) have opened up a large number of applications [1], which involve the recognition of sports activities [2–4]. Soccer, also known as football in some countries, is one of the most popular sports in the world with numerous professionals and an even larger number of nonprofessional practitioners. Soccer players train a specific set of well-defined motions (e.g., shooting and passing) to consolidate them into muscle memory and lower their reaction time during a game. The correct execution of the exercises and training has a steep learning curve. Quantity and quality information about soccer motions like shooting and passing an exhaustive performance evaluation is indispensable for coaches and players in training sessions and competitions.

Traditionally, the training of professional soccer players is human-oriented, which is done by subjective guidance based on the personal experience of coaching staff or trainers. However, different coaches could have different ideas upon their experience, and there are many fast and slight soccer motions during an action that cannot be captured through their visual sense. The need for an objectively and digitized method rather than a subjective method is raised. In the last few years, there have appeared lots of research efforts in the field of professional soccer for a measurable and quantifiable

analysis of sports through information technology. The manager and coaching staff use videography to monitor the biomechanics of soccer actions and check the objective players' performance [5]. Despite this, videography is typically restricted by equipment and environment and unable to give feedback to the coaching staff or soccer players in real time due to the significant amounts of storage and computational load.

The rapid development of bluetooth low energy (BTLE) and microelectromechanical systems (MEMS) technologies implements durative motion data capture and communication in real time. There is increasing research on motion recognition by using inertial low-cost sensors worn on the body. For example, the authors of [3] use inertial sensors worn on the wrist to identify the stroke type of table tennis. In [6, 7], wearable sensing devices (WSDs) on the wrist are used to classify different volleyball and badminton actions. However, due to the nature of soccer, the ankle movements of players are complex and indicative of different soccer motions. Moreover, the high variability in the training and individual execution of each exercise makes the classification of soccer motions extremely challenging. A particular challenge we address is to classify the most fundamental soccer motions (i.e., shooting and passing) (the development process and methods we describe in this paper can be reused in more motion recognition applications). The existing approaches used in other sports are hardly applicable [8].

Therefore, an IoT system based on WSD is proposed in this work to provide objective feedback to coaches and soccer players after or during a training session to help them improve their skills. The proposed IoT system consists of wearable devices, a mobile device (e.g., a mobile phone or tablet), and a data processing platform based cloud. A WSD of MEMS is used to collect raw data from soccer players. By using BTLE technology, the material is transmitted to the cloud-based data processing platform, which analyzes the data and outputs the results to a mobile device in real time. A support vector machine (SVM) model classification algorithm with an ankle-based attitude angle model is proposed to recognize different motions and assess different skill levels.

This system is ideally suitable for the young player soccer training. A soccer training session typically includes exercises of passes, crosses, and shots For example, in a soccer club or training school, there are dozens or hundreds of young players conducting an exercise of shooting or passing. It is evident that a limited number of coaches are hard to assess all soccer players' individual execution of each exercise. Instead, by using our developed system, how many passes/shoots have been done and how professional will be shown on a coach's mobile device. The work we present enables soccer players to keep track of the training exercises they perform. The managers and coaches could get an overview of the exercises during past days, weeks, or months, which they could use to arrange future training sessions pointedly. We intend to apply the system in a realistic training environment with a focus on the recognition of basic soccer motions. This study demonstrates the trend in using the IoT framework for a new era of soccer training.

Our key novelty and main contributions are summarized as follows:

(i) We develop an IoT system for soccer motion recognition and assessment, in which a WSD with overall dimensions of 2 mm × 10.3 mm × 8.7 mm is developed to capture inertial data; a mobile application allows visualization of experimental outputs and transfers the data to the cloud-based processing platform.

(ii) We build a SVM classification algorithm with an attitude angle model to classify different soccer motions and make a distinction between the skill levels between elite and amateur players.

(iii) Our system is confirmed by the experimental results that it enables recognizing different motions, that is, shooting and passing, with an accuracy close to 90%. The talent selection distinguishes between the elite and amateur players with an accuracy of 93%.

We organize the remainder main structure of this article as follows: we give a literature review in Section 2. Then, we illustrate the system composition and data acquisition in Section 3. Next, Section 4 presents our algorithm for soccer motion recognition and assessment as well as the experimental results. Finally, Section 5 summarizes the contribution.

2. Related Work

Motion analysis and recognition by using information technology have been a field of study for decades. During this period, a variety of motion recognition applications and approaches have been proposed to assess the performance and correctness of a physical exercise of an athlete and provide feedback to users about their actions.

In early studies, motion recognition was a kind of computer vision field that recognizes human pose or action, in which video image acquisition and processing technology was adopted to identify the human body's motion behavior. In vision-based action recognition, the training method is mainly to collect videos, images, and other information to analyze the moving process. Therefore, it is necessary to place cameras and other monitors in advance in the detection environment for data collection. The popularity of the RGB camera has made it an effective auxiliary tool for the human motion recognition study and also leads to the appearance of several survey articles [9, 10] discussing various features and classifiers for human action recognition. Though image analysis technology can identify people's daily moves more accurately, image analysis technology still suffers from many shortcomings. In nature, a large number of hardware conditions are required so as to run computer vision algorithms and computationally concentrated image processing. Moreover, a large amount of image data acquired suffers from insufficient storage, leading to failure in monitoring [11].

Several human activities' recognition approaches have been presented, in which wearable sensors with accelerometers and gyroscopes are used to acquire motion data [12, 13]. A much wider field of views can be obtained by this sensor technology. Long-running recordings, calculation, and constant interaction are possible owing to the progress of inertial sensors in energy-consuming reduction and the improvement of computational power. Moreover, 3D motion data which consists of 3-axis angular velocities from gyroscopes as well as 3-axis accelerations from accelerometers can be given by wearable inertial sensors [14, 15].

Sensors are often of small size, high precision and sensitivity, and low environmental requirements and power consumption, as well as easy to wear. A universal activity recognition depth framework based on multimodal wearable sensors was proposed in [16, 17]. Shoaib et al. [18] used inertial sensors to detect the activities that involve hand gestures (smoking, eating, drinking, giving a chat, etc.). In the field of competitive sports, most movement recognition systems are dedicated to solving wrist movements or simple lower limb movements such as walking and standing [19].

There are a few studies on the recognition and assessment of soccer players' motions. Reference [20] indicated that soccer players' shin guards instrumented with sensors can measure ankle joint kinematics. Meamarbashi et al. tested whether sensors can be used to examine kinematic parameters of soccer players in [21], but no further analysis of the data. In [4], Haladjian et al. developed a smart glove with an inertial sensor and gave a machine learning algorithm to recognize the goalkeeper's training exercises, which was a good attempt to use WSD to recognize soccer motion. As far as we know, this is the first work to use a single WSD attached to an ankle to recognize different soccer players' motions (not just the goalkeepers) and assess their skill levels.

3. System Design and Data Acquisition

Though WSDs have been widely used in motion recognition, for soccer motions, the existing approaches are hardly applicable. However, the limited number of coaches are hard to assess all soccer players' individual performance. Thus, this study presents a complete system in using the IoT framework for soccer motion recognition. By using the developed system, the passes/shoots and their skill levels can be shown on a mobile device hold by a coach. The system enables soccer players to keep track of the training exercises they perform. The specific details of hardware design, interface software, and data acquisition are given in this section. Figure 1 shows an overview of our IoT system for soccer motion recognition and assessment, which is composed of wireless WSDs, mobile devices, and cloud-based data processing platform.

By using the proposed IoT system, BTLE sent mobile device data collected by WSD. It transfers the raw data to the remote execution server via cloud computing technology at the moment when the mobile device received the exercise data. Completing all data collection and processing, users can view the analysis results of soccer players through the platform.

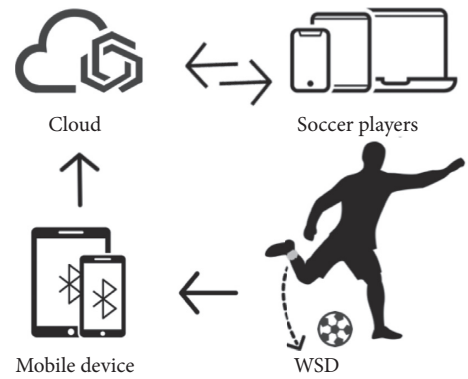


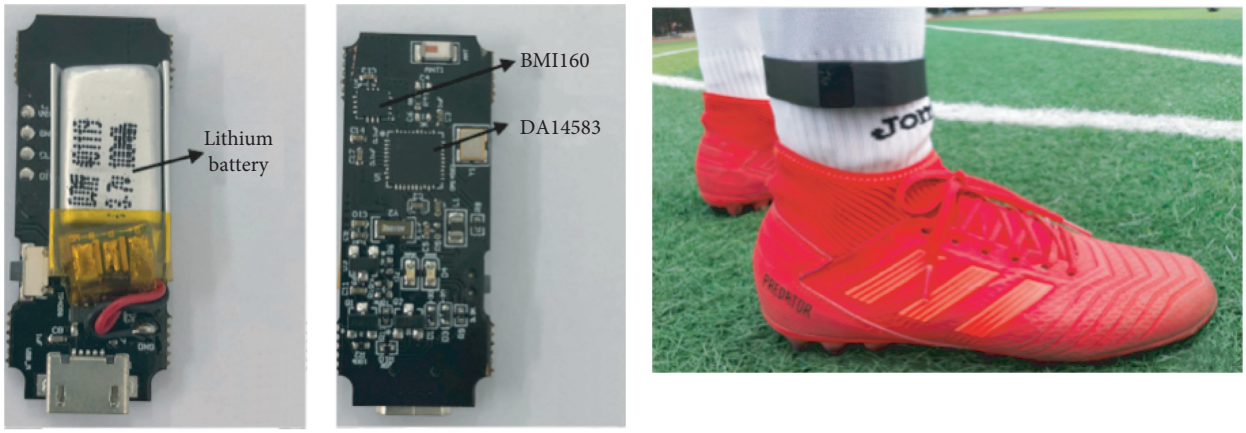
Figure 1: In the IoT system, the motion sensor is attached to the right ankle of the soccer player. The motion data is transmitted via Bluetooth to the Cloud platform for further process.

3.1. Hardware and Software System. We collect data using a wireless WSD, jointly developed by our team and the company *AI Motion Sports* (https://www.aimotionsports.com/en/). Figure 2 shows the components of the device. This wireless WSD is comprised of four major components: an MEMS motion sensing chip with a 3-axis gyroscope and accelerometer, a microprocessing unit with Bluetooth wireless newsletter function, a lithium battery, and an ON-OFF switch on the back of the board.

The 3-axis gyroscope and accelerometer in MEMS are used to sense the motions and transform the signal into raw data. DA14583 Programmed in Dialog Semiconductor, Reading, UK, baseband radio processor and entirely integrates radio transceiver for BTLE. Not only is the microprocessor with chip nRF52832 characterized by low energy consumption but also highly efficient transmission by using Bluetooth v5.0. BMI160 from BOSCH has a fitting sensor range and tiny size (2.5 mm × 3.0 mm × 0.83 mm) [22] to enable it to be adopted.

The acceleration and orientation in three dimensions can be obtained through the high-integration and low-power BMI160. The entire WSD keep in size within 10.3 mm × 8.7 mm × 2 mm, using this IMU chip, weighing only 2 g. We develop a mobile application for receiving and visualizing data transmitted from WSD in real time. The application developed by the Java Script consists of three functional modules: wireless connection, real-time captured data demonstration, and data synchronization to the cloud. All data sensed are display synchronously and saved locally. A cloud-based approach is also used to save data received to a remote server. Once the data collection process is completed, user clients can access the analytical results.

3.2. Data Collection. We conduct the experiments at the Chinese Football Association (CFA) Youth Academy (Wuhan) (Chinese National Football Team Training Center) and Sports Center of South-Central University for



(a) (b) (c)

Figure 2: (a) A lithium battery on the back of the board. (b) Circuit board of the WSD. (c) The WSD is attached to the right ankle of a soccer player.

Nationalities. We recruited 11 male soccer players, including 5 elites and 6 amateurs who are right-footed players. Among them, elite players had represented their club as participants in more than ten 10 national games from CFA Youth Academy (Wuhan); amateurs were beginners from university. As soccer is a foot-based sport, our small and light WSD is attached to each right-footed subject' right ankle when performing soccer basic training to make sure that major inertial data of soccer motions can be captured without obstruction. During the experiment, each subject performed 20 passings with the inside of the foot and 20 shootings with foot-arch at the soccer field of CFA Youth Academy (Wuhan). The subject performed the actions at the same position. Each subject's passings and shootings are required to be performed within a certain range of speed and precision; otherwise, we considered it was an invalid action and we did not count it. Figure 3 illustrates our experimental scene.

Figure 4 displays the six-axis synchronal raw data captured by the WSD from elites by performing different soccer motions. The angular velocity as well as acceleration for passing are displayed in Figures 4(a) and 4(b), while Figures 4(c) and 4(d) show the related data information for shooting.

4. Soccer Motion Recognition and Assessment

The process of the soccer motion recognition system, as shown in Figure 5, can be divided into five steps: data preprocessing, attitude angle modeling, feature extraction and selection, dimensionality reduction, and classification. To be specific, the ankle-based attitude angles' features act a pivotal part in improving the accuracy of classification, since soccer as a type of foot-based sport is more complicated than other sports like racket sports. Principle component analysis (PCA) is adopted in our system to balance the complexity and the accuracy. Finally, we propose a SVM-based classification algorithm to recognize the soccer motions.

In data preprocessing, a 3-point moving average filter is used to lessen the noisy interference from raw data. By locating the peak of the signal, the segmentation of the motion signal can be processed automatically. The sliding-window algorithm [23] can be applied to process the signals in real time.

4.1. Angle Trajectory Model. As shown in Figure 6, the intuitive signal of $\{\omega_x, \omega_y, \omega_z\}$ (acceleration from three dimensions: x -axis, y -axis, and z -axis) and $\{\alpha_x, \alpha_y, \alpha_z\}$ (angular velocity from three dimensions) are not significant. We present an attitude angle model for extracting more useful features.

For sports motion recognition, angular velocity and acceleration are important features. Besides, the ankles' rotations are specially important for soccer motion recognition and assessment [16]. A three-dimensional rotation problem is typically addressed by a rotation matrix. Mining the information of attitude angle can improve characterizing the action. The attitude angle is expressed by three Euler angles: yaw, pitch, and roll. Figure 7 shows the specific transformation.

The quaternion [24] as the quotient of two directed lines in a three-dimensional space is used to solve angle in the angle trajectory model. The representation of the quaternion is $p = p_\omega + p_x i + p_y j + p_z k$ with a real part p_ω and three imaginary parts parameters p_x, p_y, and p_z. Calculate the initial attitude angle:

$$\begin{cases} \theta = \arcsin\left(a_{x0}\right), \\ \varphi = \arctan\dfrac{magy_0}{magx_0}, \\ \psi = \arctan\left(-\dfrac{a_{y0}}{a_{z0}}\right), \end{cases} \tag{1}$$

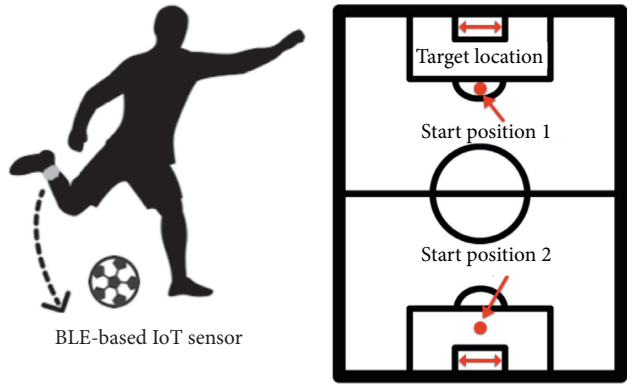


Figure 3: The sensor placement and experimental site setting for raw data collection. (a) The WSD is attached to a subject's or player's right ankle. (b) Positions of soccer players to perform passing and shooting.

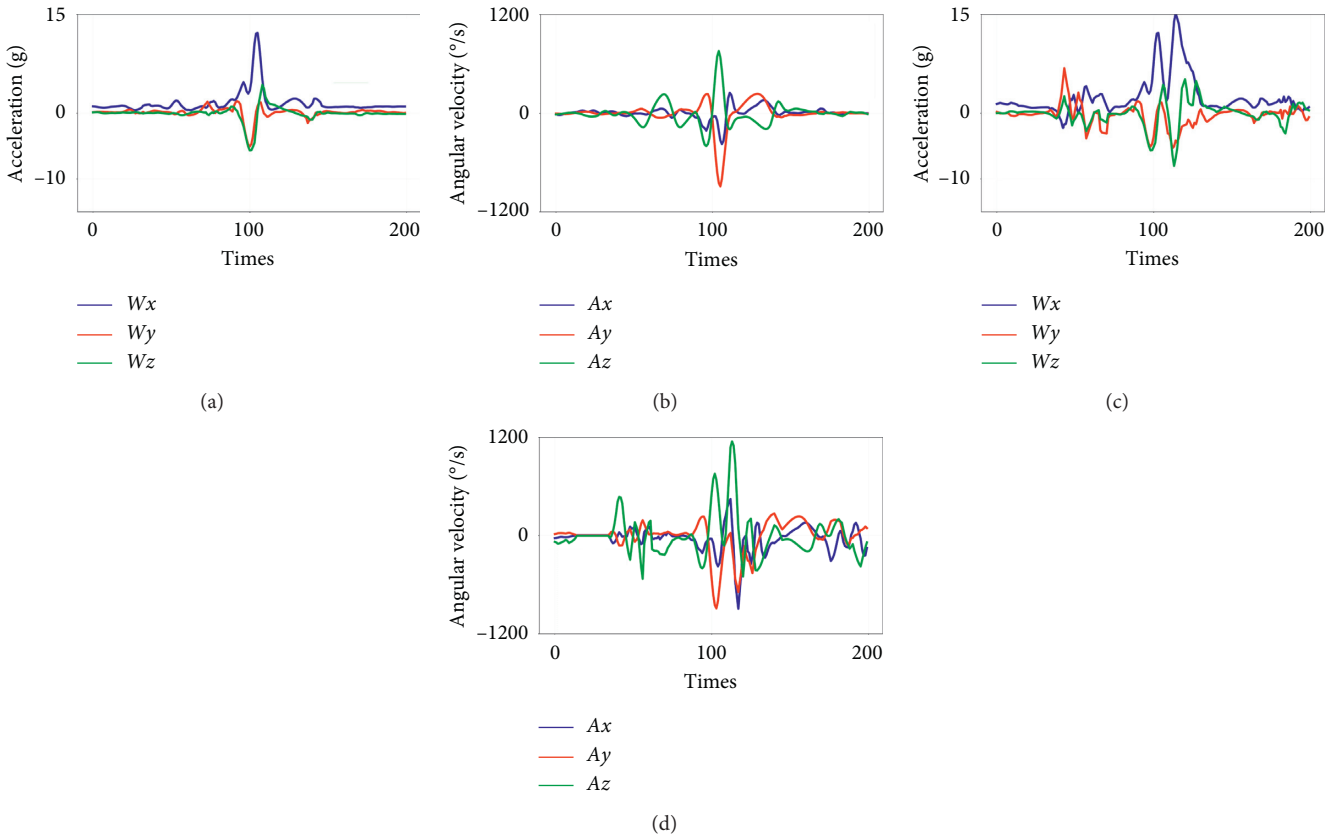


Figure 4: Passing and shooting signal collected from an elite subject. (a) Acceleration of a passing. (b) Angular velocity of a passing. (c) Acceleration of a shooting. (d) Angular velocity of a shooting.

where a_{x0}, a_{y0}, a_{z0} represent the initial acceleration; θ, ψ, and φ represent the initial yaw, roll, and pitch, respectively. $\text{mag}x_0$ and $\text{mag}y_0$ are calculated by a_{x0}, a_{y0}, a_{z0}, specifically, $\text{mag}x_0 = a_{x0}a_{y0} + a_{x0}a_{z0}/1 - 2(a_{y0}^2 + a_{z0}^2)$, and $\text{mag}y_0 = a_{x0}a_{y0} + a_{y0}a_{z0}/1 - 2(a_{x0}^2 + a_{z0}^2)$ [25]. The initial four elements are calculated from the initial angle:

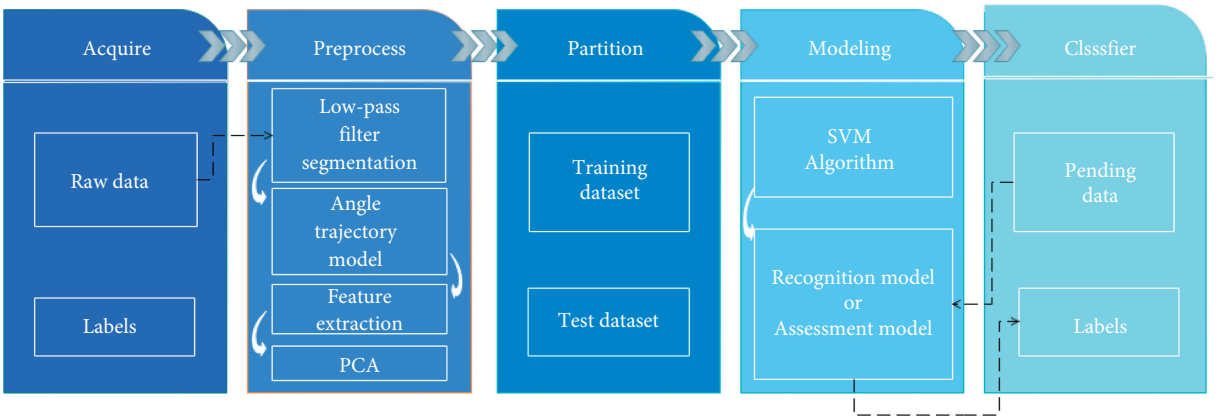


Figure 5: The process of soccer motion recognition and assessment system.

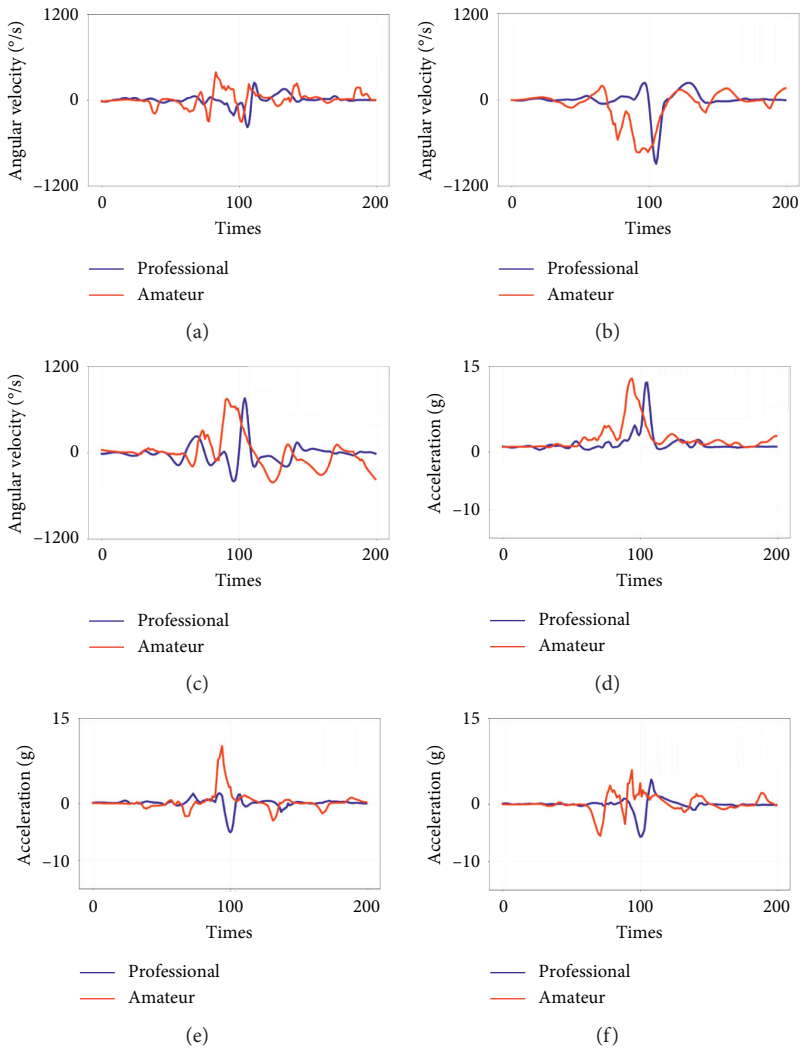


Figure 6: The elites' and amateurs' passing motion data comparison result: (a), (b), and (c) are the contrast of acceleration on the 3-axes and (d), (e), and (f) are the contrast of angular velocity on 3-axes.

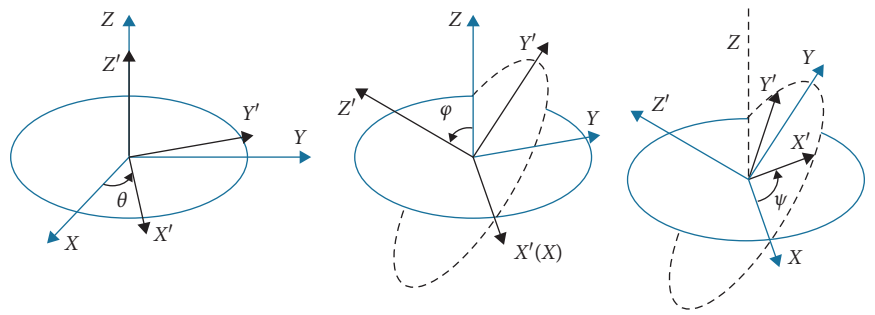


Figure 7: Attitude angle: an exploded view of the process of the attitude angle rotation.

$$\begin{bmatrix} p_0 \\ p_1 \\ p_2 \\ p_3 \end{bmatrix} = \begin{bmatrix} \sin\frac{\psi}{2}\sin\frac{\theta}{2}\sin\frac{\varphi}{2} + \cos\frac{\psi}{2}\cos\frac{\theta}{2}\cos\frac{\varphi}{2} \\ \sin\frac{\psi}{2}\cos\frac{\theta}{2}\cos\frac{\varphi}{2} - \cos\frac{\psi}{2}\sin\frac{\theta}{2}\sin\frac{\varphi}{2} \\ \cos\frac{\psi}{2}\sin\frac{\theta}{2}\cos\frac{\varphi}{2} + \sin\frac{\psi}{2}\cos\frac{\theta}{2}\sin\frac{\varphi}{2} \\ \cos\frac{\psi}{2}\cos\frac{\theta}{2}\sin\frac{\varphi}{2} - \sin\frac{\psi}{2}\sin\frac{\theta}{2}\cos\frac{\varphi}{2} \end{bmatrix}. \quad (2)$$

The Runge–Kutta method [26] which mainly eliminates the complicated process of solving differential equations when the derivatives and initial value information of the equation are known is used to solve four elements:

$$\begin{bmatrix} p_0 \\ p_1 \\ p_2 \\ p_3 \end{bmatrix}_{t+\triangle t} = \begin{bmatrix} p_0 \\ p_1 \\ p_2 \\ p_3 \end{bmatrix}_t + \frac{\triangle t}{2} \begin{bmatrix} -\omega_x \cdot p_1 - \omega_y \cdot p_2 - \omega_z \cdot p_3 \\ +\omega_x \cdot p_0 - \omega_z \cdot p_2 - \omega_y \cdot p_3 \\ +\omega_y \cdot p_0 - \omega_z \cdot p_1 + \omega_x \cdot p_3 \\ -\omega_z \cdot p_0 + \omega_y \cdot p_1 - \omega_x \cdot p_2 \end{bmatrix}, \quad (3)$$

where t means the current time value and $\triangle t$ represents the interval of next data. The following uses the quaternion to solve the attitude angle. The coordinate transformation matrix D of the known object coordinate system to the Earth coordinate system is

$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + 2\cos\frac{\theta}{2} \begin{bmatrix} 0 & -n\sin\frac{\theta}{2} & m\sin\frac{\theta}{2} \\ n\sin\frac{\theta}{2} & 0 & -l\sin\frac{\theta}{2} \\ -m\sin\frac{\theta}{2} & l\sin\frac{\theta}{2} & 0 \end{bmatrix} + 2 \begin{bmatrix} -(m^2+n^2)\sin^2\frac{\theta}{2} & lm\sin^2\frac{\theta}{2} & nl\sin^2\frac{\theta}{2} \\ lm\sin^2\frac{\theta}{2} & -(l^2+n^2)\sin^2\frac{\theta}{2} & mn\sin^2\frac{\theta}{2} \\ nl\sin^2\frac{\theta}{2} & mn\sin^2\frac{\theta}{2} & -(m^2+l^2)\sin^2\frac{\theta}{2} \end{bmatrix}. \quad (4)$$

m, l, and n mean the direction vector projected onto the geographic coordinate along the direction of rotation vector. The rotation of the object corresponds to the rotation around the axis. The trigonometry of a quaternion is $\cos\theta/2 + (l\vec{i} + m\vec{j} + n\vec{k})\sin\theta/2 = \cos\theta/2 + \vec{u}\sin\theta/2$, in which $\vec{i}, \vec{j}, \vec{k}$ denote the unit direction vector in the geographic graticule system:

$$\begin{cases} p_0 = \cos\frac{\theta}{2}, \\ p_1 = l\sin\frac{\theta}{2}, \\ p_2 = m\sin\frac{\theta}{2}, \\ p_3 = n\sin\frac{\theta}{2}. \end{cases} \quad (5)$$

The constructed quaternion describes the fixed-point rotation problem of the substance. Correspondingly, the object system is formed by one-time equivalent rotation of the Earth system. Equation (5) is substituted into D_1:

$$D_1 = \begin{bmatrix} p_0^2 + p_1^2 - p_2^2 - p_3^2 & 2(p_1p_2 - p_0p_3) & 2(p_0p_2 + p_1p_3) \\ 2(p_0p_3 + p_1p_2) & p_0^2 - p_1^2 + p_2^2 - p_3^2 & 2(p_2p_3 - p_0p_1) \\ 2(p_1p_3 - p_0p_2) & 2(p_0p_1 + p_2p_3) & p_0^2 - p_1^2 - p_2^2 + p_3^2 \end{bmatrix}. \quad (6)$$

Let the unit vector in the coordinate system X be $(ex_1, ey_1, ez_1)^T$ and correspond to $(ex_2, ey_2, ez_2)^T$ in Y, and the direction of the projection of the $(ex_2, ey_2, ez_2)^T$ is C_b^n. Suppose there is a vector R whose magnitude on X is $(x, y, z)^T$ and on Y is $(x_2, y_2, z_2)^T$; then, coordinate transformation is $(x, y, z)^T = C_b^n(x_2, y_2, z_2)^T$. The attitude angle transformation matrix is also converted into the coordinate transformation matrix form of object coordinate system to a geographic coordinate system and C_b^n is obtained as follows:

$$
\begin{aligned}
C_b^n &= Rot(z,\psi)Rot(y,\theta)Rot(x,\varphi) \\
&= \begin{bmatrix} \cos\psi & -\sin\psi & 0 \\ \cos\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{bmatrix} \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix} \\
&= \begin{bmatrix} \cos\psi\cos\theta & -\sin\psi\cos\varphi+\cos\psi\sin\theta\sin\varphi & \sin\psi\sin\varphi+\cos\psi\sin\theta\cos\varphi \\ \sin\psi\cos\theta & \cos\psi\cos\varphi+\sin\psi\sin\theta\sin\varphi & -\cos\psi\sin\varphi+\sin\psi\sin\theta\cos\varphi \\ -\sin\theta & \cos\theta\sin\varphi & \cos\theta\cos\varphi \end{bmatrix}.
\end{aligned} \tag{7}
$$

Both D_1 and C_b^n are matrices transforming the object coordinate system attitude into the geographic coordinate system; that is,

$$
\begin{cases}
2(p_1p_3 - p_0p_2) = -\sin\theta = g_1, \\
2(p_0p_1 + p_2p_3) = \sin\varphi\cos\theta = g_2m, \\
p_0^2 - p_1^2 - p_2^2 + p_3^2 = \cos\theta\cos\varphi = g_3, \\
2(p_0p_3 + p_1p_2) = \sin\psi\cos\theta = g_4, \\
p_0^2 + p_1^2 - p_2^2 - p_3^2 = \cos\psi\cos\theta = g_5.
\end{cases} \tag{8}
$$

The equation for solving the attitude angle by the quaternion is

$$
\begin{cases}
\text{roll} = \arctan\dfrac{2(p_0p_3 + p_1p_2)}{1 - 2\left(p_0^2 + p_1^2\right)}, \\
\text{pitch} = \arctan\dfrac{2(p_1p_3 - p_0p_2)}{4(p_0p_3 + p_1p_2)^2 + \left(1 - 2\left(p_0^2 + p_1^2\right)\right)^2}, \\
\text{yaw} = \arctan\dfrac{2(p_0p_1 + p_2p_3)}{1 - 2\left(p_2^2 + p_3^2\right)}.
\end{cases} \tag{9}
$$

The frequency of the sensor repetition rate is 100Hz. Then, the angular velocity value is substituted into the solution attitude angle A_{n-i}. Figure 8 shows the comparison of attitude angle of elites and amateurs.

By combining A_{n-i} with S_{n-i} as a part of R_{n-i}, we extract the basic and morphology features f_{ni} of segments $R(n)_i$, as shown in Table 1.

$F(n) = (f_1, \ldots, f_{mn})$ are merged into one large matrix F of size $n \times m$. The updated data form is shown in Figure 9.

4.2. Principle Component Analysis. The cloud server benefits from reducing memory requirements, computing load, and necessary bandwidth during model implementation. To represent identified variables in compact feature variables, we perform PCA processing. The model with PCA can achieve higher accuracy than other nonlinear dimensionality reduction methods [7]. PCA is essentially a base transformation that makes transformed data have the largest variance, which means the variance between one axis (spindle) and the point is minimized by the rotation of the coordinate axis and the translation of the coordinate origin. Suppose the matrix F with the dimension of $n \times m$, which means that there is total n samples in m-dimensions. F can be decomposed into $U\Sigma V^T$, where U has the same size as F, and the orthogonal matrix V is $m \times m$, Σ is a diagonal matrix of the same size as V. Then, $Y_r U\Sigma_r = F(\Sigma V^T)^{(-1)}\Sigma_r$.

We extracted a total of 34-dimensional features for a vector characterizing, where $F = [F_1, F_2, \ldots, F_{34}]$. After PCA, the obtained new features can be expressed as $Y_r = [Y_{r1}, Y_{r2}, \ldots, Y_{rm}]$ and m represents the calculated dimension:

$$
Y_{rm} = F_1a_{ij} + F_2a_{ij} + \cdots + F_ma_{ij}, \tag{10}
$$

where a_{ij} are eigenvalues of the covariance matrix. Equation (10) can be simplified as follows:

$$
Y_{rm} = F_1a_1 + F_2a_2 + \cdots + F_ma_m. \tag{11}
$$

4.3. Support Vector Machine. SVM is widely applied in the domain of machine learning and pattern recognition as a tool to solve the classification problem. The basic model of SVM aims to maximize the distance of the closet samples which is called support vector to the hyperplane on the eigenspace [27]. In SVM, a training data point is regarded as a p-dimensional vector; the objective is to separate such points using a $(p-1)$-dimensional hyperplane [19]. That is, the training points are mapped to points in space to maximize the width of the gap between the two and several categories by SVM. Compared to other supervised algorithms such as logistics regression (LR) model or Naïve Bayes algorithm [28], the SVM algorithm is more suitable to deal with high-dimensional and linear inseparable problems by freely selecting the parametric model and only use support vector as the classification basis of hyperplane which meets our expectations as our problem is a small sample linear inseparable problem.

The training dataset is set as $T = \{(X_k, y_k) | X \in R^m, y_k \in \{0,1\}_{k=1}^n\}$. Here are total n samples. X represents an m-dimensional matrix, the value of the classification label y_k is 0 or 1, and the current sample is X_k. Therefore, the representation of the optimization objective function is

$$
\begin{aligned}
&\min_{\omega,b} \frac{1}{2}\|\omega\|^2, \\
&\text{s.t.} \quad y_k(\omega \cdot x_k + b) - 1 \geq 0, k = 1, 2, \ldots, N.
\end{aligned} \tag{12}
$$

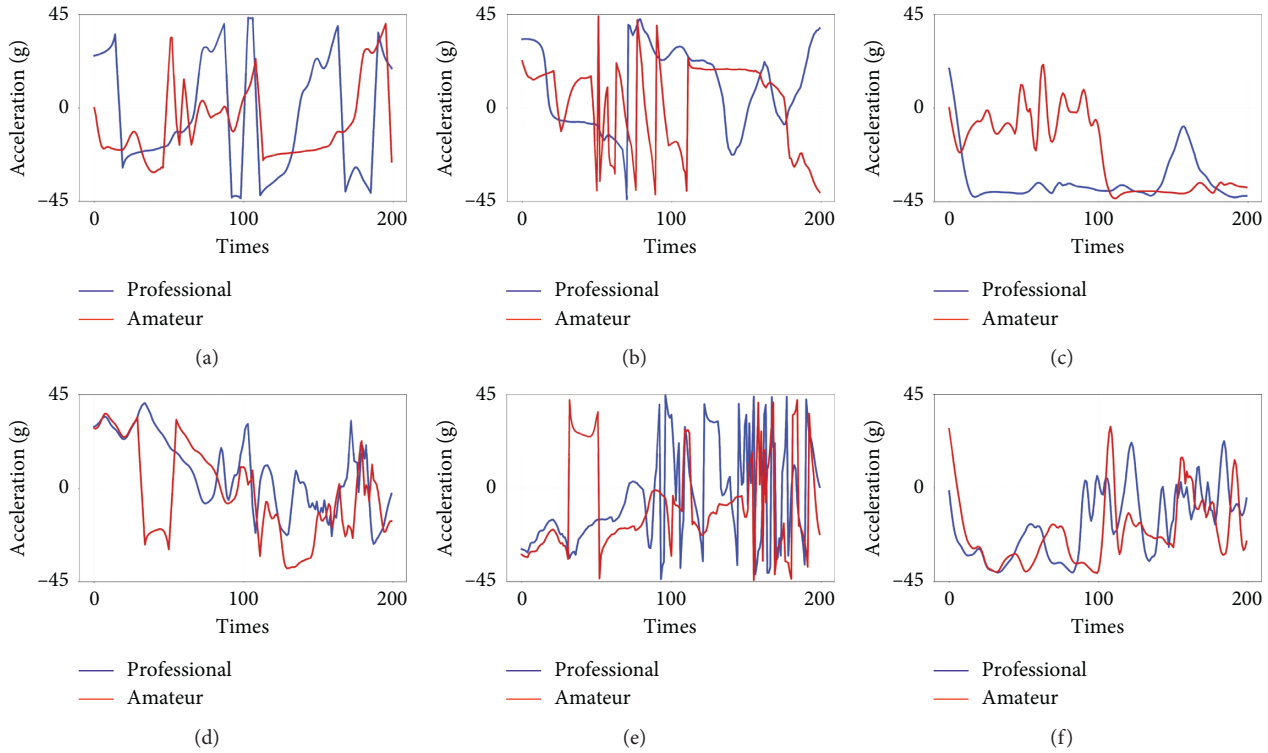


Figure 8: Comparison of elite's and amateur's attitude angular: (a) is the contrast of pitch, (b) is the contrast of roll, (c) is the contrast of yaw when passing, (d) is the contrast of pitch, (e) is the contrast of roll, and (f) is the contrast of yaw when shooting.

Here, ω is a vector on a hyperplane and the offset b of the superflat is along ω from the origin. Problem (12) is a convex quadratic programming problem [29]. According to the convex optimization theory, transform problem (12) into an unconstrained problem. The optimization function can be denoted as

$$L(\omega, b, \alpha) = \frac{1}{2}\|\omega\|^2 - \sum_{k=1}^{N} \alpha_k y_k (\omega \cdot x_k + b) + \sum_{k=1}^{N} \alpha_k, \quad (13)$$

where α_k is a Lagrangian multiplier, $\alpha_k \geq 0$, $(k = 1, 2, 3, \ldots, n)$. The original problem is expressed as

$$\max_{\alpha} \min_{\omega b} L(\omega, b, \alpha). \quad (14)$$

α_k is a Lagrangian multiplier where $\alpha_k \geq 0$, $(k = 1, 2, 3, \ldots, n)$. Solving ω and b as a minimum problem, we can get the value of ω and b:

$$\begin{cases} \omega = \sum_{k=1}^{N} \alpha_k y_k x_k, \\ \sum_{k=1}^{N} \alpha_k y_k = 0. \end{cases} \quad (15)$$

Substituting the obtained solution into the Lagrangian function for a minimize problem, the following optimization function can be obtained after substituting:

$$s.t. \min_{\alpha} \frac{1}{2} \sum_{k=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j y_k y_j (x_k \cdot x_j) - \sum_{k=1}^{N} \alpha_k \sum_{k=1}^{N} \alpha_k y_k = 0 \alpha_k \geq 0, k = 1, 2, \ldots, N. \quad (16)$$

Then, introduce a slack variable ζ_k for (x_k, y_k) for some abnormal sample points make the training set linearly inseparable. The penalty parameter $C \geq 0$. The original issue is described as

Table 1: Basic and morphology features.

Number	Symbol	Description
1	R_{ay}	Root mean square (RMS) of y-axis acceleration
2	R_{wx}	RMS of x-axis angular velocity
3	R_{wy}	RMS of y-axis angular velocity
4	R_{wz}	RMS of z-axis angular velocity
5	D_{ax}	Variance of x-axis acceleration
6	D_{ay}	Variance of y-axis acceleration
7	D_{wy}	Variance of y-axis angular velocity
8	Max_{xy}	Maximum of x-axis acceleration
9	Max_{ay}	Maximum of y-axis acceleration
10	Min_{ay}	Minimum of y-axis acceleration
11	Min_{wz}	Minimum of y-axis angular velocity
12	S_{ax}	Skewness of x-axis acceleration
13	S_{ay}	Skewness of y-axis acceleration
14	S_{az}	Skewness of z-axis acceleration
15	S_{wx}	Skewness of x-axis angular velocity
16	S_{wy}	Skewness of y-axis angular velocity
17	S_{wz}	Skewness of z-axis angular velocity
18	IQR_{ax}	Interquartile range of x-axis acceleration
19	IQR_{ay}	Interquartile range of y-axis acceleration
20	IQR_{az}	Interquartile range of z-axis acceleration
21	IQR_{wx}	Interquartile range of x-axis angular velocity
22	IQR_{wy}	Interquartile range of y-axis angular velocity
23	IQR_{wz}	Interquartile range of z-axis angular velocity
24	Std_{ax}	Standard deviation of x-axis acceleration
25	Std_{wy}	Standard deviation of y-axis angular velocity
26	M_p	Mean value of pitch
27	M_r	Mean value of roll
28	M_y	Mean value of yaw
29	R_p	Root mean square of pitch
30	R_r	Root mean square of roll
31	R_y	Root mean square of yaw
32	Std_p	Standard deviation of pitch
33	Std_r	Standard deviation of raw
34	Std_y	Standard deviation of yaw

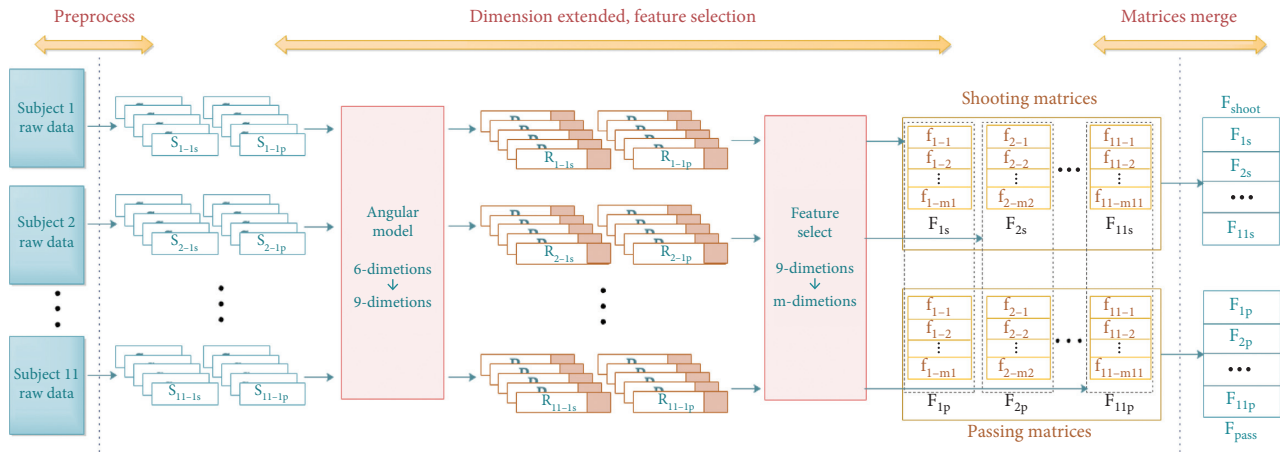


Figure 9: With preprocessing, we obtain effective data S_{n-is} for shooting and S_{n-ip} for passing. Add attitude angle for a suitable data form R_{n-is} for shooting and R_{n-ip} for passing. Select key features F_{is} and F_{ip} that contribute to maximizing classifier success rates. Combine all samples to obtain F_s and F_p.

Table 2: Parameters setting of SVM.

Penalty parameter (C)	γ	Kernel
1	10^{-5}	Linear
500	5×10^{-5}	Polynomial
10^3	10^{-4}	RBF
5×10^3	5×10^{-4}	Sigmoid
10^4	10^{-3}	
5×10^3	5×10^{-3}	

Table 3: Soccer motion recognition accuracies comparison.

Algorithms	Parameter	Accuracy
SVM + angle trajectory model	$C=1,\ \gamma=10^{-4}$	0.9
SVM	$C=1,\ \gamma=5\times10^{-4}$	0.88
KNN	N_Neighbors = 4	0.85
Decision Tree	Min_Samples Split = 3, max depth = 6	0.86

$$\text{s.t.}\min_{\omega,b,\zeta}\frac{1}{2}\|\omega\|^2 + C\sum_{k=1}^{N}\zeta_k y_k\left(\omega\cdot x_k + b\right)\geq 1-\zeta_k, k=1,2,\ldots,N\zeta_k\geq 0, k=1,2,\ldots,N. \tag{17}$$

The ultimate function is obtained after conversion:

$$\text{s.t.}\min_{\alpha}\frac{1}{2}\sum_{k=1}^{N}\sum_{j=1}^{N}\alpha_k\alpha_j y_k y_j\left(x_k\cdot x_j\right) - \sum_{k=1}^{N}\alpha_k\sum_{k=1}^{N}\alpha_k y_k = 00\leq\alpha_k\leq C, k=1,2,\ldots,N. \tag{18}$$

There are two labels: passing and shooting represented by 0 and 1, respectively, in the action classification experiment. We obtained 264 sets for passing and 250 sets for shooting. The datasets are randomly divided into training and test sets, 80% of them used for training and the rest used as test data. Randomly selected parameters are shown in Table 2. The selectable parameter C ranges from 1 to 50000; γ range is from 0.00001 to 0.05, and the kernel function includes Linear, RBF, Sigmoid, and Polynomial. In order to solve overfitting, 3-fold cross is adopted. Experimental results show that $C = 1$, γ is 0.0001, and linear kernel function is used in the optimal model.

We use the proposed SVM-based classifier to recognize the shooting and passing. Decision Tree [30] and K-nearest neighbor algorithm (KNN) [31] are the competitors in our experiments. The accuracy of different models for soccer motion recognition can be found in Table 3. Apparently, our proposed SVM-based algorithm has better performance than other competitors. The one with the attitude angle has a better performance than the others.

Table 4 shows the recognition and assessment results of our proposed system. The recognition accuracies for passing and shooting are 85.7% and 88.5% which illustrated a clear distinction between passing and shooting. The accuracy of classifying different behaviors can be 87.1% on average. This result suggests our system works.

4.4. Talent Selection System. Similar to the above analysis, the labels are elites and amateurs in the talent selection experiment. The selectable parameters are still randomly selected from Table 2. Thre-fold cross-validation is used again to prevent overfitting. Finally, the best performing classifiers are selected. Experimental results show that $C = 1$ and linear kernel function applies to all optimal models. As for γ, different from motion recognition, the value is 0.00005.

As shown in Table 5, the accuracies of shooting level classifiers are higher than others. There is a clear difference between different level participants when shooting. The results show that the SVM with the attitude angle model also has satisfactory accuracy.

The results in Table 6 are all inferior to others, which means there might be little difference between participants when passing. As listed in Tables 3, 5, and 6, comparing with KNN and Decision Tree, SVM demonstrated superiority in dealing with the linear inseparable problem, and the one with attitude angle is better than the other.

As shown in Table 7, the assessment results of proposed system. Our model with attitude angle features obviously performs better than typical model. The professionals and amateurs are 88.7% and 85.1%, respectively, for passing. For the shooting action, the accuracy of the model is 93%, which also signifies that, for the shooting action, the skill gap between the professional and the amateur is bigger than the passing action.

Table 4: The motion recognition result.

Criterion	Motion Classification		Average
	Passing	Shooting	
Accuracy	0.857	0.885	0.871
Recall	0.96	0.92	0.94
$F1$-score	0.906	0.902	0.904

Table 5: Shooting level assessment accuracies comparison.

Algorithms	Parameter	Accuracy
SVM + angle trajectory model	$C = 1,\ \gamma = 5 \times 10^{-5}$	0.93
SVM	$C = 1,\ \gamma = 10^{-4}$	0.91
KNN	N_Neighbors = 4	0.87
Decision tree	Min_Samples Split = 3, max depth = 6	0.86

Table 6: Passing level classification accuracies comparison.

Algorithms	Parameter	Accuracy
SVM + angle trajectory model	$C = 1,\ \gamma = 5 \times 10^{-5}$	0.87
SVM	$C = 1,\ \gamma = 10^{-4}$	0.86
KNN	N_Neighbors = 4	0.84
Decision Tree	Min_Samples Split = 3, max depth = 6	0.79

Table 7: The assessment result for passing and shooting.

Criterion	Skill level of passing		Average	Skill level of shooting		Average
	Elites	Amateurs		Elites	Amateurs	
Accuracy	0.887	0.851	0.869	0.959	0.902	0.931
Recall	0.94	0.8	0.87	0.94	0.92	0.93
$F1$-score	0.913	0.825	0.869	0.949	0.911	0.93

5. Conclusion and Discussion

An AIoT system to recognize different soccer (football) motions and assess the skill levels of soccer players was proposed in this paper. The proposed IoT system consists of wearable devices, a mobile device (e.g., a mobile phone or tablet), and a cloud-based data processing platform. In this proposed system, a WWSD of MEMS motion sensors is used to collect raw data from soccer players. By using BTLE technology, the data is transmitted to the cloud-based data processing platform, which analyzes the data and outputs the results to a mobile device in real-time. A SVM model classification algorithm with an ankle-based attitude angle model is proposed to recognize different motions and assess different skill levels. This intelligent system can be a new paradigm and emerging technology based on wearable sensors, attaining the experience-driven to data-driven transition in the field of sports training and talent selection, and can be easily extended to analyze other foot-related sports motions and skill levels.

In this paper, we propose a system for recognizing soccer motions based on IoT devices, which can recognize users' soccer motions and evaluate the quality of the motions; however, this system still needs to be improved in terms of storage performance and computing speed in the cloud system. The increasing popularity of IoT applications demands the computing power of IoT systems, and edge computing is one of the main methods to enhance the computing speed of IoT applications [32]. In future research, we will try to combine this recognition system with edge computing as well as [33–36], in which authors used edge computing in combination with IoT devices to increase the computational speed of the system and thus reduce its response time.

Data Availability

The soccer motion dataset used to support the findings of this study are currently under embargo while the research findings are commercialized. Requests for data, 12 months after publication of this article, will be considered by the corresponding author.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported in part by the Natural Science Foundation of China under grant 61702336, the Hubei

Provincial Natural Science Foundation of China under grant 2020CFB629, the Fundamental Research Funds for the Central Universities under grants CZT20027 and CSQ21018, and the Research Start-Up Funds of South-Central University for Nationalities under grant YZZ18006.

References

[1] M. Bhatia, "Iot-inspired framework for athlete performance assessment in smart sport industry," *IEEE Internet of Things Journal*, vol. 8, no. 12, pp. 9523–9530, 2020.

[2] Q. Zhang, Y. Zhang, C. Li, C. Yan, Y. Duan, and H. Wang, "Sport location-based user clustering with privacy-preservation in wireless iot-driven healthcare," *IEEE Access*, vol. 9, pp. 12906–12913, 2021.

[3] A. Rajšp and I. Fister, "A systematic literature review of intelligent data analysis methods for smart sport training," *Applied Sciences*, vol. 10, no. 9, p. 3013, 2020.

[4] J. Haladjian, D. Schlabbers, S. Taheri, M. Tharr, and B. Bruegge, "Sensor-based detection and classification of soccer goalkeeper training exercises," *ACM Transactions on Internet Technology*, vol. 1, no. 2, pp. 1–20, 2020.

[5] X. Bu, "Human motion gesture recognition algorithm in video based on convolutional neural features of training images," *IEEE Access*, vol. 8, pp. 160 025–160 039, 2020.

[6] Y. Wang, M. Chen, X. Wang, R. H. M. Chan, and W. J. Li, "Iot for next-generation racket sports training," *IEEE Internet of Things Journal*, vol. 5, no. 6, pp. 4558–4566, 2018.

[7] Y. Wang, Y. Zhao, R. H. M. Chan, and W. J. Li, "Volleyball skill assessment using a single wearable micro inertial measurement unit at wrist," *IEEE Access*, vol. 6, pp. 13 758–813 765, 2018.

[8] N. F. Ghazali, N. Shahar, N. A. Rahmad, N. A. J. Sufri, M. A. As'ari, and H. F. M. Latif, "Common sport activity recognition using inertial sensor," in *Proceedings of the IEEE 14th International Colloquium on Signal Processing Its Applications (CSPA)*, pp. 67–71, Malaysia, March 2018.

[9] M. Ramanathan, W.-Y. Yau, and E. K. Teoh, "Human action recognition with video data: research and evaluation challenges," *IEEE Transactions on Human-Machine Systems*, vol. 44, no. 5, pp. 650–663, 2014.

[10] J. K. Aggarwal and M. S. Ryoo, "Human activity analysis," *ACM Computing Surveys*, vol. 43, no. 3, pp. 1–43, 2011.

[11] M. Stikic, D. Larlus, S. Ebert, and B. Schiele, "Weakly supervised recognition of daily life activities with wearable sensors," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 12, pp. 2521–2537, 2011.

[12] J. Wang, Y. Chen, S. Hao, X. Peng, and L. Hu, "Deep learning for sensor-based activity recognition: a survey," *Pattern Recognition Letters*, vol. 119, pp. 3–11, 2019.

[13] K. Wang, J. He, and L. Zhang, "Attention-based convolutional neural network for weakly labeled human activities' recognition with wearable sensors," *IEEE Sensors Journal*, vol. 19, no. 17, pp. 7598–7604, 2019.

[14] C. Chen, R. Jafari, and N. Kehtarnavaz, "Improving human action recognition using fusion of depth camera and inertial sensors," *IEEE Transactions on Human-Machine Systems*, vol. 45, no. 1, pp. 51–61, 2015.

[15] K. Kui Liu, C. Chen Chen, R. Jafari, and N. Kehtarnavaz, "Fusion of inertial and depth sensor data for robust hand gesture recognition," *IEEE Sensors Journal*, vol. 14, no. 6, pp. 1898–1903, 2014.

[16] A. Mannini, S. S. Intille, M. Rosenberger, A. M. Sabatini, and W. Haskell, "Activity recognition using a single accelerometer placed at the wrist or ankle," *Medicine & Science in Sports & Exercise*, vol. 45, no. 11, pp. 2193–2203, 2013.

[17] A. Mannini, M. Rosenberger, W. L. Haskell, A. M. Sabatini, and S. S. Intille, "Activity recognition in youth using single accelerometer placed at wrist or ankle," *Medicine & Science in Sports & Exercise*, vol. 49, no. 4, pp. 801–812, 2017.

[18] M. Shoaib, S. Bosch, O. Incel, H. Scholten, and P. Havinga, "Complex human activity recognition using smartphone and wrist-worn motion sensors," *Sensors*, vol. 16, no. 4, p. 426, 2016.

[19] J. P. Varkey, D. Pompili, and T. A. Walls, "Human motion recognition using a wireless sensor-based wearable system," *Personal And Ubiquitous Computing*, vol. 16, pp. 897–910, 2016.

[20] J. S. Akins, N. R. Heebner, M. Lovalekar, and T. C. Sell, "Reliability and validity of instrumented soccer equipment," *Journal of Applied Biomechanics*, vol. 31, no. 3, pp. 195–201, 2015.

[21] A. Meamarbashi and S. Hossaini, "Application of novel inertial technique to compare the kinematics and kinetics of the legs in the soccer instep kick," *Journal of Human Kinetics*, vol. 23, no. 1, pp. 5–13, 2010.

[22] BMI160 Datasheet, *Document Bst-Bmi160-Ds000-07*, Bosch Sensortec, Reutlingen, Germany, 2015.

[23] G. Okeyo, L. Chen, H. Wang, and R. Sterritt, "Dynamic sensor data segmentation for real-time knowledge-driven activity recognition," *Pervasive and Mobile Computing*, vol. 10, pp. 155–172, 2014.

[24] S. Ioannidou and G. Pantazis, "Helmert transformation problem. from euler angles method to quaternion algebra," *ISPRS International Journal of Geo-Information*, vol. 9, no. 9, p. 494, 2020.

[25] N. Jouppi, C. Young, N. Patil, and D. Patterson, "Motivation for and evaluation of the first tensor processing unit," *IEEE Micro*, vol. 38, no. 3, pp. 10–19, 2018.

[26] H. Ranocha, M. Sayyari, L. Dalcin, M. Parsani, and D. I. Ketcheson, "Relaxation Rkutta methods: fully discrete explicit entropy-stable schemes for the compressible euler and Navier–Stokes equations," *SIAM Journal on Scientific Computing*, vol. 42, no. 2, pp. A612–A638, 2020.

[27] A. M. Andrew, *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods by Nello Christianini and John Shawe-Taylor*, Cambridge University Press, Cambridge, UK, 2000.

[28] X. Yang and Y. L. Tian, "Eigen joints-based action recognition using Naïve-Bayes-nearest-neighbor," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, pp. 14–19, Providence, RI, USA, June 2012.

[29] T. V. Gestel, B. Baesens, P. V. Dijcke, J. A. K. Suykens, and T. Alderweireld, "Linear and nonlinear credit scoring by combining logistic regression and support vector machines," *Social Science Electronic Publishing*, vol. 1, no. 4, 2005.

[30] A. Jindal, A. Dua, K. Kaur, M. Singh, N. Kumar, and S. Mishra, "Decision tree and SVM-based data analytics for theft detection in smart grid," *IEEE Transactions on Industrial Informatics*, vol. 12, no. 3, pp. 1005–1016, 2016.

[31] P. Thanh Noi and M. Kappas, "Comparison of random forest, k-nearest neighbor, and support vector machine classifiers for land cover classification using sentinel-2 imagery," *Sensors*, vol. 18, no. 1, p. 18, 2018.

[32] H. Li, K. Ota, and M. Dong, "Learning iot in edge: deep learning for the internet of things with edge computing," *IEEE network*, vol. 32, no. 1, pp. 96–101, 2018.

[33] W. Rafique, L. Qi, I. Yaqoob, M. Imran, R. U. Rasool, and W. Dou, "Complementing iot services through software defined networking and edge computing: a comprehensive survey," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 3, pp. 1761–1804, 2020.

[34] L. Yang, H. Yao, J. Wang, C. Jiang, A. Benslimane, and Y. Liu, "Multi-UAV-enabled load-balance mobile-edge computing for iot networks," *IEEE Internet of Things Journal*, vol. 7, no. 8, pp. 6898–6908, 2020.

[35] C. Gong, F. Lin, X. Gong, and Y. Lu, "Intelligent cooperative edge computing in internet of things," *IEEE Internet of Things Journal*, vol. 7, no. 10, pp. 9372–9382, 2020.

[36] B. C. Kavitha, R. Vallikannu, and K. S. Sankaran, "Delay-aware concurrent data management method for iot collaborative mobile edge computing environment," *Microprocessors and Microsystems*, vol. 74, Article ID 103021, 2020.

Hindawi
Journal of Healthcare Engineering
Volume 2021, Article ID 5534607, 17 pages
https://doi.org/10.1155/2021/5534607

Research Article

Environmental Design Strategies to Decrease the Risk of Nosocomial Infection in Medical Buildings Using a Hybrid MCDM Model

Lei Xiong,[1] **Ge Sheng**,[1] **Zi-Mu Fan**,[2] **Hua Yang**,[3] **Feng-Jang Hwang**,[4] **and Bo-Wei Zhu**[2]

[1]*School of Architecture and Applied Art, Guangzhou Academy of Fine Arts, Guangzhou 511400, China*
[2]*Faculty of Humanities and Arts, Macau University of Science and Technology, Macau 999078, China*
[3]*Mural and Complex Material Painting Department, Hubei Institute of Fine Arts, Wuhan 430205, China*
[4]*School of Mathematical and Physical Sciences, University of Technology Sydney, Sydney, Australia*

Correspondence should be addressed to Bo-Wei Zhu; zhubowei301@gmail.com

Received 30 January 2021; Accepted 27 May 2021; Published 16 June 2021

Academic Editor: Chia-Yu Lin

The prevention and control of nosocomial infection (NI) are becoming increasingly difficult, and its mechanism is becoming increasingly complex. A globally aging population means that an increasing proportion of patients have a susceptible constitution, and the frequent occurrence of severe infectious diseases has also led to an increase in the cost of prevention and control of NI. Medical buildings' spatial environment design for the prevention of NI has been a hot subject of considerable research, but few previous studies have summarized the design criteria for a medical building environment to control the risk of NI. Thus, there is no suitable evaluation framework to determine whether the spatial environment of a medical building is capable of inhibiting the spread of NI. In the context of the global spread of COVID-19, it is necessary to evaluate the performance of the existing medical building environment in terms of inhibiting the spread of NI and to verify current environmental improvement strategies for the efficient and rational use of resources. This study determines the key design elements for the spatial environment of medical buildings, constructs an evaluation framework using exploratory factor analysis, verifies the complex dominant influence relationship, and prioritizes criteria in the evaluation framework using the decision-making trial and evaluation laboratory- (DEMATEL-) based analytical network process (ANP) (DANP). Using representative real cases, this study uses the technique for order preference by similarity to ideal solution (TOPSIS) to evaluate and analyze the performance with the aspiration level of reducing the NI risk. A continuous and systematic transformation design strategy for these real cases is proposed. The main contributions of this study include the following: (1) it creates a systematic framework that allows hospital decision-makers to evaluate the spatial environment of medical buildings; (2) it provides a reference for making design decisions to improve the current situation using the results of a performance evaluation; (3) it draws an influential network relation map (INRM) and the training of influence weights (IWs) for criteria. The sources of practical problems can be identified by the proposed evaluation framework, and the corresponding strategy can be proposed to avoid the waste of resources for the prevention of epidemics.

1. Introduction

Nosocomial Infection (NI) is a nonprimary infection that is contracted during hospital stay. It is also known as healthcare-acquired infection [1]. In China and the United States, the annual number of NI-related deaths is 1.25–3.5 million [2], and the direct economic loss due to Nosocomial Infection is about US$28–33 billion per year [3]. Nosocomial Infection increases the burden on the public health system and increases the medical risk and recovery time for patients. The prevention and control of Nosocomial Infection is becoming more difficult and the occurrence mechanism is becoming more complex. The abuse of antibiotics in many countries has led to an increase in multidrug resistance for bacteria, so the prevention and treatment of Nosocomial Infection is more difficult [4]. A globally aging population has also led to a continuous increase in the proportion of patients with a susceptible constitution [5]. The complex

treatment process increases the risk of exposure to Nosocomial Infection and the frequent occurrence of severe infectious diseases increases the cost of the prevention and control of Nosocomial Infection [6]. In terms of SARS and COVID-19, Nosocomial Infection is the main mode of contagious infection in the early stage, so there are a large number of infections among medical staff. Between January 1, 2020, and January 28, 2020, 138 patients who were diagnosed with New Coronavirus pneumonia in Zhongnan Hospital in Wuhan University. It is estimated that about 41% of these infections were Nosocomial Infections and 29% involved medical staff.

In terms of architectural design, reducing the risk of NI is a main safety concern of medical architectural design. Since the 19th century, the ventilation, light, cleaning, noise, and disinfection of building space have been the basic elements that are used to control Nosocomial Infection in hospital buildings [7]. In a Nightingale ward, there is spatial separation between patients and medical staff, a specific distance between patients' beds, air convection by opening windows on both sides of the room, and waste is treated separately [8]. Due to the widespread use of antibiotics, there was not much progress in hospital infection prevention and control in medical building design after this original strategy.

Since the beginning of the 21st century, Nosocomial Infection has been increasingly caused by drug-resistant bacteria in the West. Good environmental design can prevent Nosocomial Infection, so a large number of interdisciplinary studies have been conducted. Previous studies show that medical buildings play an important role in the intervention of Nosocomial Infection, which involves almost the entire cycle of Nosocomial Infection [9–11]. Many practical cases show that good environmental design reduces Nosocomial Infection.

In terms of global COVID-19 prevention and control, many areas must construct Disaster Prevention Hospital Buildings using extremely limited resources. The existing medical buildings must also be transformed in terms of design to address new challenges [12–14]. From the perspective of comprehensive architectural design and public health, there is a relative lack of literature on medical architectural design elements that reduce the risk of Nosocomial Infection. Previous studies mostly focus on hospital management and safety behavior management [15, 16], or pathological analysis from the perspective of epidemiology and contagious infection [17]. At present, no systematic architectural design elements have been proposed, so it is impossible to construct an evaluation framework to determine whether the spatial environment in medical buildings reduces the risk of NI. However, with the continuous improvement in NI prevention and control, the occurrence mechanism for NI is becoming increasingly complex. The effectiveness and efficiency of the spatial environment of medical buildings in terms of NI intervention must be determined to verify whether the required level of prevention and intervention is attainable. In terms of the global spread of COVID-19, it is more necessary to study design decision-making for scenarios using the results of current research, in order to improve the spatial environment in medical buildings and minimize the NI risk in medical buildings.

This study proposes that the construction of this evaluation model requires the embedding of a mutually supportive process for dynamic thinking. A systematic and mutual influence perspective is required to establish and confirm the effect and priority for each evaluation criterion. The contributions of the evaluation study are confined to the ranking and selection of actual cases and provide a basis of reference for further formulation of an improvement strategy [18, 19].

This study determines the key design elements for the spatial environment of medical buildings to reduce the risk of NI and constructs the evaluation criterion framework. In terms of the prevention of the current epidemic, the complex interaction relationship and priority for elements in the evaluation framework are verified. Representative actual medical building cases are used to reduce the risk of NI as the aspiration level for performance evaluation analysis. For each case, a continuous and systematic design strategy is proposed.

This study first summarizes the environmental design elements that affect the NI risk for medical buildings in an analysis of the key literature and uses exploratory element analysis (EFA) to define the core design elements and element categories to build the evaluation framework. DANP is then used to assign Influence Weights (IWs) to the evaluation elements and clarify the Influential Network Relationship Map (INRM) for the evaluation criteria. TOPSIS is used to evaluate and analyze the performance for actual cases. The results of evaluation and analysis show the influence network relationship and priority for elements, and are used to identify the source of key problems and to formulate continuous and systematic environmental transformation design strategies for actual cases.

The main contributions of this study are as follows:

(1) The planning and design of spatial environment can affect the behavior of individuals in medical buildings and the activity of infection sources. It can also reduce the risk of Nosocomial Infection in medical buildings. This study uses a review of the relevant literature and analysis of expert interviews to summarize the environmental design criteria for medical buildings that can reduce the risk of Nosocomial Infection. These criteria are the prevention and control measures for Nosocomial Infection.

(2) In order to control the risk of Nosocomial Infection, this study confirms the reliability of content and the structure of the environmental design criteria, and constructs a hierarchical evaluation framework to determine the current actual environmental planning and design status of medical buildings.

(3) To allow interaction between criteria in the same dimension, this study draws an Influential Network Relationship Map (INRM) and assigns Influence Weights (IWs) to each criterion. In terms of the reconstruction of medical buildings, this allows architects to identify significant problems and the source of actual problems.

(4) Using three general hospitals in the Pearl River Delta region of China as actual cases, this study constructs an evaluation framework to evaluate the risk of Nosocomial Infection in actual cases, from the perspective of building environment planning and design. The performance evaluation results and the Influential Network Relationship Map (INRM) and Influence Weights (IWs) are used to study current environmental improvement design strategies for each case from the "source of impact" issues due to limited resources.

2. Literature Review

2.1. Nosocomial Infection in Medical Buildings. The process of infection is the process by which pathogens enter the human body from the outside, break through its immune system, and cause disease. From the source of infection, infection includes endogenous and exogenous infection. Nosocomial Infection is mainly an exogenous infection [20]. In previous studies, the triangular chain of "infection source environment host" is identified as the main process for Nosocomial Infection. This process involves the source of infection entering the medical building, then entering the human body in a specific way to form an infection, and finally the infected person becomes a carrier and a secondary source of infection (respiratory, body fluid, and environmental pollution).

The main mechanisms for Nosocomial Infection in medical buildings are air, contact, and water transmission [11, 21, 22]. Lenfestey et al. [3] determined that contact transmission is the most common form, accounting for about 80% of all Nosocomial Infections. Air transmission and water transmission are also factors. However, these three forms of transmission often occur intermittently. The main sites of infection are the skin system, the respiratory system, and the digestive system.

Contact transmission can involve direct transmission, indirect transmission, or droplet transmission. Direct transmission occurs through direct contact between individuals, indirect transmission involves an environmental surface as a carrier, and infection occurs when a susceptible person frequently contacts the contaminated surface (with pathogens) [23]. The distance over which droplet transmission operates is very small. It mainly infects the surrounding population through droplets that are discharged from the mouth and nose of the carrier. The liquid particles are large, so they are not suspended in the air for a long time. Research shows that their range is within 1.8 m of the infected person. In contrast to droplet transmission, air transmission occurs after the pathogen particles in the air enter the human body. It may be suspended for a long time and spread with the airflow. Therefore, this route of transmission is the most difficult to prevent, especially in confined spaces. The pathogens for this route of transmission can come from carriers, or the environment, in water, dust, or biological excreta [24]. These pathogens are mainly transmitted by aerosol and dust. The former are generally dry-weather-resistant pathogens, such as measles virus and tuberculosis. Dust transmission is relatively rare, often involves fungal pathogens, and mostly occurs during building renovation.

Although there are only few cases in which Nosocomial Infection is caused by water transmission, pathogens breed easily in water. The ideal temperature is different for different pathogens [25]. The specific mechanisms for water transmission include water contact, drinking, and aerosol infection. The pathogen for this route of transmission travels mainly through daily contact during washing, showering, or drinking, and by aerosol inhalation in the body [26, 27]. These three transmission routes mainly occur in nursing units, wards, emergency halls, and other places. Contact transmission can also occur in infusion centers, and outpatient and laboratory departments. Air transmission can also occur in infection departments and water transmission can occur in public toilets.

2.2. The Effect of the Spatial Environment on Nosocomial Infection. The spatial environment in medical buildings is the main carrier of Nosocomial Infection sources. Physical conditions, such as light, temperature, humidity, air flow, and cleanliness, have a significant effect on the activity of pathogens. Previous studies show that ultraviolet light in natural light reduces the survival time for most bacteria and inhibits reproduction [28]. COVID-19 has been the subject of many studies since its outbreak [29, 30]. The effect of environmental temperature and humidity on COVID-19 activity has been studied. Goswami et al. [29] collected daily data for New Coronavirus infection cases and climatic elements in India. Regression analysis shows that the interaction between mean temperature and mean relative humidity has a specific effect on the occurrence of COVID-19. The environmental temperature and humidity in medical buildings can be precisely controlled using effective design measures. Changes in air distribution and composition in hospital air also affect the reproduction and activity of NCV [31]. A clean environmental interface for public buildings is also a basic strategy to reduce the risk of new crown infection at a construction site. Recently, many hotels, schools, and hospitals have also updated their cleaning procedures and now use advanced cleaning technologies to improve disinfection (e.g., electrostatic spray or ultraviolet technology).

The spatial environment is also affected by people's behavior. Streamlining and functional design affect the behavior path and unnecessary contact and crossing can be controlled [8]. This is an effective way to reduce the incidence of Nosocomial Infection. More specifically, the chances of contact with the source of infection can be lessened by affecting individuals' behavior and it is the key element of design to restrict the moves of the carriers and susceptible people. This controls Nosocomial Infections and is a basic principle of modern hospital design. It involves the separation of doctor and patient, the separation of cleaning and contamination, and the separation of internal and external. Controlling the scope of action and protecting the susceptible population can also reduce the risk of exposure to Nosocomial Infection.

3. Methodology and Process

The research procedure for this study is shown in Figure 1, which shows the research methods and technical routes that are used in different stages of this study and the corresponding research problems to be solved. The evaluation framework was initially established through a literature review, and an exploratory element analysis (EFA) was used for pretesting to determine whether the structure of the evaluation framework is accurate and whether the evaluation criteria are effective, in order to construct the evaluation framework to determine the risk of Nosocomial Infection in hospital buildings. DANP is used to determine the impact relationship between the evaluation criteria using expert domain knowledge, an Influential Network Relation Map (INRM) is drawn, and Influential Weights (IWs) are assigned to each criterion. In terms of actual case evaluation and analysis, this study uses three hospital buildings in Guangzhou as actual cases to conduct performance evaluation and analysis. Using the performance evaluation questionnaire, TOPSIS is used to determine the current risk of Nosocomial Infection risk for each case and the performance of each case in terms of each criterion. Using the results for each research stage, the influence network relationship and priority of elements are determined to identify the source behind the problems and to formulate continuous and systematic environmental transformation design strategies for these cases.

3.1. Exploratory Factor Analysis (EFA). EFA is used to determine the number of elements that affect variables and to analyze which variables go together [32]. The hypothesis for EFA is that there are m common latent elements to be discovered in the data set, and the goal is to identify the smallest number of common elements that account for the correlations [33]. The dependent variables are surface attributes and the underlying structures (factors) are internal attributes [34]. Common elements are those that affect more than one of the surface attributes, and specific elements are those which only affect a particular variable [34]. The main procedure for principal component analysis is described in the following steps when applying exploratory element analysis:

Step (E1): determine the correlation matrix **R** or variance–covariance matrix for the objects to be assessed.

Step (E2): determine the eigenvalues λ_k, k= $1, 2, ..., m$ and eigenvectors $\beta_k = [\beta_{1k}, ..., \beta_{ik}, ..., \beta_{pk}]$ to assess the element loading $a_{ik} = \sqrt{\lambda_k}\beta_{ik}$ and the number of elements m.

Step (E3): consider the eigenvalue ordering $\lambda_1 > ... > \lambda_k > ... > \lambda_m$, where $\lambda_m > 1$, to determine the number of common elements, and specify the number of common elements to be extracted using a predetermined criterion.

Step (E4): according to [35], use the varimax element to determine the rotated element loading matrix, which provides additional insights for the rotation of the element-axis.

Step (E5): name the element in terms of the combination of manifest variables.

3.2. The DEMATEL-Based ANP (DANP) Method. In previous studies, the analysis methods to determine the relative importance of evaluation criteria for a spatial environment use expert domain knowledge, such as an Analytic Hierarchy Process (AHP) or Analytic Network Process (ANP), and public preference judgment, such as choice-based joint analysis. Many traditional analysis techniques that are used in previous studies must allocate the weight of criteria based on the assumption that the evaluation criteria are independent of each other; so, the influence source of real problems can be neglected when formulating improvement strategies using the analysis results for the current situation evaluation. In order to broaden the assumption of independence, the field of operational research uses the DEMATEL for a multicriteria decision-making model. In combination with the basic concept of ANP, it is used to train the weight of evaluation indexes by allowing the mutual influence between various evaluation indexes [18, 19, 36]. The main procedure for this method involves the following steps:

Step (D1): establish the direct influence relation matrix **E**. Data are obtained using a questionnaire and the scales for the questionnaire involve an integer score of 0, 1, 2, 3, or 4, where 0 represents absolutely no influence and 4 represents very high influence, according to the natural language element in linguistics. The respondent is assumed to be an expert in the field, and pairwise comparison is used to determine the degree of influence of the element and show the degree to which each element i affects each other element j. This matrix must be an $n \times n$ nonnegative matrix. The results for H experts are used to construct the direct influence relation matrix **E**, as shown in Equation (1), and the direct influence relation matrix for each expert is $\mathbf{E}^h = [e_{ij}^h]_{n\times n}$, $h = 1, 2, ..., H$.

$$\mathbf{E} = \begin{bmatrix} e_{11} & \cdots & e_{1j} & \cdots & e_{1n} \\ \vdots & & \vdots & & \vdots \\ e_{i1} & \cdots & e_{ij} & \cdots & e_{in} \\ \vdots & & \vdots & & \vdots \\ e_{n1} & \cdots & e_{nj} & \cdots & e_{nn} \end{bmatrix}. \quad (1)$$

Step (D2): construct the average direct influence relation matrix **A**. The average scores for the H experts **are** $a_{ij} = (1/H)\sum_{h=1}^{H} e_{ij}^h$. The average matrix is called the average direct influence relation matrix **D** and represents the degree of influence that one criterion exerts on another and the degree of influence that the criterion receives from another, as shown in

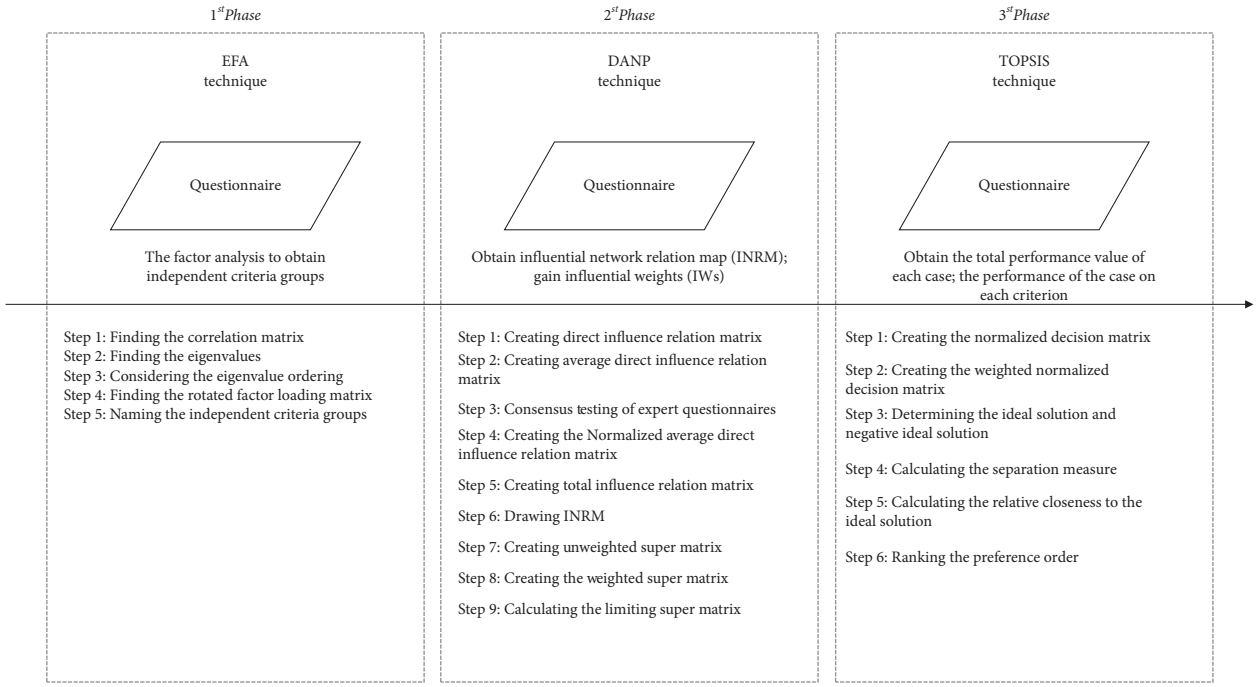


Figure 1: Research procedure for the multiple attribute decision-making model.

$$\mathbf{A} = \begin{bmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{bmatrix}. \tag{2}$$

Step (D3): determine the consensus. The value for consensus is estimated using Equation (3), which represents the degree of consensus for the experts. The statistical threshold for the average gap ratio is 5% and a value less than 5% implies a confidence level of more than 95%, which also represents a stable system. If a system is unstable, the first phase must be implemented again to verify whether data collection is correct and whether the number of experts is sufficient.

$$\text{average gap-ratio in consensus } (\%) = \frac{1}{n(n-1)} \sum_{i=1}^{n} \sum_{j=1}^{n} \left(\frac{\left| a_{ij}^{H} - a_{ij}^{H-1} \right|}{a_{ij}^{H}} \right) \times 100\%. \tag{3}$$

Step (D4): formulate the normalized average direct influence relation matrix **D**. The matrix **D**, which is acquired by normalizing the matrix **A**, is derived from Equations (4) and (5), where all principal diagonal elements are equal to 0:

$$\mathbf{D} = b \cdot \mathbf{A}, \tag{4}$$

$$b = \min \left\{ \frac{1}{\max_{1 \le i \le n} \sum_{j=1}^{n} a_{ij}}, \frac{1}{\max_{1 \le j \le n} \sum_{i=1}^{n} a_{ij}} \right\}. \tag{5}$$

Step (D5): construct the total influence relation matrix **T**. There is a continuous decrease in the indirect effects of problems with the coefficients of the matrix **D**, e.g., $\mathbf{D}^2, \ldots, \mathbf{D}^{\infty}$, and $\lim_{q \longrightarrow \infty} \mathbf{D}^q = [0]_{n \times n}$, for Equation (6), where **I** is a $n \times n$ unit matrix. The total influence relation matrix **T** is a $n \times n$ matrix that is defined by $\mathbf{T} = [t_{ij}]_{n \times n}$, as shown in Equation (7):

$$\lim_{q \longrightarrow \infty} \left(\mathbf{I} + \mathbf{D} + \mathbf{D}^2 + \ldots + \mathbf{D}^q \right) = (\mathbf{I} - \mathbf{D})^{-1}. \tag{6}$$

$$\begin{aligned} \mathbf{T} &= \mathbf{D}(\mathbf{I} - \mathbf{D})^{-1}, \\ \lim_{q \longrightarrow \infty} \mathbf{D}^q &= [0]_{n \times n}. \end{aligned} \tag{7}$$

Step (D6): generate the Illustration for the INRM. The total influence relation matrix **T** for the INRM is **calculated** using Equations (8) and (9), which, respectively, generate each row sum and column sum in the matrix **T**.

$$\mathbf{o} = (o_i)_{n\times 1} = \left[\sum_{j=1}^{n} t_{ij}\right]_{n\times 1} = (o_1, \ldots, o_i, \ldots, o_n). \tag{8}$$

$$\mathbf{r} = (r_i)_{n\times 1} = \left(r_j\right)'_{1\times n} = \left[\sum_{i=1}^{n} t_{ij}\right]'_{1\times n} = \left(r_1, ..., r_j, ..., r_n\right)'. \tag{9}$$

Step (D7): calculate the unweighted supermatrix $\mathbf{W}^{\alpha}$. Normalize the total influence relation matrix $\mathbf{T}_C$ using dimensions (called clusters), as shown in

$$\boldsymbol{T}_C^{\alpha} = \begin{array}{c} \\ D_1 \\ \vdots \\ D_i \\ \vdots \\ D_m \end{array} \begin{array}{c} \begin{array}{cccccc} & \boldsymbol{D_1} & & \boldsymbol{D_j} & & \boldsymbol{D_m} \\ & c_{11}\ldots c_{1m_1} & \cdots & c_{j1}\ldots c_{jm_j} & \cdots & c_{n1}\ldots c_{mm_m} \end{array} \\ \begin{array}{c} c_{11} \\ c_{12} \\ \vdots \\ c_{1m_1} \\ c_{i1} \\ c_{i2} \\ \vdots \\ c_{im_i} \\ c_{m1} \\ c_{m2} \\ \vdots \\ c_{mm_m} \end{array} \begin{bmatrix} \boldsymbol{T}_c^{\alpha 11} & \cdots & \boldsymbol{T}_c^{\alpha 1j} & \cdots & \boldsymbol{T}_c^{\alpha 1m} \\ \vdots & & \vdots & & \vdots \\ \boldsymbol{T}_c^{\alpha i1} & \cdots & \boldsymbol{T}_c^{\alpha ij} & \cdots & \boldsymbol{T}_c^{\alpha im} \\ \vdots & & \vdots & & \vdots \\ \boldsymbol{T}_c^{\alpha m1} & \cdots & \boldsymbol{T}_c^{\alpha mj} & \cdots & \boldsymbol{T}_c^{\alpha mm} \end{bmatrix}_{n\times n|m<n,\, \sum_{j=1}^{m} m_j = n} \end{array}, \tag{10}$$

where $\mathbf{T}_C^{\alpha}$ denotes the normalizing total influence relation matrix for criteria by dimensions, and $\mathbf{T}_c^{\alpha 14}$ is derived from Equations (11) and (12). Similarly, $\mathbf{T}_c^{\alpha mm}$ can be obtained:

$$t_i^{14} = \sum_{j=1}^{m_4} t_{ij}^{14}, \quad i = 1, 2, \cdots, m_1. \tag{11}$$

$$\mathbf{T}_c^{\alpha 14} = \begin{array}{c} \\ c_{11} \\ \vdots \\ c_{1i} \\ \vdots \\ c_{1m_1} \end{array} \begin{array}{c} \begin{array}{cccccc} & c_{41} & \cdots & c_4 j & \cdots & c_{4m_4} \end{array} \\ \begin{bmatrix} t_{11}^{14}/t_1^{14} & \cdots & t_{1j}^{14}/t_1^{14} & \cdots & t_{1m_4}^{14}/t_1^{14} \\ \vdots & & \vdots & & \vdots \\ t_{i1}^{14}/t_i^{14} & \cdots & t_{ij}^{14}/t_i^{14} & \cdots & t_{im_4}^{14}/t_i^{14} \\ \vdots & & \vdots & & \vdots \\ t_{m_1^1}^{14}/t_{m_1}^{14} & \cdots & t_{m_1 j}^{14}/t_{m_1}^{14} & \cdots & t_{m_1 m_4}^{14}/t_{m_1}^{14} \end{bmatrix} \end{array} = \begin{bmatrix} t_{11}^{\alpha 14} & \cdots & t_{1j}^{\alpha 14} & \cdots & t_{1m_4}^{\alpha 14} \\ \vdots & & \vdots & & \vdots \\ t_{i1}^{\alpha 14} & \cdots & t_{ij}^{\alpha 14} & \cdots & t_{im_4}^{\alpha 14} \\ \vdots & & \vdots & & \vdots \\ t_{m_1 1}^{\alpha 14} & \cdots & t_{m_1 j}^{\alpha 14} & \cdots & t_{m_1 m_4}^{\alpha 14} \end{bmatrix}. \tag{12}$$

Using pairwise comparisons of the criteria and the basic concept of ANP, the unweighted supermatrix $\mathbf{W}^{\alpha}$ is obtained by transposing the normalized influence relation matrix $\mathbf{T}_C^{\alpha}$ by dimensions (clusters); that is, $\mathbf{W}^{\alpha} = (\mathbf{T}_C^{\alpha})$, as shown in

$$\boldsymbol{W}^{\alpha} = (\boldsymbol{T}_C^{\alpha})' = \begin{array}{c} \\ D_1 \\ \vdots \\ D_j \\ \vdots \\ D_m \end{array} \begin{array}{c} \begin{array}{cccccc} & D_1 & & D_i & & D_m \\ & c_{11}\ldots c_{1m_1} & \cdots & c_{i1}\ldots c_{im_i} & \cdots & c_{m1}\ldots c_{mm_m} \end{array} \\ \begin{array}{c} c_{11} \\ c_{12} \\ \vdots \\ c_{1m_1} \\ c_{j1} \\ c_{j2} \\ \vdots \\ c_{jm_j} \\ c_{m1} \\ c_{m2} \\ \vdots \\ c_{mm_m} \end{array} \begin{bmatrix} \boldsymbol{W}^{11} & \cdots & \boldsymbol{W}^{i1} & \cdots & \boldsymbol{W}^{m1} \\ \vdots & & \vdots & & \vdots \\ \boldsymbol{W}^{1j} & \cdots & \boldsymbol{W}^{ij} & \cdots & \boldsymbol{W}^{mj} \\ \vdots & & \vdots & & \vdots \\ \boldsymbol{W}^{1m,} & \cdots & \boldsymbol{W}^{im} & \cdots & \boldsymbol{W}^{mm} \end{bmatrix}_{n\times n|m<n,\, \sum_{j=1}^{m} m_j = n} \end{array}. \tag{13}$$

Step (D8): calculate the weighted supermatrix. The normalized total influence-relation matrix of dimension $\mathbf{T}_D^{\alpha}$ is obtained by dividing the total influence-relation matrix $\mathbf{T}_D$ by $d_i = \sum_{j=1}^{m} t_{ij}$, $i = 1, 2, ..., m$, as shown in Equation (14):

$$\mathrm{T}_D^{\alpha} = \begin{bmatrix} t_{11}/\mathrm{d}_1/ & \cdots & t_{1j}/\mathrm{d}_1 & \cdots & t_{1m}/\mathrm{d}_1 \\ \vdots & & \vdots & & \vdots \\ t_{i1}/\mathrm{d}_i & \cdots & t_{ij}/\mathrm{d}_i & \cdots & t_{im}/\mathrm{d}_i \\ \vdots & & \vdots & & \vdots \\ t_{m1}/\mathrm{d}_m & \cdots & t_{mj}/\mathrm{d}_m & \cdots & t_{mm}/\mathrm{d}_m \end{bmatrix}_{m\times m} = \begin{bmatrix} t_{11}^{\alpha D} & \cdots & t_{1j}^{\alpha D} & \cdots & t_{1m}^{\alpha D} \\ \vdots & & \vdots & & \vdots \\ t_{i1}^{\alpha D} & \cdots & t_{ij}^{\alpha D} & \cdots & t_{im}^{\alpha D} \\ \vdots & & \vdots & & \vdots \\ t_{m1}^{\alpha D} & \cdots & t_{mj}^{\alpha D} & \cdots & t_{mm}^{\alpha D} \end{bmatrix}_{m\times m}. \tag{14}$$

The matrix $\mathbf{T}_D^{\alpha}$, the unweighted supermatrix $\mathbf{W}^{\alpha}$, and the weighted supermatrix $\mathbf{W}$ are obtained using Equation (15), where $t_{ij}^{\alpha D}$ is a scalar and $\sum_{j=1}^{m} m_j = n$:

$$\boldsymbol{T}_D^{\alpha} = \begin{bmatrix} t_{11}/d_1 & \cdots & t_{1j}/d_1 & \cdots & t_{1m}/d_1 \\ \vdots & & \vdots & & \vdots \\ t_{i1}/d_i & \cdots & t_{ij}/d_i & \cdots & t_{im}/d_i \\ \vdots & & \vdots & & \vdots \\ t_{m1}/d_m & \cdots & t_{mj}/d_m & \cdots & t_{mm}/d_m \end{bmatrix}_{m\times m} = \begin{bmatrix} t_{11}^{\alpha D} & \cdots & t_{1j}^{\alpha D} & \cdots & t_{1m}^{\alpha D} \\ \vdots & & \vdots & & \vdots \\ t_{i1}^{\alpha D} & \cdots & t_{ij}^{\alpha D} & \cdots & t_{im}^{\alpha D} \\ \vdots & & \vdots & & \vdots \\ t_{m1}^{\alpha D} & \cdots & t_{mj}^{\alpha D} & \cdots & t_{mm}^{\alpha D} \end{bmatrix}_{m\times m}. \tag{15}$$

Step (D9): limit the weighted supermatrix. Limit the weighted supermatrix by raising it to the zth power until the supermatrix converges and becomes a stable supermatrix. The global priority vectors and the global weight $\mathbf{w}^g$—which are called the IWs of DANP—are obtained, such as $\lim_{z\longrightarrow\infty}(\mathbf{W})^z$, **where** z represents any power. By summing the IWs for each criterion in every dimension, the local weight of dimension $\mathbf{w}_D^l$ is obtained. The global weight of each criterion is divided by the local weight of its own dimension to yield the local weight of criteria $\mathbf{w}_c^l$.

3.3. Technique for Order Preference by Similarity to an Ideal Solution (TOPSIS). TOPSIS is described in [37], with reference to [38]. TOPSIS is a multiple attribute method that is used to identify solutions from a finite set of alternatives. The chosen alternative must be closest to the positive ideal solution and farthest from the negative ideal solution. The procedure for TOPSIS involves a series of steps:

Step**(T1)**: calculate the normalize decision matrix. The normalized value n_{ij} is calculated using

$$n_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^{m} x_{ij}^2}}, \quad i = 1, ...m,\ j = 1, ...n. \tag{16}$$

Step (T2): calculate the weighted normalized decision matrix. The weighted normalized value v_{ij} is calculated as

$$v_{ij} = w_{ij} n_{ij}, \quad i = 1, ...m,\ j = 1, ...n. \tag{17}$$

where w_{ij} is the weight of the $i-$th attribute, and $\sum_{i=1}^{m} w_j = 1$.

Step (T3): determine the positive ideal and negative ideal solution:

$$\begin{aligned} A^+ &= \{v_1^+, ..., v_n^+\} = \left\{\left(\frac{\max_j v_{ij}}{i \in I}\right), \left(\frac{\min_j v_{ij}}{i \in J}\right)\right\}, \\ A^- &= \{v_1^-, ..., v_n^-\} = \left\{\left(\frac{\min_j v_{ij}}{i \in I}\right), \left(\frac{\max_j v_{ij}}{i \in J}\right)\right\}. \end{aligned} \tag{18}$$

where I is associated with a benefit attribute, and J is associated with a cost attribute.

Step (T4): calculate the separation from the positive ideal solution as

$$\mathrm{d}_i^+ = \left\{\sum_{j=1}^{n} \left(v_{ij} - v_j^+\right)^2\right\}^{1/2}, \quad i = 1, \ldots, m. \tag{19}$$

Similarly, the separation from the negative ideal solution is

$$\mathrm{d}_i^- = \left\{\sum_{j=1}^{n} \left(v_{ij} - v_j^-\right)^2\right\}^{1/2}, \quad i = 1, \ldots, m. \tag{20}$$

Step (T5): calculate the relative proximity to the ideal solution. The relative proximity of the alternative A_i with respect to A^+ is defined as

$$R_i = \frac{d_i^-}{\left(d_i^+ + d_i^-\right)}, \quad i = 1, \ldots, m. \tag{21}$$

Since $d_i^- \geq 0$ and $d_j^+ \geq 0$, then $R_i \in [0, 1]$.

Step (T6): rank the preference order for ranking alternatives using this index. Alternatives are ranked in decreasing order.

4. Results and Discussion

4.1. Environmental Design Elements to Control the Risk of Nosocomial Infection in Medical Buildings. This study focuses on Nosocomial Infection due to risk elements in the environment of medical buildings. The environmental design criteria for medical buildings that reduce the risk of Nosocomial Infection are determined using a review of relevant literature and an analysis of expert interviews. For this study, these guidelines are the prevention and control measures for Nosocomial Infection. The planning and design of a spatial environment can affect the behavior of individuals in medical buildings and the activity of infection sources, and reduce the risk of Nosocomial Infection in medical buildings.

4.1.1. Optimization of Sanitary Ware Layout and Design. Noncompliance with hand hygiene standards is one of the main high-risk behaviors that result in Nosocomial Infection. Many Nosocomial Infections are transmitted through the hands of medical staff [39], so hand hygiene is key to reducing the risk of Nosocomial Infection. Hand contact occurs during almost all medical processes. Practical experience shows that poor arrangement of sanitary ware, insufficient quantities, and poor location in medical buildings reduces compliance with hand hygiene protocols. Hugonnet et al. [40] noted that optimizing the layout of sanitary ware improves compliance with hand hygiene protocols. In a medical building, the number of sanitary stations must fully meet the needs of the individuals that use the building, so the layout of sanitary ware must match the distribution of people and entrances, and waiting areas and other places with a frequent flow of people should have sufficient cleaning stations. Health compliance is related to visual cues. Medical staff often ignore the position of sanitary ware at work so compliance with hand hygiene is increased by the use of induction lamps and significant visual warnings [41]. Noncontact technology also reduces the demand for hand hygiene. Inductive paper drawing, wash basins, and disinfectors reduce the probability of contact, and doors and windows, elevators and other frequently contacted parts can also use inductive technology.

4.1.2. Comfortable and Efficient Public Space. In medical buildings, patients must gather in public spaces for a long time, which increases the risk of contact in the hospital. The indoor environment of medical buildings has a high concentration of pathogens so long stays and frequent transfers significantly increase the risk of Nosocomial Infection. The crowd density in an enclosed space is positively correlated with the infection rate, and poor design can increase the time that patients remain in hospital [42, 43]. Insufficient waiting space, complex and tortuous streamline design, long distances between departments, and poor guide design increase the risk of infection. Azuma et al. [43] noted that in some narrow, crowded, and poorly ventilated indoor environments, aerosol transmission in close contact through some small, atomized particles is combined with respiratory droplets and contact transmission. Although there is little evidence of the impact of aerosol transmission, preventive measures are necessary to control the distribution of individuals in public spaces in medical buildings because COVID-19 has seriously affected global public health, the community, and the social economy.

4.1.3. Control the Crossing and Gathering of Crowd Movement Lines. The design for medical buildings must separate different types of traffic routes to control Nosocomial Infection [44]. The flow lines for common, susceptible, and high-risk groups must be distinguished in terms of the contact risk elements, and the range of movement within the hospital must be controlled [45]. The design for moving lines must use the path-finding characteristics of patients because difficulty in identifying a location is a common reason for unnecessary contact between patients. Specific measures include simplifying the paths, arranging rooms according to patients' path-finding habits, and reducing invalid space transfer [46]. Practical experience shows that a space can be classified according to the risk of Nosocomial Infection and cleanliness. There are ordinary areas, high-risk areas, and buffer areas. Significant buffer areas can be established in different cleanliness conversion areas and materials and colors can be used to emphasize the level of risk.

4.1.4. Appropriate Selection of Environmental Interface Materials. The interior of medical buildings is a closed space, so pathogens can survive in this internal environment for a long time [47]. Some frequently contacted environmental surfaces are significant vectors for Nosocomial Infection. The use of inappropriate environmental interface materials increases the risk of Nosocomial Infection. Carpet allows bacteria to breed and is difficult to clean, and the fabric is easily soaked and can become moldy [48, 49].

Pathogens also adhere to air particles and propagate for a long time, and volatilize into the air to form suspended pathogenic particles [50]. Therefore, environmentally friendly PVC is suited for wall decoration in hospitals. It has unique active antibacterial and self-cleaning characteristics. The surface coating is maintenance-free and surface dust can be directly cleaned using clean water. It has a minimal friction coefficient and experiences minimal dust fouling. Depending on the characteristics of hospital space, new

materials, such as bactericidal color latex paint, nano coatings, and antibacterial glass may be appropriate. The selection of materials must be informed by the rationality and practicability of the materials. Decorative materials in toilets must be waterproof and easily cleaned, operating rooms should use antibacterial glass partitions, and the corridors and wards in the hospital should be decorated with bactericidal color emulsion paint.

4.1.5. Higher Proportion of Single Compartments. Nursing units, operating rooms, and wards in medical buildings should achieve better space partition, and the proportion of single rooms should be increased to reduce the probability of contact transmission [51]. Practical experience shows that increasing the proportion of single compartments increases air quality (filtration, ventilation, and air flow control) in medical buildings. Previous studies have shown that under the same environmental conditions, the risk of Nosocomial Infection in a single room is significantly lower than that in a room with multiple occupants or an open ward [52], and the risk of water transmission in the toilet in a single room is also lower, especially in ICU and infection wards. Single room can also become isolation wards in case of an outbreak of an infectious disease. In practice, however, we are often faced with limited space, which can be tackled using the spatial design techniques to dismantle a cluster of units and increase the physical distance between them.

4.1.6. Correct Air Circulation and Purification. Hospital buildings are usually designed to have windows that face the sun and high ceilings because fresh air and good lighting reduce the risk of Nosocomial Infection. Previous studies show that a building's properties, especially the source of ventilated air and the airflow rate, are related to the diversity and composition of indoor bacterial communities [53]. Hobday and Dancer [54] noted that buildings are designed to increase exposure to outdoor air and sunshine to inhibit the survival and transmission of indoor infectious agents. However, many hospitals rely on mechanical ventilation, so air flow and filtering must be designed to prevent Nosocomial Infection. The concentration of pathogens in outdoor natural air is generally lower than that in indoor air, and natural air has a good microbial balance. So, many countries or regions require hospitals to use passive natural ventilation, rather than relying mainly on air conditioning or mechanical ventilation [54]. Previous studies show that opening windows in an ICU significantly reduces the probability of respiratory tract infection. The infection rate for SARS is related to the area of opening windows in a room [55]. Designers should actively use natural ventilation and mechanical ventilation to control indoor air flow and cleanliness, in order to reduce the risk of Nosocomial Infection.

4.1.7. Good Planning of Negative Pressure Isolation Areas. For severe respiratory infectious diseases, such as COVID-19 and SARS, negative pressure wards are used to control Nosocomial Infection [56]. Since the spread of COVID-19, many countries have noted a lack of vacuum isolation rooms, which increases the risk of Nosocomial Infection, aggravates patients' condition, and increases mortality [57]. The isolation area in a hospital has a ventilation system that features negative pressure to prevent infectious diseases from spreading to adjacent rooms or areas. Viruses are isolated and cross-infection is reduced [58]. The air distribution in a negative pressure isolation ward must control the direction of air flow in the ward such that the medical staff are on the windward side of the room and patients are on the leeward side.

4.1.8. Avoid the Accumulation of Water on the Surfaces of an Environment. Water is a breeding environment for the majority of pathogens, and water indoors can volatilize into the air to form pathogenic particles through contact with human skin and respiratory tract infection. Hospital Aspergillus can cause fatal infection in patients with low immune function, so Anaisie et al. [59] conducted a three-year study and determined that air and the hospital's water supply system are the main sources of hospital Aspergillus. Poor environmental design produces spaces that are difficult to clean, and if there is long-term water accumulation, the area can become a breeding ground for infection. These areas include water storage tanks, wash basins, shower heads, and decorative water systems [60].

In order to curb the breeding of pathogens, medical buildings should feature reduced water flow. Water vapor is also a potential source of infection and ventilation, and light controls the propagation of pathogens in water vapor. So, hospital buildings should use natural ventilation and light because ultraviolet light in sunlight inhibits mold and fungal pathogens.

4.2. Establish an Evaluation Framework for the Design of Medical Buildings' Spatial Environment to Curb the Risk of Nosocomial Infection. The general qualitative analysis of the relevant literature and expert interviews show that there are eight key environmental design elements. This study uses these environmental design elements to inhibit infection in medical buildings and uses a Likert Five-Point scale questionnaire. Subjects assigned a degree of approval for each item on the basis of actual experience in the prevention of Nosocomial Infection in medical buildings or the design of medical buildings' spatial environment. The scale has five levels: agree very much, agree, neither agree nor disagree, disagree, and disagree very much. The respective scores are 5, 4, 3, 2, and 1. A question that is not answered has a missing value. The subjects of the questionnaire included nurses in the Department of Nosocomial Infection, doctors, and researchers in the field of public health and who design medical buildings' spatial environment. All of the respondents had more than 3 years' work experience and 67 have had a master's degree or higher qualification. Sixty-one were female and 54 were male. A total of 115 questionnaires were distributed for this study and103 were valid.

This study uses EFA to determine the content validity and potential structure of environmental design elements, and then constructs an evaluation framework. Principal component analysis (PCA) and maximum variation (varimax) are used to select a characteristic value greater than 1, and items with elemental loads of less than 0.6 are deleted. Hundred three samples were suitable for element analysis (KMO = 0.729). In terms of reliability, the Cronbach α for the eight items is 0.718. The results show that two common elements have eigenvalues greater than 1 and the total cumulative explained variance is 82.431%. These eight evaluation criteria are summarized into two dimensions and the elemental load of each criterion is greater than 0.6. The two dimensions are guiding favorable behavior (D_1) and reducing the source of infection (D_2) (see Table 1). For the guidance of favorable behavior (D_1) dimension, there are four evaluation criteria: the optimization of sanitary ware layout and design (C_{11}); the control of the intersection and aggregation of crowd moving lines (C_{12}); a higher proportion of single compartments (C_{13}); and a comfortable and efficient public space (C_{14}). These feature good internal consistency (Cronbach $\alpha = 0.931$). There are also four evaluation criteria for reducing the source of infection (D_2): correct air circulation and purification (C_{21}); selection of appropriate environmental interface materials (C_{22}); avoiding water accumulation on environmental surfaces (C_{23}); and good planning of negative pressure isolation areas (C_{24}). These also feature good internal consistency (Cronbach $\alpha = 0.887$).

4.3. Determine the Influential Network Relationship Map (INRM) and Influence Weights (IWs) for the Evaluation Criteria. Exploratory element analysis gives a hierarchical evaluation framework with two dimensions and eight criteria. Weights are assigned to each criterion to allow interaction between the criteria in the same dimension. DANP is used to analyze the evaluation framework, draw the Influential Network Relationship Map (INRM) for each criterion in the same dimension, and assign the Influence Weights (IWs) for each criterion. The subjects for the DANP questionnaire were experts and scholars in the fields of nursing, public health, and medical architectural design. They all have many years of practical and scientific research experience (more than 5 years), and most have PHD degrees (79%). A total of 40 expert questionnaires were distributed, 34 valid questionnaires were returned, and the consistency of expert opinions was verified.

The results show that (as shown in Figure 2) in the D_1 dimension, the most influential criterion is a high proportion of single compartments (C_{13}), followed by the cross and aggregation of crowd moving lines (C_{12}), and the weakest criterion is the optimization of sanitary ware layout and design (C_{11}). In the D_2 dimension, the most influential criterion is the good planning of negative pressure isolation areas (C_{24}) and the others are: correct air circulation and purification (C_{21}), appropriate environmental interface materials (C_{22}), and avoiding water accumulation on environmental surfaces (C_{23}). Figure 2 shows that for the prevention of Nosocomial Infection in medical buildings, some evaluation criteria have strong Influence Weights (IWs), but their dominant influence on other criteria is weak. These include comfortable and efficient public space in D_1 (C_{14}) and avoiding water accumulation on environmental surfaces in D_2 (C_{23}).

4.4. Determine the Risk of Nosocomial Infection for Real Cases from the Perspective of Medical Buildings' Spatial Environment Design

4.4.1. Description of Real Cases. COVID-19 has affected many aspects of daily life. Different types of public buildings and open spaces in cities must undergo new inspection protocols to meet requirements for epidemic prevention and control. Medical buildings across China are facing unprecedented challenges. In China, high-quality medical resources are mostly concentrated in the first-tier cities, but even in these cities, there is still a gap between the overall shortage of medical services and the growing demand. The Pearl River Delta region has an advanced manufacturing base and a modern service industry base. The region is one of the main portals for China's foreign exchange and cooperation, and the cities with the highest annual net population inflow are in this region.

Three tertiary hospitals in the Pearl River Delta region are selected as actual cases, with case numbers SYUH, GZMH, and ZCWH. The first two hospitals are located in Guangzhou and the third is in Zhuhai. These three hospitals comprise multiple medical buildings with complex functions. They are also expanding on extremely limited land. They must prevent the current epidemic, but the hospitals are undergoing rapid development.

Case SYUH is in Tianhe District, Guangzhou. At present, the hospital area is roughly divided into three functional areas: the north side is the diagnosis and treatment area; the south side is the living area; and between the diagnosis and treatment area and the living area is the central green space of the hospital. In the medical buildings, the first floor to the fourth floor of the outpatient building, the medical technology complex building, the experimental building, and the inpatient building are connected by a corridor and the underground part of each building is also connected by an underground passage. The main entrance to the north side of the hospital is short of land, crowded, and mixed. The first-floor emergency hall is spacious and bright, and meets the spatial needs of emergency rescue. The emergency area is clearly divided. On weekdays, there are too many patients in the waiting space on each floor of the medical technology complex. There are several nursing units in the medical technology complex and the proportion of single rooms is low. Most wards in the inpatient building have windows that face south and the sanitary ware in the public space is insufficient and the design is old. The negative pressure isolation ward in the hospital was originally an ordinary ward and now uses a high-power exhaust fan for ventilation.

Case GZMH is in Changgang District, Guangzhou City, with a total floor area of 111000 square meters and a building

Table 1: The EFA results for environmental design elements.

($N = 103$) Evaluation criteria	Evaluation dimension	
	Guiding favorable behavior (D_1)	Reducing the source of infection (D_2)
The optimization of sanitary ware layout and design (C_{11})	0.947	
The control of the intersection and aggregation of crowd moving lines (C_{12})	0.920	
A higher proportion of single compartments (C_{13})	0.889	
A comfortable and efficient public space (C_{14})	0.880	
Correct air circulation and purification (C_{21})		0.954
Selection of appropriate environmental interface materials (C_{22})		0.926
Avoiding water accumulation on environmental surfaces (C_{23})		0.904
Good planning of negative pressure isolation areas (C_{24})		0.801
Cronbach α	0.931	0.887
Cumulative % of variance	82.431%	

area of 219300 square meters. The hospital area comprises an outpatient area to the north, an emergency area to the south, and an inpatient department in the middle. The fever clinic is located on the first floor of the administration building, adjacent to the inpatient building. The hall in the outpatient building is high, spacious, and bright, and waiting patients and accompanying people gather or stay for a long time in the public space. In contrast, the entrance hall of the emergency building is low and there is a clear moving line. The area of public toilets in the inpatient department and the catering department of the hospital district center is relatively common, and sanitary ware is insufficient. In the inpatient ward of the infectious diseases department, the proportion of single rooms is insufficient (generally 5 people), the indoor air humidity is high, and the ventilation depends on a high-power exhaust fan and natural ventilation.

As the third practical case for this study, ZCWH is in Xiangzhou District of Zhuhai City, adjacent to Gongbei port. The hospital covers an area of 42000 square meters and has a building area of 68000 square meters. The hospital area comprises an outpatient building in the east, a physical examination center building in the west, and an inpatient department in the middle. The emergency area is located on the first floor of the west side of the outpatient building and is connected to the inpatient department by a corridor. There are nursing units on each floor of the outpatient building, and it is not common for people to gather or stay in the waiting area. The first floor of the inpatient department is a registration hall with a high lift and payment hall, and the second floor and above house wards for different departments. Most wards have opening windows to the south. The proportion of single rooms is not high, but a large number of wards are double rooms with an independent bathroom and a public toilet. The negative pressure isolation area is divided based on the current situation of the hospital. The ventilation and purification of indoor air relies on natural ventilation, combined with air conditioning and an exhaust fan.

4.4.2. Evaluating the Risk of Nosocomial Infection for Real Cases. Based on the analysis results of the first two research stages, this study uses TOPSIS to determine the performance of these real cases in terms of building environment planning and design, and determines the risk of Nosocomial Infection for these real cases. A questionnaire (0–10 scale) was distributed to experts who have both architectural design and public health knowledge. An evaluation team of 15 experts from Guangdong, Hong Kong, Macau, and Taiwan was established. These individuals have a master's degree or a higher qualification, have accumulated more than 3 years of working experience, and have participated in scientific research projects that are related to medical architectural design engineering and public health management. The experts in the evaluation team are all familiar with the three real cases, and in the past year, most experts have conducted field surveys in these three hospitals. Some of the experts were hospitalized in the three hospitals and observed and recorded for a long time.

The performance evaluation results for these real cases (Table 2) show that from the perspective of guiding favorable behaviors to curb the risk of Nosocomial Infection, the risk of Nosocomial Infection in the three real cases is ranked from low to high as: ZCWH; GZMH; SYUH. From the perspective of reducing the source of infection to controlling the risk of Nosocomial Infection, the performance ranking for the three real cases is: ZCWH; SYUH; GZMH. The results in Figure 2 show that ZCWH, which has the highest comprehensive performance ranking, achieves high expert scores for the two criteria with the highest influence weight in the D_1 and D_2 dimensions. SYUH has the worst performance in guiding favorable behaviors to curb the risk of Nosocomial Infection. The hospital does not separate different types of traffic routes and indoor crowd crossing and unnecessary gathering is the most serious. Compared with ZCWH, SYUH's registration hall, outpatient waiting area, and other public spaces are not spacious and bright, and the distribution efficiency is low.

GZMH has the worst performance in reducing the source of infection to control the risk of Nosocomial Infection. There is a significant accumulation of water on the surfaces of the hospital environment, the air humidity in the wards is high, and the indoor air cleanliness is low.

4.4.3. Explore the Improvement Strategies for Building Environment Design for the Real Cases to Control the Risk of Nosocomial Infection. Based on the interaction between the

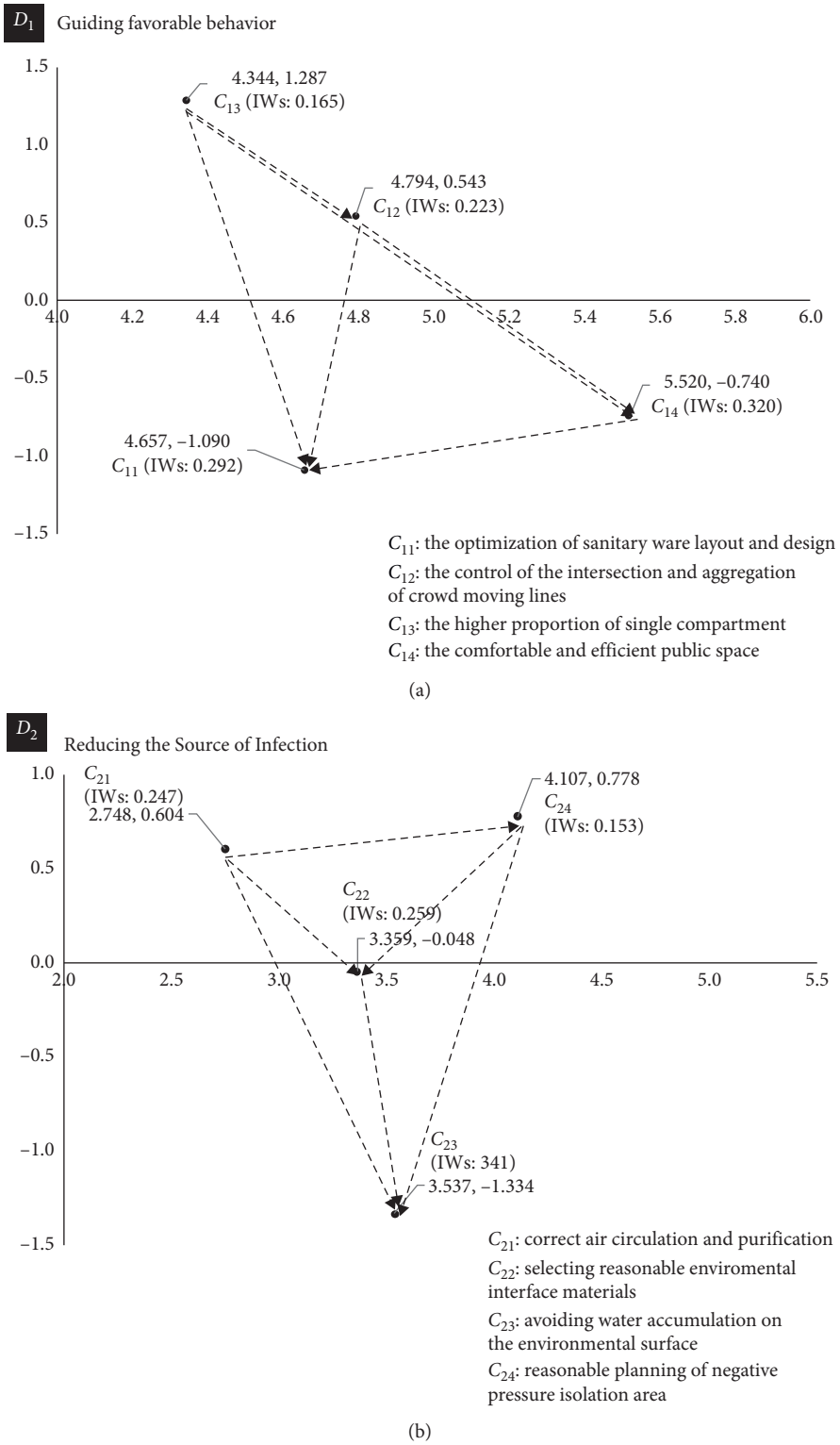


Figure 2: The Influential Network Relationship Map (INRM) for each evaluation criterion in the dimension.

evaluation criteria (Figure 2), combined with the performance evaluation results for real cases (Table 2), this study identifies the source of real problems for each case and develops continuous improvement strategies for each case. In order to promote and guide the behaviors that are conducive for restraining the risk of Nosocomial Infection in medical buildings, this study finds that both SYUH and GZMH must focus on existing resources to improve the proportion of single rooms in the hospital. If the space is deconstructed and reorganized, all types of unit compartments can be close to single rooms. It is common to use a flexible partition to split and combine horizontally. If the two cases can split the existing functional groups and design modular standard units, the proportion of single rooms will be increased and the nursing efficiency and adaptability of the building itself will undergo a beneficial functional transformation. If the unit compartments are transformed, the crowd movement line can be indirectly controlled and a one-way streamline and subcenter can be realized. Flexible partitions can be used to create a semi-enclosed group waiting mode, which controls the seat distance and personnel density and improves the comfort and distribution efficiency of the public space.

SYUH and GZMH have poor performance in terms of reducing the source of infection. According to the risk factors for airborne transmission, SYUH and GZMH must modify the design of the hospital air-conditioning and ventilation system to control and reduce the flow direction and concentration of airborne particles. Air must flow in an appropriate direction and from clean areas to semi-polluted areas to buffer area and then to polluted areas. In terms of the overall planning and design of the hospital, the areas where the most vulnerable patients are located (such as operating rooms, transplant facilities, and intensive care units) or the areas for patients with infectious diseases (such as infectious disease rooms or isolation wards), it is necessary to use a more powerful mechanical ventilation system to remove potential pathogenic biological aerosols, in order to reduce the risk of infection transmission.

ZCWH has the best performance of the three real cases, but Figure 2 and Table 2 show that the hospital could improve its performance in terms of then C_{13} and C_{12} criteria by optimizing hospital sanitary ware layout and design. This study finds that ZCWH should use data analysis for patients' path-finding characteristics to plan and design a moving line and simplify the path and reduce the unnecessary contact for patients in hospital. The layout for sanitary ware on the moving path of hierarchical planning avoids the waste of resources and allows a hierarchical design for sanitary ware and a guide system. The patients' routing habits mean that the layout of the unit compartments in the hospital can also be adjusted and invalid space transfer of patients in the hospital can be further improved, in order to improve the collection and distribution efficiency of the public spaces. The waiting area and entrance are the key elements for streamline control. It is difficult to identify high-risk groups, so the layout of the waiting area should allow a safe personnel density and control seat spacing. The fully open waiting area is not conducive for the control of Nosocomial Infection and the semi-enclosed group waiting mode can be used (as shown in Figure 3). Flexible partitions could be used to establish an isolation waiting area in pediatrics and obstetrics. When patients enter the hospital, their activities are difficult to control; so, it is necessary to increase control at the entrance and exit, to control the number of entrances and exits and to establish observation and registration facilities.

In terms of reducing the source of infection, ZCWH is significantly worse than the other cases in terms of the most influential C_{24} criterion. This hospital should transform some simple negative pressure isolation wards through safe and accessible technical means, in order to reduce the risk of cross-infection. In order to ensure a unidirectional flow of air from clean areas to polluted areas, the air pressure inside the ward must be lower than that outside the room. This is achieved by ensuring reasonable air distribution and keeping the door closed. Although space resources are extremely limited, the hospital should also use three areas and two channels to divide the polluted areas, semi-polluted areas, and clean areas, and clarify the medical channels and patient channels. In the negative pressure isolation ward, transfer windows with interlocking inner and outer windows and doors should be established on the wall of the potential pollution area in the adjacent corridor and the transfer window structure should be closed. The hospital should also plan a buffer area in the isolation area, increase disinfection and sterilization of the environment, and increase air purification. An ultraviolet sterilizer should be used as an auxiliary means to kill pathogens and a chilled beam can also be used as an indoor cooling source to separate the temperature regulation and ventilation functions, in order to reduce airborne infection.

Hospital designers can eliminate infection sources and restrict human behavior by planning the real space in the medical buildings to curb the risk of Nosocomial Infection. The public spaces in medical buildings, where people stay for a long time, are key areas for the prevention of Nosocomial Infection. This is a complex system of a variety of blocks with different functional attributes (outpatient department, inspection and treatment unit). By considering personnel density, residence time, physical environment data, and other relevant information, designers should establish a zoned layout and a design for supporting facilities at the site, to increase the distribution efficiency, safety, and comfort of public spaces. The results in Figure 2 show that designers should coordinate the planning and design of public spaces by considering unit layout and the organization of moving lines. A one-way streamline and a subcenter are ideal for medical buildings. Designers should make full use of space resources and maximize the proportion of single rooms. Flexible partitions can be used to divide polluted areas and semi-polluted areas, and to increase the size of buffer rooms and decontamination dressing rooms. The pedestrian flow line should be one-way and inflow and outflow should be separated. The direction in which doors open and the design of doors can be used to prevent the backflow of pedestrian flow from a polluted area to a clean area. Finally, designers should set corresponding functional attribute units in terms

Table 2: The performance evaluation for the case study using TOPSIS.

The evaluation framework		SYUH	GZMH	ZCWH	
Dimensions	Criteria	Influential weights	The normalize value n_{ij}		
Guiding favorable behavior (D_1)	The optimization of sanitary ware layout and design (C_{11})	0.292	0.551	0.689	0.470
	The control of the intersection and aggregation of crowd moving lines (C_{12})	0.223	0.407	0.598	0.690
	A higher proportion of single compartment (C_{13})	0.165	0.516	0.377	0.769
	A comfortable and efficient public space (C_{14})	0.320	0.420	0.596	0.685
Reducing the source of infection (D_2)	Correct air circulation and purification (C21)	0.247	0.623	0.430	0.653
	Selection of appropriate environmental interface materials (C22)	0.259	0.465	0.734	0.494
	Avoiding water accumulation on environmental surfaces (C23)	0.341	0.542	0.425	0.725
	Good planning of negative pressure isolation areas (C24)	0.153	0.710	0.522	0.473
Calculating the relative proximity to the ideal solution		R_i $(\boldsymbol{D_1})$	**0.214**	**0.565**	**0.660**
		R_i $(\boldsymbol{D_2})$	**0.433**	**0.369**	**0.619**

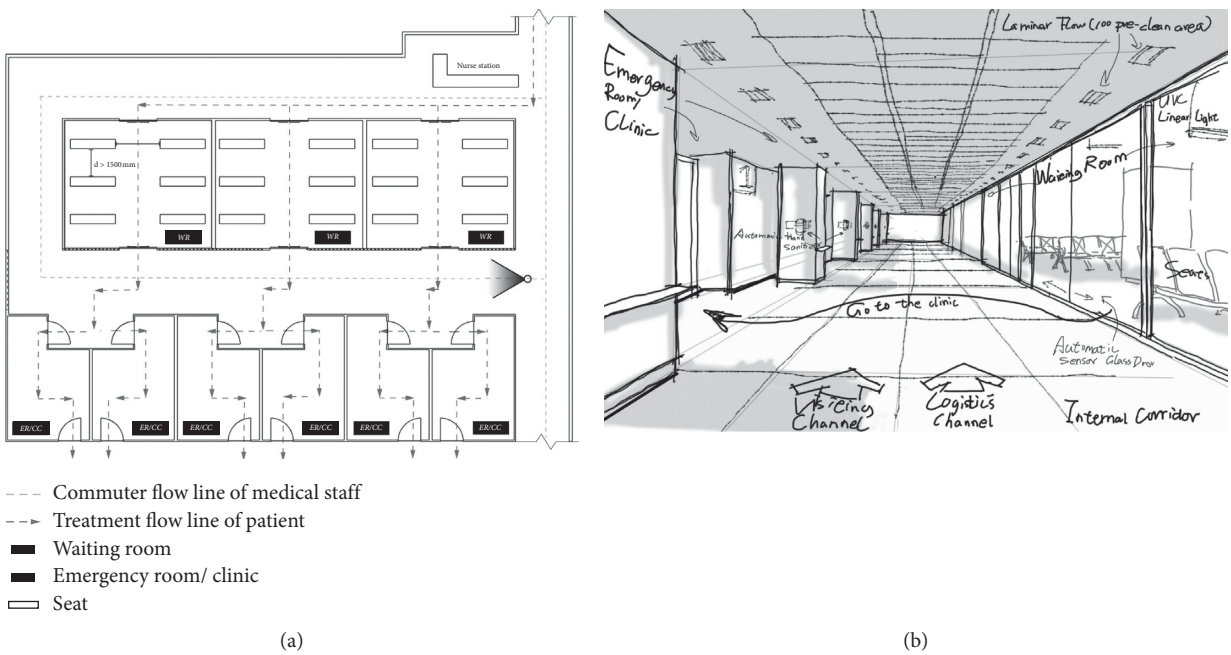


Figure 3: Example of outpatient department design.

of space division and organization. An independent physiological laboratory and waste treatment system are located in the infectious disease unit to reduce the risk of Nosocomial Infection during waste transfer for on-site pretreatment.

This study finds that reducing indoor water is the most important design criterion for reducing pathogen breeding and Nosocomial Infection risk. The antibacterial and cleaning properties of environmental interface materials are important. Materials that are easy to clean inhibit the sources of infection. Smooth and nonporous materials should be used for frequent contact elements and fabrics, and fiber materials should not be used. Floor materials should be seamless, and stainless steel skirting lines should be used in rooms in which there is a risk of blood and body fluid splashing, and internal corners should be avoided at junctions. The end of water flow should also be reduced. Figure 2 shows that designers should use simple room shapes, control ventilation and light to reduce pathogen reproduction in water vapor, give priority to natural ventilation and light, and use ultraviolet light to inhibit mold and fungal pathogens. In terms of the planning and design of a negative pressure isolation ward, designers should achieve a path for the inflow and outflow of air based on the zoning planning of the front room and ward of the negative pressure isolation ward, introduce sufficient fresh air from outside, and

maintain the air exchange rate. Particles to which pathogens attach for transport can be used as tracking materials and the particle removal efficiency can be used as a performance evaluation index for a negative pressure isolation ward to determine the effect of ward size, ventilation times, negative pressure value, and pathogen generation locations, in order to eliminate pathogens as quickly as possible.

5. Conclusions

Nosocomial Infection has a negative effect on social health, which directly and indirectly affects all of society. Hospitals are a gathering place for various infection sources, which leads to the risk of Nosocomial Infection. Nosocomial Infection is the early transmission mode for many epidemic infectious diseases, so its prevention and control affects the progress of an epidemic. Good Nosocomial Infection prevention and control mechanisms can prevent the further spread of an epidemic. Drug disinfection has been the main method to keep the hospital environment clean, so the source of infection has become increasingly resistant to drugs and increasingly difficult to control. These elements greatly increase the risk of Nosocomial Infection and the burden of epidemic prevention for the public health system. Nosocomial Infection due to COVID-19 means that hospitals must devise effective methods to prevent and control Nosocomial Infection.

Long-term effective reduction of Nosocomial Infection requires buildings with the correct spatial environment. Using a review of relevant literature, this study determines eight environmental design elements that reduce the risk of Nosocomial Infection: the optimization of sanitary ware layout and design (C_{11}); the control of the intersection and aggregation of crowd moving lines (C_{12}); a higher proportion of single compartments (C_{13}); a comfortable and efficient public space (C_{14}); correct air circulation and purification (C_{21}); selection of appropriate environmental interface materials (C_{22}); avoiding water accumulation on environmental surfaces (C_{23}); and good planning of negative pressure isolation areas (C_{24}). In order to transform these design elements into evaluation criteria and construct an evaluation framework, this study uses EFA to determine the effectiveness of the design elements and to clarify the categories (common elements) of the criteria. The evaluation framework that is constructed for this study has two dimensions: guiding favorable behavior (D_1) and reducing the source of infection (D_2). There are four evaluation criteria for each dimension. DANP is used to plot the Influential Network Relationship Map (INRM) and assign Influence Weights (IWs) for each criterion in the same category.

This study uses three general hospitals in the Pearl River Delta of China as real cases and evaluates the risk of Nosocomial Infection for each case from the perspective of the design of medical buildings' spatial environment. An analysis of the performance in these cases and the results in Figure 2 and Table 2 allow designers to determine the most efficient way to transform and update existing medical buildings or construct new medical buildings.

Using the results for real cases, this study embeds a system dynamic, which is an interactive point of view to determine the source of problems, which avoids the waste of resources, and develops a continuous and systematic design strategy for the real case. The main limitation of this study is that the use of EFA to select evaluation criteria and construct an evaluation framework means that the common factors (evaluation dimensions) that are obtained by analysis are independent of each other, so there is no interaction between D_1 and D_2. This study does not use DANP to clarify the INRM and IWs between the evaluation dimensions. The TOPSIS method that is used for this study to assess performance levels for the case studies is an additive method but circumstances are often nonadditive, so future research might use nonadditive methods to assess performance.

Data Availability

The data that are used to support the findings of this study are included in the article.

Conflicts of Interest

The authors declare no conflicts of interest.

Authors' Contributions

Lei Xiong and Ge Sheng contributed equally to this study.

Acknowledgments

The authors are grateful to all of the medical staff and designers who took part in this study. The authors would like to thank Hai-Miao Fan, the research assistant, for making the questionnaire survey a success. The authors are also immeasurably grateful to the alumni association of the Environmental Art Design, School of Architecture and Applied Art, Guangzhou Academy of Fine Arts, for their support during the research. This study was funded by the Graduate Education Innovation Project (GEIP) of Guangdong Province. The project number is 2019JGXM36. At the same time, this study was funded by the art projects supported by National Social Science Foundation, and the project number is 18BG08. This work was also supported by the Graduate Teaching Reform Project of Guangzhou Academy of Fine Arts (Grant number: 6040119012).

References

[1] J.-Y. Liu and J. K. Dickter, "Nosocomial infections," *Gastrointestinal Endoscopy Clinics of North America*, vol. 30, no. 4, pp. 637–652, 2020.

[2] X. Feng and L. Zhao, "Essential of hand hygiene to patients′ safety in hospital," *Chinese Journal of Nosocomiology*, vol. 12, 2006.

[3] N. F. Lenfestey, M. E. Denham, K. K. Hall, and D. B. Kamerow, "Expert opinions on the role of facility design in the acquisition and prevention of healthcare-associated infections," *HERD: Health Environments Research & Design Journal*, vol. 7, no. 1_suppl, pp. 31–45, 2013.

[4] R. J. Lubelchek and R. A. Weinstein, "Antibiotic resistance and nosocomial infections," *The Social Ecology of Infectious Diseases*, vol. 28, pp. 241–274, 2008.

[5] P. Spigaglia, "COVID-19 and Clostridioides difficile infection (CDI): possible implications for elderly patients," *Anaerobe*, vol. 64, Article ID 102, 2020.

[6] K. Yokoo, F. Sugaya, S. Matsuzaka et al., "The first case of COVID-19 occurring as community-acquired pneumonia in Hokkaido, Japan and our preventive measures against nosocomial infection," *Respiratory Medicine Case Reports*, vol. 30, Article ID 101078, 2020.

[7] M. Murphy and J. Mansfield, "Can architecture heal? Building as instruments of health," *Architectural Design*, vol. 87, no. 2, pp. 82–89, 2017.

[8] H. A. Gilbert, "Florence Nightingale's Environmental Theory and its influence on contemporary infection control," *Collegian*, vol. 27, no. 6, pp. 626–633, 2020.

[9] S. Rao, "Designing hospital for better infection control: an experience," *Medical Journal Armed Forces India*, vol. 60, no. 1, pp. 63–66, 2004.

[10] Y. U. H. G. O. Yamaguchi, "Better healing from better hospital design," *Harvard Business Review*, vol. 5, 2015.

[11] N. A. Megahed and E. M. Ghoneim, "Indoor Air Quality: rethinking rules of building design strategies in post-pandemic architecture," *Environmental Research*, vol. 193, Article ID 110471, 2020.

[12] H. Luo, J. Liu, C. Li, K. Chen, and M. Zhang, "Ultra-rapid delivery of specialty field hospitals to combat COVID-19: lessons learned from the Leishenshan Hospital project in Wuhan," *Automation in Construction*, vol. 119, Article ID 103345, 2020.

[13] K.-W. Wang, J. Gao, X.-X. Song et al., "Fangcang shelter hospitals are a One Health approach for responding to the COVID-19 outbreak in Wuhan, China," *One Health*, vol. 10, Article ID 100167, 2020.

[14] D. C. Anderson, T. Grey, S. Kennelly, and D. O'Neill, "Nursing home design and COVID-19: balancing infection control, quality of life, and resilience," *Journal of the American Medical Directors Association*, vol. 21, no. 11, pp. 1519–1524, 2020.

[15] S. Donetto, C. Penfold, J. Anderson, G. Robert, and J. Maben, "Nursing work and sensory experiences of hospital design: a before and after qualitative study following a move to all-single room inpatient accommodation," *Health & Place*, vol. 46, pp. 121–129, 2017.

[16] A. Sharifi and A. R. Khavarian-Garmsir, "The COVID-19 pandemic: impacts on cities and major lessons for urban planning, design, and management," *Science of The Total Environment*, vol. 749, Article ID 142391, 2020.

[17] C. Schröder, M. Behnke, C. Geffers, and P. Gastmeier, "Hospital ownership: a risk factor for nosocomial infection rates?" *Journal of Hospital Infection*, vol. 100, no. 1, pp. 76–82, 2018.

[18] G. H. Tzeng and K. Y. Shen, *New Concepts and Trends of Hybrid Multiple Criteria Decision Making*, CRC Press, Boco Raton, FL, USA, 2017.

[19] S.-K. Hu and G.-H. Tzeng, "A hybrid multiple-attribute decision-making model with modified PROMETHEE for identifying optimal performance-improvement strategies for sustainable development of a better life," *Social Indicators Research*, vol. 144, no. 3, pp. 1021–1053, 2019.

[20] R. A. Weinstein, "Epidemiology and control of nosocomial infections in adult intensive care units," *The American Journal of Medicine*, vol. 91, no. 3, pp. S179–S184, 1991.

[21] X. Wang, X. Jiang, Q. Huang et al., "Risk factors of SARS-CoV-2 infection in healthcare workers: a retrospective study of a nosocomial outbreak," *Sleep Medicine: X*, vol. 2, Article ID 100028, 2020.

[22] B. K. Decker and T. N. Palmore, "Hospital water and opportunities for infection prevention," *Current Infectious Disease Reports*, vol. 16, no. 10, p. 432, 2014.

[23] N. Zhang, W. Chen, P. T. Chan, H. L. Yen, J. W. T. Tang, and Y. Li, "Close contact behavior in indoor environment and transmission of respiratory infection," *Indoor Air*, vol. 30, no. 4, pp. 645–661, 2020.

[24] M. F. Yassin and S. Almouqatea, "Assessment of airborne bacteria and fungi in an indoor and outdoor environment," *International Journal of Environmental Science & Technology*, vol. 7, no. 3, pp. 535–544, 2010.

[25] A. C. Lowen, S. Mubareka, J. Steel, and P. Palese, "Influenza virus transmission is dependent on relative humidity and temperature," *PLoS Pathogens*, vol. 3, no. 10, p. e151, 2007.

[26] E. J. Anaissie, S. R. Penzak, and M. C. Dignani, "The hospital water supply as a source of nosocomial infections," *Archives of Internal Medicine*, vol. 162, no. 13, pp. 1483–1492, 2002.

[27] J. Rello, J. J. Rouby, C. Sole-Lleonart et al., "Key considerations on nebulization of antimicrobial agents to mechanically ventilated patients," *Clinical Microbiology and Infection*, vol. 23, no. 9, pp. 640–646, 2017.

[28] A. Rezaie, G. G. S. Leite, G. Y. Melmed et al., "Ultraviolet a light effectively reduces bacteria and viruses including coronavirus," *PLoS One*, vol. 15, no. 7, Article ID e0236199, 2020.

[29] K. Goswami, S. Bharali, and J. Hazarika, "Projections for COVID-19 pandemic in India and effect of temperature and humidity," *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, vol. 14, no. 5, pp. 801–805, 2020.

[30] Y. Diao, S. Kodera, D. Anzai, J. Gomez-Tames, E. A. Rashed, and A. Hirata, "Influence of population density, temperature, and absolute humidity on spread and decay durations of COVID-19: a comparative study of scenarios in China, England, Germany, and Japan," *One Health*, vol. 12, Article ID 100203, 2021.

[31] M. Guo, P. Xu, T. Xiao, R. He, M. Dai, and Y. Zhang, "Review and comparison of HVAC operation guidelines in different countries during the COVID-19 pandemic," *Building and Environment*, vol. 187, Article ID 107368, 2020.

[32] J. DeCoster, *Overview of Factor Analysis*, http://www.stathelp.com/notes.html, 1998.

[33] R. P. McDonald, *Factor Analysis and Related Methods*, Psychology Press, East Sussex, UK, 1985.

[34] L. R. Tucker and R. C. MacCallum, *Exploratory Factor Analysis*, Unpublished manuscript, Ohio State University, Columbus, OH, USA, 1997.

[35] H. F. Kaiser, "The varimax criterion for analytic rotation in factor analysis," *Psychometrika*, vol. 23, no. 3, pp. 187–200, 1958.

[36] C.-L. Lin and G.-H. Tzeng, "A value-created system of science (technology) park by using DEMATEL," *Expert Systems with Applications*, vol. 36, no. 6, pp. 9683–9697, 2009.

[37] S.-J. Chen and C.-L. Hwang, "Fuzzy multiple attribute decision making methods," *Lecture Notes in Economics and Mathematical Systems*, Springer, Berlin, Germany, pp. 289–486, 1992.

[38] C.-L. Hwang and K. Yoon, "Methods for multiple attribute decision making," in *Multiple Attribute Decision Making*, pp. 58–191, Springer, Berlin, Germany, 1981.

[39] S. O. Samuel, O. O. Kayode, O. I. Musa et al., "Nosocomial infections and the challenges of control in developing

countries," *African Journal of Clinical and Experimental Microbiology*, vol. 11, no. 2, 2010.

[40] S. Hugonnet, T. V. Perneger, and D. Pittet, "Alcohol-based handrub improves compliance with hand hygiene in intensive care units," *Archives of Internal Medicine*, vol. 162, no. 9, pp. 1037–1043, 2002.

[41] L. Diegel-Vacek and C. Ryan, "Promoting hand hygiene with a lighting prompt," *HERD: Health Environments Research & Design Journal*, vol. 10, no. 1, pp. 65–75, 2016.

[42] P. Sun, L. Zhao, M. Yu, and M. Zhang, "Prevention and control of the novel coronavirus in the stomatological hospital," *European Journal of Preventive Medicine*, vol. 8, no. 2, pp. 12–15, 2020.

[43] K. Azuma, U. Yanagi, N. Kagi, H. Kim, M. Ogata, and M. Hayashi, "Environmental factors involved in SARS-CoV-2 transmission: effect and role of indoor environmental quality in the strategy for COVID-19 infection control," *Environmental Health and Preventive Medicine*, vol. 25, no. 1, pp. 66–16, 2020.

[44] J. H. Bowie, R. W. Tonkin, J. S. Robson, and A. A. Dixon, "The control of hospital infection by design," *The Lancet*, vol. 284, no. 7374, pp. 1383–1387, 1964.

[45] H. K. Ki, S. K. Han, J. S. Son, and S. O. Park, "Risk of transmission via medical employees and importance of routine infection-prevention policy in a nosocomial outbreak of Middle East respiratory syndrome (MERS): a descriptive analysis from a tertiary care hospital in South Korea," *BMC Pulmonary Medicine*, vol. 19, no. 1, p. 190, 2019.

[46] S.-Y. Tzeng and J.-S. Huang, "Spatial forms and signage in wayfinding decision points for hospital outpatient services," *Journal of Asian Architecture and Building Engineering*, vol. 8, no. 2, pp. 453–460, 2009.

[47] M. Forrester, A. Pettitt, and G. Gibson, "Bayesian inference of hospital-acquired infectious diseases and control measures given imperfect surveillance data," *Biostatistics*, vol. 8, no. 2, pp. 383–401, 2007.

[48] D. D. Harris, A. Pacheco, and A. S. Lindner, "Detecting potential pathogens on hospital surfaces: an assessment of carpet tile flooring in the hospital patient environment," *Indoor and Built Environment*, vol. 19, no. 2, pp. 239–249, 2010.

[49] K. Diba, F. Jangi, K. Makhdoomi, N. Moshiri, and F. Mansouri, "Aspergillus diversity in the environments of nosocomial infection cases at a university hospital," *Journal of Medicine and Life*, vol. 12, no. 2, pp. 128–132, 2019.

[50] S. A. Fernando, T. J. Gray, and T. Gottlieb, "Healthcare-acquired infections: prevention strategies," *Internal Medicine Journal*, vol. 47, no. 12, pp. 1341–1351, 2017.

[51] T. Ueno and N. Masuda, "Controlling nosocomial infection based on structure of hospital social networks," *Journal of Theoretical Biology*, vol. 254, no. 3, pp. 655–666, 2008.

[52] A. P. R. Wilson and G. L. Ridgway, "Reducing hospital-acquired infection by design: the new University College London Hospital," *Journal of Hospital Infection*, vol. 62, no. 3, pp. 264–269, 2006.

[53] S. W. Kembel, E. Jones, J. Kline et al., "Architectural design influences the diversity and structure of the built environment microbiome," *The ISME Journal*, vol. 6, no. 8, pp. 1469–1479, 2012.

[54] R. A. Hobday and S. J. Dancer, "Roles of sunlight and natural ventilation for controlling infection: historical and current perspectives," *Journal of Hospital Infection*, vol. 84, no. 4, pp. 271–282, 2013.

[55] A. R. Escombe, C. C. Oeser, R. H. Gilman et al., "Natural ventilation for the prevention of airborne contagion," *PLoS Medicine*, vol. 4, no. 2, p. e68, 2007.

[56] H. Kim, D. Kim, C. Paul, and C. K. Lee, "The spatial allocation of hospitals with negative pressure isolation rooms in Korea: are we prepared for new outbreaks?" *International Journal of Health Policy and Management*, vol. 9, no. 11, p. 475, 2020.

[57] S.-W. Jeon and L. A. Williams, "The way to manage a lack of NPIR for COVID-19 patients," *Frontiers in Health Informatics*, vol. 9, no. 1, p. 43, 2020.

[58] M.-Y. Yen, Y. E. Lin, I.-J. Su et al., "Using an integrated infection control strategy during outbreak control to minimize nosocomial infection of severe acute respiratory syndrome among healthcare workers," *Journal of Hospital Infection*, vol. 62, no. 2, pp. 195–199, 2006.

[59] E. J. Anaissie, S. L. Stratton, M. C. Dignani et al., "Pathogenic Aspergillus species recovered from a hospital water system: a 3-year prospective study," *Clinical Infectious Diseases*, vol. 34, no. 6, pp. 780–789, 2002.

[60] K. T. Van, *Enhancing Appropriate Environmental Design in Healthcare Facilities for the Inhibition of Nosocomial Infection*EDP Sciences, Les Ulis, France, 2019.

Hindawi
Journal of Healthcare Engineering
Volume 2021, Article ID 5547749, 20 pages
https://doi.org/10.1155/2021/5547749

Research Article

Exploring Strategies for Improving Green Open Spaces in Old Downtown Residential Communities from the Perspective of Public Health to Enhance the Health and Well-Being of the Aged

Si-Jie Li,[1,2] Yu-Feng Luo,[1,2] Zi-Chuan Liu,[2] Lei Xiong,[3] and Bo-Wei Zhu[1]

[1]*Faculty of Humanities and Arts, Macau University of Science and Technology, Avenida Wai Long, Taipa 999078, Macau, China*
[2]*Academy of Fine Arts, South China Normal University, Guangzhou 510631, China*
[3]*School of Architecture and Allied Art, Guangzhou Academy of Fine Arts, Guangzhou 511400, China*

Correspondence should be addressed to Lei Xiong; kmt20005@gmail.com and Bo-Wei Zhu; zhubowei301@gmail.com

Received 29 January 2021; Revised 9 April 2021; Accepted 27 April 2021; Published 14 June 2021

Academic Editor: Chia-Yu Lin

Based on the trend of global aging, people are paying more and more attention to the health of the elderly and the improvement of green open spaces. However, few studies have focused on strategies to improve green spaces in response to this trend. Especially, with the outbreak of COVID-19, an urgent need to develop a sustainable system strategy to improve the health of the elderly in residential communities in old districts has emerged. Traditional improvement strategies based on current situation evaluation often focus on the most prominent practical problems. Therefore, the objective of this study was to provide theoretical research and practical improvement strategies for green open spaces in old downtown residential communities to improve the health and well-being of the elderly. In response to this problem, this research proposes an alternative method based on causality (FDM-DANP-mV model), by extracting 23 green open space elements that affect the health of the elderly and dividing them into three dimensions, to form a preliminary evaluation framework. On this basis, the more effective and feasible standard elements are screened out, and the influence relationship behind the elements is clarified. Then, the sustainable development strategy is systematically discussed in three practical cases. This allows for the analysis of the present situation to not only identify the current significant problems but also to capture the source of the influence behind the real problems based on the clarification of the dominant influence relationship. The actual value of this study is to provide a key design decision basis for the improvement of the green open spaces in old downtown residential communities, aiming at avoiding waste to the greatest extent under the premise of limited resources and gradually promoting the improvement of the urban built environment to promote the health and well-being of the elderly.

1. Introduction

With the aging of the global population, the World Health Organization (WHO) has put forward a policy of "healthy aging" [1], with "providing the aged with safe, inclusive and barrier-free public spaces" [2] as the goal of global aging health strategy and action. And, China has become an important representative of healthy aging [3]. When taking public interest into account, incorporating the growing health needs of the aged into the development of public open spaces (POS) has become a major challenge of the 21st century [4]. In addition, as said by the WHO, whether long life will become a social burden depends on the health of the aged to a large extent [2]. At present, the sudden outbreak and wide coverage of COVID-19 brings new challenges to urban planning and architectural design [5]. How to reduce the dissemination of the virus and restore economic growth through optimal urban planning and architectural design has become an urgent topic in the post-epidemic age [6]. Having low autoimmunity and mostly suffering from underlying diseases, the aged are susceptible to the virus and once infected are difficult to cure. As reported, most of the infected and critically ill patients are vulnerable elderly people [7].

As the center of early commercial and cultural activities in many developing countries, most old urban areas have superior geographical location [8], still retain the imprint of the times, and show unique characteristics of the times. Because these communities have been around for long time, a large number of elderly people have long lived in these areas, and the phenomenon of community aging has become serious. Generally, the green space in these communities is small and disordered, chaotic, and the infrastructure is old and in need of maintenance. Thus, the quality of public open space is low, and the social function has been neglected. As such, these spaces cannot meet the demands of the elderly for green open space [9]. The outbreak of the novel coronavirus also had an indirect impact. One survey showed that China's overall emotional health decreased by 74% [10]. Moreover, the majority of elderly people have long been at high risk of experiencing mental health problems [11]. Therefore, people have begun to pay attention to health problems and maintain health by participating in activities [12]. Ekstrom [13] pointed out that the uncertainty of the new environment has caused pressure on the elderly, and resettlement will have a negative impact on their mental health, resulting in increased pressure and loss of social support. Urban renewal design aims to solve the problem of unplanned urban aging and promote the sustainable development of cities, while helping to preserve the identity and sense of belonging of the elderly to the original community, emphasize the inclusiveness of the society, and promote the social participation and active aging of the elderly [14]. As a result, providing a green landscape space for the elderly to enhance health and well-being should not be done simply by reprogramming and updating the full picture [15]. At the same time, the outbreak of the epidemic poses a new challenge to urban and architectural design, which is an important part of the national economy, [16]. Epidemics are more likely to erupt in urban areas, arousing our reflection and new exploration in the direction of epidemic prevention through urban construction [17].

In a previous study, Owen Douglas et al. [18] proposed to sort out and review the relationship between health, happiness, and green open space using the life cycle method. In the whole life cycle, they suggested, the key points of planning green open space that is conducive to enhancing the health and well-being of the aged are as follows: first, providing less-intense sports facilities to encourage social interaction and interaction with nature; second, incorporating spontaneous and leisurely opportunities for interaction with the environment into the design of green open space; third, providing sheltered seating areas, sources of drinking water, and toilets. Taking Iran as an example, Azadeh Lak et al. [19], clarified the preferences of the aged for active aging when using public places, and proposed that social environment, sense of belonging, cultural background, and ultimate life satisfaction were the top-priority factors. Perception and preference of the aged to urban space can be independently understood from each of these factors. Furthermore, the design strategies that can promote the health of the aged include: (1) functional quality (such as safety, aesthetics, convenience, cleanliness, comfort, density and urban landscape, and the putting quality of the place first); (2) preferred quality (such as wayfinding ability, subjective aesthetics, safety and control of the fear of falling, which will make older people prefer open spaces); and (3) an environment that has an active and significant impact on preferences of the aged, especially including the social environment (social interaction and civic participation), cultural environment (cultural and religious beliefs), sense of belonging, and life satisfaction in different contexts. With respect to the disorder of green spaces in old urban areas, Chen et al. [20] put forward the following improvement strategies for the purpose of increasing the humanistic and ecological benefits of street landscapes: (1) strengthening the standardized management of municipal sanitation and engineering construction and paying attention to garbage removal, greening, maintenance, and overall environmental governance in corners; (2) attaching importance to the overall ecology and landscape design of urban areas, providing tree ponds, rain gardens, and other sponge facilities as per the local conditions and taking into full consideration the integration of planting and maintenance; and (3) dynamic monitoring of the environmental greening facilities in appropriate combination with digital sensor techniques to guide smart street management. Meanwhile, street furniture and small works of art can be combined to achieve higher humanistic and ecological benefits and increase the interest and artistry of urban space.

To sum up, it is believed that based on the premise of limited resources, it is necessary to focus on improving the current situations of sites and to carry out adaptive adjustments. Obviously, this needs to base on a systematic evaluation of the current situation of a site. And, the examination of real problems and the formulation of applicable improvement strategies need to be completed in combination with a dynamic influence and complementary viewpoints. This helps to formulate a systematic improvement strategy starting from the root cause based on the actual situation of the site [21–23]. Therefore, it is urgently necessary to construct an evaluation framework for evaluating the current situations of green open spaces in downtown residential communities, and clarify the dominant influence relation between the criteria based on criterion weight training. This is the basis for making subsequent evaluation and analysis of the performance in real cases. The results of performance evaluation should reveal the gap between empirical cases and the desired level. Furthermore, a systematic continuous improvement strategy should be formulated for real cases based on the above evaluation and analysis.

The purpose of this research was to extract the construction criteria of green open spaces in old downtown residential communities based on construction of the health and well-being of the aged under the influence of the outbreak, make clear the priorities of the criteria and the dominant influence relation between the criteria, further evaluate the performance of empirical cases on this basis, and explore sustainable improvement strategies for different empirical cases. Therefore, the process of this research is as follows: first, we extract the criteria relevant to improving

green open spaces in old downtown residential communities for enhancing the health and well-being of the aged under the influence of the outbreak and initially establish an index framework system on this basis; second, we screen out the evaluation criteria and judge the stability and effectiveness of the evaluation framework using the Fuzzy Delphi Method (FDM); further, we identify the weight of each index by applying the DANP technique while clarifying the complicated and key dominant influence relations between the criteria; finally, we evaluate and analyse the performance of the empirical cases by using a modified VIKOR technique, making clear the gap between each empirical case and the desired level and further making strategies for improving green open spaces in old downtown residential communities to enhance the health and well-being of the aged. This allows for the proposed situation evaluation and analysis to not only identify the current significant problems but also to capture the source of the influence behind the real problems based on the clarification of the dominant influence relationships, and avoid waste to the greatest extent under the premise of limited resources. Therefore, the practical value of this study is to provide a direction for the transformation of residential areas in the centers of old cities, and to guide researchers and decision-makers in theory and practice on how to enhance the health and well-being of the elderly by improving the real root causes.

The structure of the rest of this paper is as follows: In Section 2, relevant elements are summarized and the research evaluation framework is initially constructed through a review of key previous studies of outdoor activities, connecting with nature, and social interaction; in Section 3, the methods and steps used in this research are introduced in detail; in Section 4, FDM is used to screen the effectiveness of the criteria, DANP is used to clarify the influence relation between and weights of the evaluation criteria, and the evaluation criteria are discussed with regard to three real cases in combination with performance evaluation and improvement based on VIKOR; Section 5 discusses the contributions and limitations of this research and suggests future research directions.

2. Establishment of an Evaluation Framework

The WHO defines healthy aging as: developing and maintaining the physical functions of the health of the aged, where physical functions depend on one's endogenous capacity (i.e., the combination of physical and mental abilities) and living environment (including physical, social, and policy environments) and interactions [1]. Based on this, scholars have analyzed the impact mechanism of urban public open spaces on the health of the aged, and have generally summarized several influential factors of open space within three dimensions, namely, outdoor activities, biological affinity, and social support and community place creation [14]. Therefore, in this work, the factors affecting the health and well-being of the aged in green open spaces in old downtown residential communities are summed up through the review of key studies of "outdoor activities," "connecting with nature," and "social interaction," and an evaluation framework is proposed based on this review (as shown in Table 1).

2.1. Outdoor Activities. Nature-based outdoor activities play an important role in improving the well-being of the aged [65]. As indicated by many scholars, friendly slow walking paths such as curved ring-type paths can expose the aged to green open spaces for longer periods of time, providing them with more options for outdoor activities, including slow walking, jogging, and brisk walking [24]. Some studies have emphasized that outdoor activities of the aged may be related to safety, walkability, park access, natural and aesthetically pleasing scenery, and recreational facilities [25]. Walkable infrastructures [19], wide barrier-free sidewalks [31], and accessible public transportation are the basic structures required for the aged to develop outdoor activities in green open spaces. A Hong Kong study reported that leisurely walking of local elderly people is positively correlated with connectivity [26]. In addition, the aged prefer to enjoy green open spaces that offer convenient facilities such as services and public toilets [18], as well as well-designed landscapes including comfortable seats for sitting and resting, plants for shade, adequate lighting, and sports related equipment [30]. Physical activity is an important and decisive factor of the health of the aged [34]. Through sports activities, it is possible to extend the time the aged stay in a green open space, further enhancing health and well-being [35]. During the outbreak of COVID-19, appropriate control measures should be taken in response to the fast dissemination and high incidence rate [39]. Studies have shown that the aged can obtain effective and full service information relevant to the surrounding environment from readable signage systems such as air quality monitors, landmarks, or hazard indicators in green open spaces [25]. Meanwhile, some artificial measures such as outdoor ventilators and purifiers can be set up to create a healthy environment. Moreover, separating water supply systems from sewage systems can prevent and delay infection [37]. As confirmed by Thompson and Aspinall [66], outdoor leisure activities can bring benefits to the aged, including physiological benefits for maintaining and enhancing their physical health. All the benefits of nature-based outdoor activities have significant impacts on the well-being of the aged [30].

2.2. Touching the Nature (Biological Affinity). Thompson [66] and Jessica Finlay et al. [44] indicated that connecting with nature is essential for physical and mental health. As jointly shown in several studies, incorporating appropriate features of natural water [37], such as lakes, ponds, and fountains, superior greening vegetation [44], flowers and plants with fragrant smells [42], beautiful outdoor landscapes [30] ,such as small artworks, outdoor furniture, and greening into green open space, can provide the aged with sufficient auditory relief, olfactory treatment, and human immune functions and further improve their active emotions [43]. According to the human-oriented principle, the health and well-being of the aged can be promoted by incorporating biological affinity into green open spaces and

Table 1: Index factors for pretest.

Dimensions	Criteria	Description	Reference source
Outdoor activities (D_1)	Friendly slow walking (C_1)	Providing continuous circular or curved walking paths for the aged to do more outdoor activities such as slow walking, jogging, and brisk walking; eliminating any form of slope or puddle on the walking path, and constructing walk-friendly infrastructure to get rid of the fear of falling of the aged; providing barrier-free sidewalks with sufficient width to meet the needs of special elderly people.	[19, 24, 25]
	Public transport connectivity and accessibility (C_2)	Ensuring the public traffic routes adjacent to the green open space are clear and well-connected so that it is easy to find and enter the space; and routes should be within suitable walking distance from the space.	[26–28]
	Readable guidance systems and hazard indicators (C_3)	Setting up information exchange boards, air-quality test displays, landmark or hazard indicators to provide the aged with effective and adequate service information about the surrounding environment.	[25, 29, 30]
	Suitable lighting (C_4)	Choosing applicable lamps, setting the brightness and installing the lamps in appropriate positions to ensure suitable brightness transition and uniformity, and clear vision, so as to enhance safety in night activities.	[19, 25]
	Adequate seating (C_5)	Providing comfortable seats with backrests and armrests for the aged to stay and rest and communicate with each other (the seats should be installed as per the overall green open space design, facing wide views; the backs should be shaded by tree shadows. The seats can be small benches with low soft walls or sheltered landscape structures such as gallery frames).	[30–33]
	Convenient infrastructure (C_6)	Providing drinking water sources, public toilets and the like provisions for convenience and environmental cleanliness, which are considered as the most important needs of the aged.	[18,21]
	Sports entertainment facilities (C_7)	Recreational facilities can be used, such as outdoor sports equipment, golf courses, or bicycle lanes. The aged prefer recreational facilities with guidance and emergency braking systems to avoid injury.	[24, 34, 35]
	Separation of rainwater and sewage (C_8)	Historically, constructive measures and standards for preventing and delaying epidemics mainly include urban and rural water supplies and sewage systems (separation of rainwater and sewage).	[36–38]
	Outdoor ventilator/air cleaning device (C_9)	COVID-19 can spread through droplets under windless conditions; hence, it is necessary to take some manual measures to enhance ventilation efficiency, quickly remove air pollutants, and further prevent the spread of the virus.	[39–41]

Table 1: Continued.

Dimensions	Criteria	Description	Reference source
Touching the nature (D_2)	Features of natural water (C_{10})	Appropriate features of natural water such as lakes, ponds, and fountains should be provided, as water elements in nature can lead to higher therapeutic responses and preferences and improve active emotions in the aged.	[19, 33]
	High-quality plants (C_{11})	Vegetation planted in suitable positions and suitable greening allocation throughout the four seasons can offer sufficient visual relief effects through adjustment of natural colors and the sun. Planting flowers and plants with fragrant smells is conducive to improving the olfactory and immune functions of the aged. To avoid pollen allergies, it is necessary to choose flower varieties with less pollen.	[42–44]
	Interactive leisure landscape (C_{12})	Incorporating opportunities for incidental and leisure interaction with the environment into the design of green open spaces, such as providing operable landscape installations to encourage interaction.	[18]
	Suitable soundscape (C_{13})	Providing pleasant natural sounds such as birdsongs (trees suitable for perching birds can be planted) and the sounds of water and wind (waterscapes or microclimates can be created by virtue of altitude difference) to stimulate auditory sense and promote faster physiological recovery.	[45–48]
	Access to fruits and vegetables (C_{14})	Making available fruit and vegetable plants though gardening and DIY so that the aged can obtain tactile feedback from real plant life in nature, taste fresh fruits and vegetables, and have more positive feelings.	[49, 50]
	Beautiful outdoor landscape (C_{15})	Providing beautiful outdoor landscapes in harmony with the natural environment, such as the combination of greening and street furniture and small artworks to achieve humanistic ecological benefits and increase the interest and artistry of the landscapes.	[19, 51]
	No noise pollution (C_{16})	The aged in high-density urban areas prefer quiet green open spaces. No noise is one of the most important factors to ensure smooth communication, especially for the aged with poor hearing. The increase in noise may negatively affect the communication frequency of the aged.	[52–55]
Social Interaction (D_3)	Security management (C_{17})	Including daily patrols for security, installation of monitoring TV, and maintenance of infrastructures such as changing guardrails along waterfronts and replacing old exercise equipment to avoid crime and vandalism.	[32, 56, 57]
	Gathering social square (C_{18})	Forming a large square space for the aged to carry out various entertainment activities, and adding wide spaces for promoting their participation in and supporting community communication.	[[25, 30, 33]
	Open multi-functional spaces (C_{19})	The wider the adjacent street is, the stronger the sense of openness is, and the lower the surrounding buildings are, the stronger the sense of openness is. This can help increase the frequency of social interaction.	[40,58]
	Optimization of children's playground (C_{20})	The children's play area should have good visibility and a space suitable for the aged to rest. This can drive inter-generational exchanges.	[59]
	Pet activity space (C_{21})	Giving full consideration to pet activity space to increase the outdoor activities of the aged and drive the aged to take walks with their pets in the space to enhance their physical and mental health; this has been found to be a way to lower the blood pressure of the aged and help strengthen community exchange and interaction.	[60–62]
	Preservation of community culture (C_{22})	The full application of elements with cultural and historical information of the original community to the design of green open spaces can enhance local identity and sense of belonging.	[25, 63, 64]
	Public health service station (C_{23})	This kind of station functions to quickly and conveniently deal with the emergency health conditions of the aged while disseminating daily disease prevention knowledge, and facilitates good connectivity and accessibility. Necessary space for emergency shelters should be established to deal with sudden disasters, and obvious route maps should be posted.	[6, 37]

giving the aged a sensory experience through interaction with nature [42]. Ratcliffe et al. [46] and Alvarsson et al. [45] stated that pleasant natural sounds such as the sounds of water and wind and chirping of bird songs promote auditory stimulation of the aged; by gardening and planting, the aged not only obtain fruits and vegetables but also experience nature in man-made environments and obtain tactile feedback from real plant life [49]. Douglas et al. [18] mentioned that incorporating leisure and interactive landscapes such as operable landscape installations into the design of green open spaces provides opportunities for the aged to interact with each other, relieves pressure, and produces more active feelings. According to the concept of biological affinity, connecting with nature plays a fundamental role in maintaining the physical and mental health of humans, showing the inner link between humans and nature [67]. In addition, for the aged with weak hearing, no noise is one of the most important factors to ensure smooth communication. Noise increase has a negative impact on communication frequency of the aged, and greatly reduces their expectations from green open space [26].

2.3. Social Interaction (Social Support and Community Place Creation). It has been determined that social interaction, social environment, and social relation (especially in social environments of neighborhood parks) are significantly related with the health of the aged [68]. They are potentially crucial mechanisms for public open space to intervene in the relationship between the environment and the social health of the aged [14]. One of the functions of green open space is to create an outdoor environment for the aged to relax and socialize, and improve their health, well-being, and quality of life [58]. As emphasized in some studies, if the aged can live in more age-appropriate communities with multifunctional open spaces supporting interaction, it is helpful to promote the positive social interaction of the aged [9], reduce depression and anxiety, and improve their quality of life [30]. Studies have shown that in planning, attention should be paid to preserving the elements and cultural heritage of the original community, which is conducive to increasing the sense of belonging and identity [19]. Wong and Yu [64] found that the aged tend to sit down when engaging in social activities. Therefore, providing enough friendly seats and multifunctional spaces and other built-in environmental functions can enable the aged to improve their well-being and sense of self-control through contact and conversation, allowing them more opportunities to develop their social lives, and further establishing strong social connection [25]. Chen [30] pointed out that different levels of social interaction can be achieved through various activities in public open spaces such as dancing, walking, and exercising. Social activities on streets and in community squares are one of the main reasons for the aged to enter public open spaces as they need space to walk and do other sports activities with friends living nearby. Considering that public square space should be available to do interactive activities, it is necessary to make appropriate planning [30]. Social activities of the aged also include keeping pets, which can help release stress in life, reduce depression, and bring a sense of security and love [61]. In particular, walking with pets in public open spaces can help enhance the physical and mental health and social interaction of the aged [69]. Hence, consideration can be taken to set up space for pet activity when designing green open spaces. At the same time, the aged like to visit strongly inclusive places (such as parks or neighborhood open spaces) to watch children playing [59]. Thus, planners should provide the aged with resting areas and wide visibility of children's facilities. This can play a positive role in encouraging the aged to visit parks and such planning allows children and their grandparents to feel safe and enjoy the nature-based entertainment.

3. FDANP-mV Model

The idea of this research is shown in Figure 1. This figure presents the research methods and technical routes applied in different stages of the research and the corresponding research problems. Firstly, an evaluation framework is initially established through a literature review; then, pretest is made using the Fuzzy Delphi Method (FDM) to judge the accuracy of the structure of the framework and the effectiveness of each evaluation criterion; further, the final evaluation framework is constructed. Subsequently, the DANP technique is applied to clarify the influence of relations between various evaluation criteria based on expert knowledge, draw an Influential Network Relation Map (INRM), and allocate the Influential Weights (IWs) of the criteria. In the phase of real case evaluation and analysis, performance evaluation and analyses are made of three old communities in central Guangzhou. Based on the performance evaluation questionnaire, a modified VIKOR technique is applied in this study to determine the gap value between each of the three cases and the total performance value. Above all, analysis results of the research phases are taken as the basis to explore a sustainable and systematic design strategy for improving public open spaces in old downtown neighborhood communities for the purpose of enhancing the health and well-being of the aged.

3.1. Fuzzy Delphi Method. Proposed by Dalkey and Helmer [70] in 1960, the Delphi method is a procedural method for systematically expressing the opinions of expert groups. The fuzzy set was incorporated in the Delphi method by Murray, Pipino, and Gigch [71] for the first time. Ishikawa et al. [72] integrated the opinions of experts into fuzzy numbers based on the concepts of cumulative frequency distribution and fuzzy integrals, which is called the Fuzzy Delphi Method (FDM). This method can be used as a tool for screening evaluation criteria. Compared with traditional Delphi method, FDM has the following advantages: (1) the frequency of investigation is reduced; (2) it can completely express the opinions of experts; (3) expert knowledge will be more in line with rationality and practical needs by applying a fuzzy set; and (4) it is more economical in terms of time and cost.

Generally, FDM can be used to carry out the following three main steps: (1) establish a set of evaluation factors that affect decision-making; (2) collect opinions of experts or

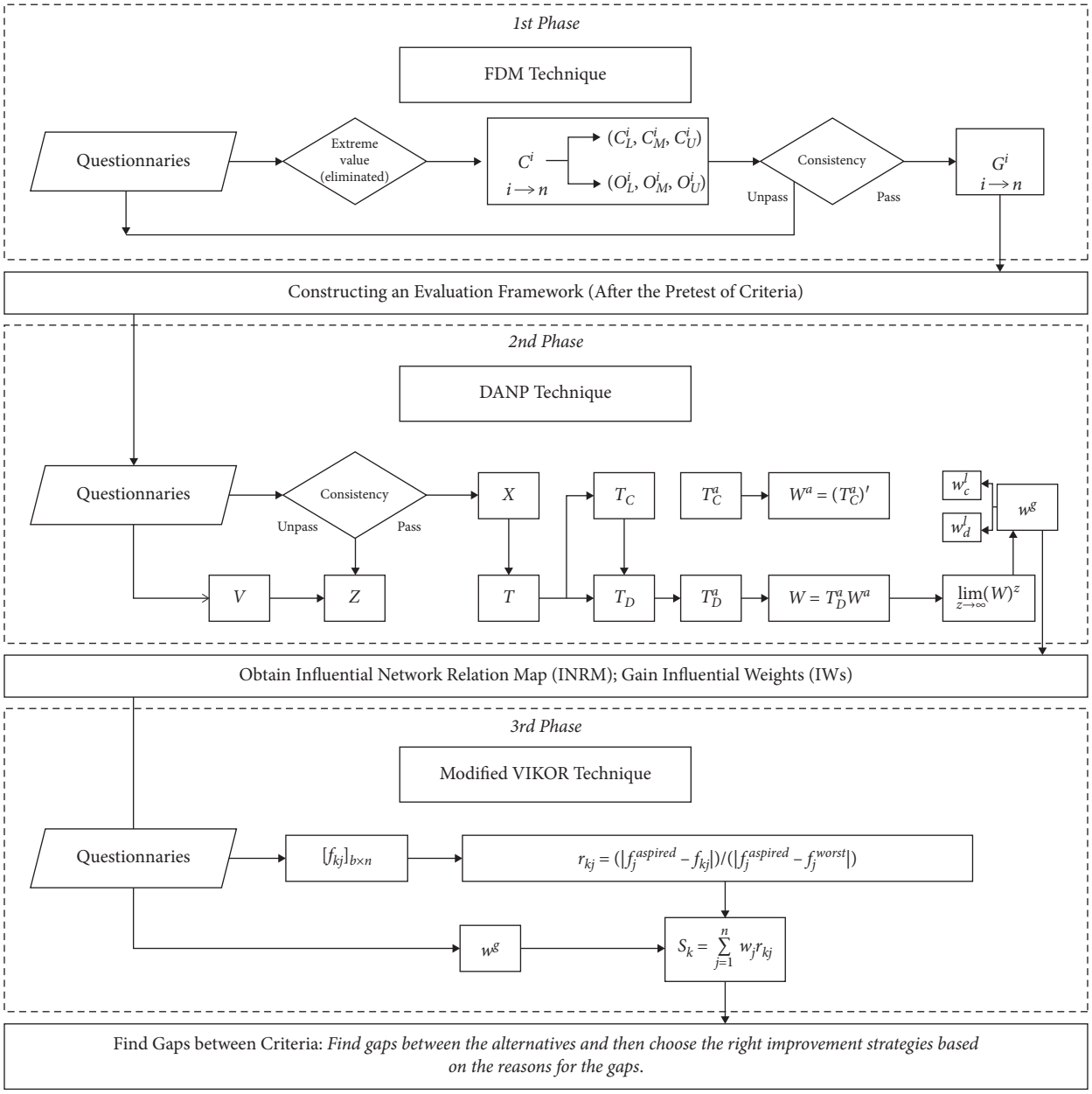


Figure 1: Modeling procedure of FDANP-mV.

decision-making groups; and (3) calculate the evaluation value in the FDM method. Zheng Cangbin [73] integrated expert opinions on the basis of a "double-triangular fuzzy number," and the "grey zone test method" is more effective for testing whether expert opinions are consistently convergent (that is, reaching a consensus). Therefore, in this study, FDM was used to screen out important evaluation criteria for public open spaces in neighborhood communities to improve the health and well-being of the aged, and then to construct an evaluation framework to lay a foundation for subsequent evaluation and analysis. The application and steps of the test method are described in detail as follows:

Step 1. Count the "conservative value $\boldsymbol{C}_i$," the "optimistic value $\boldsymbol{O}_i$," and the "single value $\boldsymbol{A}_i$."

Count the interval value given by the expert for each evaluation criterion in all expert questionnaires. The "minimum value" among the interval values represents the "conservative value $\boldsymbol{C}_i$" given by the expert to the quantified score of the evaluation criterion, while the "maximum value" among the interval values represents the "optimistic value $\boldsymbol{O}_i$" given by the expert to the quantified score of the evaluation criterion. The single value refers to the "subjective cognitive value" given by the expert to the quantified score of the evaluation criterion.

Step 2. Take out extreme values that fall outside the "double standard deviation."

Count the minimum value, maximum value, and the single values given by the experts for each evaluation

criterion, and take out the extreme values that fall outside the "double standard deviation." Then, calculate the remaining minimum value, geometric mean, and maximum value in "conservative value $\boldsymbol{C}_{\mathrm{i}}$," as well as their counterparts in "optimistic value $\boldsymbol{O}$," which are not taken out.

Step 3. Check whether the expert opinions are in agreement with each other based on the "grey zone verification method."

Establish the triangular fuzzy number (C_L^i, C_M^i, C_U^i) of conservative value $\boldsymbol{C}_{\mathrm{i}}$" and that of "optimistic value $\boldsymbol{O}_{\mathrm{i}}$" (O_L^i, O_M^i, O_U^i) calculated in Step Two. As per the calculation result shown in Table 2, the interval between the value of C_U^i and that of O_L^i refers to the verification value of $\boldsymbol{M}_{\mathrm{i}}$—$\boldsymbol{Z}_{\mathrm{i}}$ for the grey zone. Where $M_i = (O_M^i - C_M^i)$, $Z_i = (C_U^i - O_L^i)$. In the case that the value of C_U^i is larger than O_L^i, and $\boldsymbol{M}_{\mathrm{i}}$—$\boldsymbol{Z}_{\mathrm{i}}$ is bigger than 0, it indicates that the expert opinions tend to be identical, and convergence is reached for the evaluation criterion; when $\boldsymbol{M}_{\mathrm{i}}$—$\boldsymbol{Z}_{\mathrm{i}}$ is negative, it means that the expert opinions are not in agreement with each other, and no convergence is reached for the evaluation criterion. At this point, the second questionnaire can be used for the evaluation criterion for which convergence has not yet been reached.

Step 4. Calculate the consensus value of the expert $\boldsymbol{G}_{\mathrm{i}}$, and identify the reasonable threshold.

The intersection node of the two lines of the C_U^i value and O_L^i value refers to the consensus value of the expert G^i, which presents an important basis for the identification of the threshold. In the meantime, the higher the value of G^i, the higher the consensus value of the expert, or, in other words, the higher the importance of the evaluation criterion. It can be seen from Table 2 that in this study, the threshold is set at 6.667, referring to the arithmetic mean of G^i, and is based on the semantic scale of a Likert questionnaire. Therefore, each criterion for which G^i is lower than the threshold of 6.667 is taken out.

3.2. Research Procedures of the DANP Technique. In 1996, Thomas L. Saaty established the Analytic Network Process (ANP) with attribute dependency and feedback relation based on the issue of improving the Analytic Hierarchy Process, AHP, thereby bringing the weight of attributes closer to the real-world condition. However, in ANP, it is assumed that the attribute weight value of each facet is equal. Though this simplifies calculation, the fact that the importance of the attribute is not equal to that in the real world is neglected. As a result, in 2008, based on the Markov Chain Process (the degree of importance is characterized by Transferability), Gwo-Hshiung Tzeng stressed the degree of importance for the calculation of attributes through the ANP mode according to the total influence relation matrix $\boldsymbol{T}$ established by DEMATEL as the influence weight. Through the empirical analysis, Tzeng showed that this method better conforms with the conditions of real society.

Step 5. Create direct relation matrix $\boldsymbol{Y}$.

The experts transform the experience of real society into the degree value of the interaction of the attributes within the model structure based on five point Likert scales (0 for no influence, 1 for very low influence, 2 for low influence, 3 for high influence, and 4 for very high influence) and establish the receptance matrix $\boldsymbol{Y}$ for the direct relation matrix with nonnegative value of $n \times n$, where y_{ij} represents the degree of the impact of the i-th criterion on the j-th criterion.

Step 6. Create initial direct relation matrix $\boldsymbol{Z}$.

Use equations (1) and (3) to integrate the direct relation matrix of $\boldsymbol{Y}$ for p experts, thereby obtaining initial direct influence relation matrix $\boldsymbol{Z}$ representing the experience of p experts, where p refers to the total number of experts and y_{ij} represents the degree of the impact of the i-th criterion on the j-th criterion.

$$Z_{ij} = \frac{1}{p} \sum_{k=1}^{p} y_{ij}^k, \tag{1}$$

$$Z = \begin{bmatrix} z_{11} & \cdots & z_{1j} & \cdots & z_{1n} \\ \vdots & & \vdots & & \vdots \\ z_{i1} & \cdots & z_{ij} & \cdots & z_{in} \\ \vdots & & \vdots & & \vdots \\ z_{n1} & \cdots & z_{nj} & \cdots & z_{mn} \end{bmatrix}. \tag{2}$$

Step 7. Examine the consensus.

The value of consensus can be estimated by equation (3), which represents the level of expert consensus. The threshold of the average gap ratio is 5%, and a value less than 5% implies a confidence level above 95%, which also represents a stable system. Conversely, if an unstable system is obtained, the first phase should be implemented again to verify whether data collection is correct and whether the number of experts is sufficient.

$$\text{average gap} - \text{ratio in consensus } (\%) = \frac{1}{n(n-1)} \sum_{i=1}^{n} \sum_{j=1}^{n} \left(\frac{\left| z_{ij}^p - z_{ij}^{p-1} \right|}{z_{ij}^p} \right) \times 100\%. \tag{3}$$

Step 8. Normalize direct influence matrix $\boldsymbol{X}$.

The initialization of direct influence relation matrix Z can build the boundary of its influence relation matrix through equations (4) and (5). All the influence values of its influence relation matrix are within the range from 0–1; after that, the standardized direct relation matrix $\boldsymbol{X}$ has a

Table 2: Fuzzy Delphi statistical analysis results.

Criteria	Conservative Value Ci		Optimistic Value Oi		Single Value Ai		Geometric Mean M			Verification values	Consensus values
	Min	Max	Min	Max	Min	Max	Ci	Oi	Ai	Mi—Zi	Gi
Friendly slow walking (C_1)	5	8.5	8	10	6.5	9	6.621	8.811	7.586	1.690	8.151
Public transport connectivity and accessibility (C_2)	6	8.2	8	10	7	9	6.891	9.005	7.869	1.914	8.087
Readable guidance systems and hazard indicators (C_3)	4	8	6	9	5	8	5.525	7.813	6.609	0.288	6.846
Suitable lighting (C_4)	5	9	7.5	10	6	9	6.357	8.644	7.449	0.787	7.953
Adequate seating (C_5)	5	9	7	10	6	10	6.447	8.528	7.335	0.081	7.749
Convenient infrastructure (C_6)	5	8	7	10	6	9	6.149	8.594	7.148	1.445	7.463
Sports entertainment facilities (C_7)	5	9	7	10	6	9	6.578	8.655	7.643	0.077	7.812
Separation of rainwater and sewage (C_8)	6	8	7	10	6	10	6.835	8.806	7.696	0.971	7.608
Outdoor ventilator/air cleaning device (C_9)	5.3	8	7	10	6	9	6.425	8.566	7.469	1.141	7.499
Natural water characteristics (C_{10})	5.5	8	7.5	10	6	9	6.978	9.277	8.091	1.799	7.817
High-quality plants (C_{11})	6	8	8	10	7	9	6.915	9.366	8.097	2.452	8.000
Interactive leisure landscape (C_{12})	5	7.5	7	10	6	9	6.652	8.950	7.796	1.797	7.348
Suitable soundscape (C_{13})	5	7	6	10	6	9	5.996	8.050	7.062	1.054	7.344
Access to fruits and vegetables (C_{14})	4	7	6	10	5	9	5.643	7.803	6.697	1.160	6.571
Beautiful outdoor landscape (C_{15})	5	8.6	8	10	7	9	7.011	9.364	8.094	1.753	8.277
No noise pollution (C_{16})	5	9	8	10	6	9	6.556	8.895	7.551	1.339	8.268
Security management (C_{17})	6	9	8	10	7	9	7.026	9.382	7.975	1.356	8.412
Gathering social square (C_{18})	4	8	7	9.6	5	8	5.611	7.769	6.517	1.158	7.243
Open multifunctional spaces (C_{19})	6	9	8	10	7	9	6.826	9.004	7.859	1.178	8.316
Optimization of children's playground (C_{20})	5	8	7	10	6	9	5.822	7.844	6.767	1.021	7.279
Pet activity space (C_{21})	5	7	6	10	5	8	5.595	7.613	6.621	1.019	6.534
Preservation of community culture (C_{22})	5	9	7	10	6	9	6.954	8.981	7.884	0.026	7.984
Public health service station (C_{23})	6	9	8.7	10	7	9	7.373	9.638	8.308	1.965	8.810
Number of evaluation elements selected				21						Threshold value	6.667

minimum row-column total value of 0, and a maximum value of 1.

$$X = m \cdot Z, \tag{4}$$

$$m = \min\left\{\frac{1}{\max_{1\le i\le n}\sum_{j=1}^{n} z_{ij}}, \frac{1}{\max_{1\le j\le n}\sum_{i=1}^{n} z_{ij}}\right\}. \tag{5}$$

Step 9. Obtain total influence relation matrix $\boldsymbol{T}$.

The standardized direct relation matrix $\boldsymbol{X}$ is able to calculate the value of multiple influences and the indirect influence of the criterion, i.e., the total influence value of the criterion, through equation (6). Therefore, when $\mathbf{o} \longrightarrow \infty$, that is, $\mathbf{N}^o = [0]_{n\times n}$, the total influence relation matrix $\boldsymbol{T}$ is obtained, where $\boldsymbol{I}$ refers to the unit matrix.

$$\begin{aligned} T &= X + X^2 + X^3 + \cdots + X^u, whereo \longrightarrow \infty, X^o = [0]_{n\times n}, \\ &= X\left(I\text{—}\mathrm{X}^{\mathrm{k}-1}\right)(I\text{—X})^{-1}, \\ &= X(I\text{—X})^{-1}. \end{aligned} \tag{6}$$

Step 10. Establish the network relationship map (NRM).

At last, the total value of attributed influence degree r_i is obtained by summing up the attribute column of the obtained total influence relation matrix $\boldsymbol{T}$. The summing up of the row refers to the total value of the degree of the influenced attribute d_i, while $r_i + d_i$ refers to the correlation degree of the attribute, which is called the prominence; r_i -d_i denotes the degree of significance for the influence nature of the attribute, which is called the relation. In addition, a method for quantifying the system structure relation is available for the efficient simplification of the degree of complexity of inter-criterion correlation, thereby making the overall perspective of the entire evaluation system easy to understand and utilize.

Step 11. Establish the unweighted supermatrix $\mathbf{W}^a$.

Normalize the total influence relation matrix $\mathbf{T}_C$ by the dimensions (called "clusters") as shown in the following equation:

$$T_c^{\alpha} = \begin{array}{c} \\ D_1 \\ \\ D_i \\ \\ D_m \end{array} \begin{array}{c} \begin{array}{cccccc} D_1 & & D_j & & D_m \\ c_{11\mathrm{L}}\, c_{1m_1} & \cdots & c_{j1\mathrm{L}}\, c_{jm_j} & \mathrm{L} & c_{n1\mathrm{L}}\, c_{mm_m} \end{array} \\ \begin{bmatrix} T_c^{\alpha 11} & \mathrm{L} & T_c^{\alpha 1j} & \mathrm{L} & T_c^{\alpha 1m} \\ \mathrm{M} & & \mathrm{M} & & \mathrm{M} \\ T_c^{\alpha i1} & \mathrm{L} & T_c^{\alpha ij} & \mathrm{L} & T_c^{\alpha im} \\ \mathrm{M} & & \mathrm{M} & & \mathrm{M} \\ T_c^{\alpha m1} & \mathrm{L} & T_c^{\alpha mj} & \mathrm{L} & T_c^{\alpha mm} \end{bmatrix}_{n\times n} \end{array} m < n, \sum_{j=1}^{m} m_j = n, \tag{7}$$

where $\mathbf{T}_C^{\alpha}$ denotes the normalizing total influence relation matrix of criteria by dimensions, and $T_c^{\alpha 14}$ is derived from equations (8) and (9). Similarly, $T_c^{\alpha mm}$ can be obtained as follows:

$$t_i^{14} = \sum_{j=1}^{m_4} t_{ij}^{14}, \quad i = 1, 2, \ldots, m_1, \tag{8}$$

$$T_c^{\alpha 14} = \begin{array}{c} \\ c_{11} \\ \mathrm{M} \\ c_{1i} \\ \mathrm{M} \\ c_{1m_1} \end{array} \begin{array}{c} \begin{array}{ccccc} c_{41} & \mathrm{L} & c_{4j} & \mathrm{L} & c_{4m_4} \end{array} \\ \begin{bmatrix} t_{11}^{14}/t_1^{14} & \mathrm{L} & t_{1j}^{14}/t_1^{14} & \mathrm{L} & t_{1m_4}^{14}/t_1^{14} \\ \mathrm{M} & & \mathrm{M} & & \mathrm{M} \\ t_{i1}^{14}/t_i^{14} & \mathrm{L} & t_{ij}^{14}/t_i^{14} & \mathrm{L} & t_{im_4}^{14}/t_i^{14} \\ \mathrm{M} & & \mathrm{M} & & \mathrm{M} \\ t_{m1}^{14}/t_{m_1}^{14} & \mathrm{L} & t_{m_1j}^{14}/t_{m_1}^{14} & \mathrm{L} & t_{m_1m_4}^{14}/t_{m_1}^{14} \end{bmatrix} \end{array} = \begin{bmatrix} t_{11}^{\alpha 14} & \mathrm{L} & t_{1j}^{\alpha 14} & \mathrm{L} & t_{1m_4}^{\alpha 14} \\ \mathrm{M} & & \mathrm{M} & & \mathrm{M} \\ t_{i1}^{\alpha 14} & \mathrm{L} & t_{ij}^{\alpha 14} & \mathrm{L} & t_{im_4}^{\alpha 14} \\ \mathrm{M} & & \mathrm{M} & & \mathrm{M} \\ t_{m_1 1}^{\alpha 14} & \mathrm{L} & t_{m_1 j}^{\alpha 14} & \mathrm{L} & t_{m_1 m_4}^{\alpha 14} \end{bmatrix}. \tag{9}$$

First, a normalization calculation was performed on the total criterion influence relationships according to different dimensions. Since the direction of pairwise comparison between the criteria is opposite to the ANP concept, the total-influenced relationship matrix needs to be transposed. The total-influenced relationship matrix is obtained from the DEMATEL, and each column is summed up for normalization. After normalization, an unweighted super matrix is established.

Step 12. Establish the weighted supermatrix $\mathbf{W}_w$.

The weighted super matrix $\mathbf{W}_w$ can be obtained by multiplying the criterion of $\mathbf{W}^a$ in the unweighted supper matrix and the facet of $\mathbf{T}_D^{\alpha}$, as shown in equation (10).

$$W_w = T_D^{\alpha} W^{\alpha} = \begin{array}{c} \\ D_1 \\ \\ \\ D_i \\ \\ \\ D_m \\ \\ \end{array} \begin{array}{c} \begin{array}{cccc} D_1 & & D_i & D_m \end{array} \\ \begin{bmatrix} c_{11} \; c_{12} \; \mathrm{M} \; c_{1m_1} & \cdots & c_{11\mathrm{L}} \; c_{im_i} & \mathrm{L} & c_{m1\mathrm{L}} \; c_{mm_m} \\ t_{11}^{\alpha D} \times W^{11} & \mathrm{L} & t_{i1}^{\alpha D} \times W^{i1} & \mathrm{L} & t_{m1}^{\alpha D} \times W^{m1} \\ \mathrm{M} & & \mathrm{M} & & \mathrm{M} \\ t_{1j}^{\alpha D} \times W^{1j} & \mathrm{L} & t_{ij}^{\alpha D} \times W^{ij} & \mathrm{L} & t_{mj}^{\alpha D} \times W^{mj} \\ \mathrm{M} & & \mathrm{M} & & \mathrm{M} \\ t_{1m}^{\alpha D} \times W^{1m} & \mathrm{L} & t_{im}^{\alpha D} \times W^{im} & \mathrm{L} & t_{mm}^{\alpha D} \times W^{mm} \end{bmatrix} \end{array}. \tag{10}$$

Step 13. Calculate the limiting supermatrix $\mathbf{W}^{\mathrm{g}}$.

Convergence and stability can eventually be achieved through self-multiplication of the weighted supper matrix $\mathbf{W}_w$, or, in other words, by obtaining the maximized super matrix $\mathbf{W}^{\mathrm{g}}$ where $\mathbf{W}_w$ refers to the weighted super matrix and l represents the number of self-multiplications.

3.3. Research Procedures of the Modified VIKOR Technique. VIKOR (VlseKriterijumska Optimizacija I Kompromisno Resenje) is a multi-criteria, decision-making method that was proposed by Professor Serafim Opricovic and Professor Gwo-Hshiung Tzeng in 1998. VIKOR is one of the optimal compromise solutions of multi-criteria decision-making methods, the basic idea of which is to first define the positive-ideal solution (optimal solution) and the negative-ideal solution (the worst solution).

Step 14. The modified approach for replacement by the aspiration level and the worst value is as follows.

The aspiration level: $f^{\text{aspired}} = (f_1^{\text{aspired}}, \ldots, f_j^{\text{aspired}}, \ldots, f_n^{\text{aspired}})$, where f_j^{aspired} is an aspiration level, which is also called the best value.

The worst values: $\mathbf{f}^{\text{worst}} = (f_1^{\text{worst}}, \ldots, f_j^{\text{worst}}, \ldots, f_n^{\text{worst}})$, where f_j^{worst} is a worst value.

In this study, performance scores ranging from 1 to 5 (very bad $\longleftarrow$ 1, 2, ..., 5 $\longrightarrow$ very good) are used with natural language in the linguistic/semantic questionnaire. Thus, the aspiration level takes the highest score of 5 and the worst value takes the value of 1. Hence, $f_j^{\text{aspired}} = 5$ is defined as the aspiration level and $f_j^{\text{worst}} = 1$ is defined as the worst value.

Step 15. Determine the mean group utility for the gap and then establish the priority improvement strategy. These values can be calculated using the following equation:

$$s_k = \sum_{j=1}^{n} w_j r_{kj} = \frac{\sum_{j=1}^{n} w_j \left(\left| f_j^{\text{aspired}} - f_{kj} \right| \right)}{\left(\left| f_j^{\text{aspired}} - f_j^{\text{worst}} \right| \right)}, \tag{11}$$

where s_k is defined as the normalized ratio (%) of distance to the aspiration level, which implies the synthesized gap of the criteria. In this case, w_j indicates the IWs for the criteria obtained from DANP.

4. Results and Discussion

4.1. Using FDM to Effectively Screen Criteria. In this research, the design of the fuzzy Delphi questionnaire was based on related literature, and the content was oriented to the design strategies to improve the health and well-being of the aged in green open spaces old downtown residential communities. Researchers summarized the relevant influential elements acquired by close reading (of the interview transcript verbatim) and considering the multiple meanings inherent in the text. In the analysis process, the meanings were described in text (such as correlation, chain, and implication) initially through continuous analysis and comparison on the modification and improvement of each influential element. Furthermore, the elements with similar meanings were combined in categories (i.e., outdoor activities, touching the nature, and social interaction). Through inductive analysis of related materials, 23 influential elements were extracted, and each design element was provided with supplementary instruction as per the coding records, as shown in Table 1.

The 23 design influential elements in Table 1 were further screened and analyzed to make clear their effectiveness in serving as evaluation criteria. Moreover, the influential elements were screened using the FDM method; experts from relevant fields were invited to fill out the fuzzy Delphi questionnaire for experts according to a scale of 1–10 as per

the importance of each influential element, wherein the experts had no less than 5 years of working experience and engaged teaching related to environment art design, scientific research, and architectural landscape design. Finally, a total of 38 questionnaires were distributed and 30 valid questionnaires were collected. In this research, FDM was used to integrate expert opinions based on "double-triangular fuzzy numbers," and the "grey zone test" was taken as an effective way to test the consistency of the results recognized by experts. Based on the semantic scale of a Likert questionnaire, 6.667 was determined as the threshold value of consistency in this research. As a result, among the 23 influential elements, 3 elements ($\mathbf{C}_3$; $\mathbf{C}_{14}$; $\mathbf{C}_{21}$) scored in the range of 6.5–7, 6 elements ($\mathbf{C}_6$; $\mathbf{C}_9$; $\mathbf{C}_{12}$; $\mathbf{C}_{13}$; $\mathbf{C}_{18}$; $\mathbf{C}_{20}$) scored in the range of 7.1–7.5, 7 elements ($\mathbf{C}_4$; $\mathbf{C}_5$; $\mathbf{C}_7$; $\mathbf{C}_8$; $\mathbf{C}_{10}$; $\mathbf{C}_{11}$; $\mathbf{C}_{22}$) scored in the range of 7.6–8, and 7 elements ($\mathbf{C}_1$; $\mathbf{C}_2$; $\mathbf{C}_{15}$; $\mathbf{C}_{16}$; $\mathbf{C}_{17}$; $\mathbf{C}_{19}$; $\mathbf{C}_{23}$) scored above 8. The second round of expert interviews only addressed the three influential elements that scored in the range of 6.5–7. On the basis of the interview, the fuzzy Delphi questionnaire was distributed again to confirm the importance of the three elements. The analysis results show that the consistence values of $\mathbf{C}_{14}$ and $\mathbf{C}_{21}$ fail to pass the threshold so that they are excluded from the effective influential criteria. The remaining 21 influential elements can be used as important evaluation criteria for reconstruction of green open spaces in old downtown residential communities. The statistical results are shown in Table 2. The expert consistence values of each criterion are listed in the rightmost column.

4.2. Using DANP to Clarify the Influence Relation between and Weights of Evaluation Criteria. Based on Fuzzy Delphi, the effectiveness of criteria was screened and $\boldsymbol{C}_{14}$ and $\boldsymbol{C}_{21}$ were evaluated so that they could be excluded from the effective criteria. In this section, the findings show how the DANP-V model can clarify the influence relationships of the screened criteria and allocate weights. This survey distributed 32 DEMATEL questionnaires. Some of these questionnaires were collected through expert group meetings and were based on the results of expert discussions. The consistency reached more than 96% by the end of the calculation. Therefore, 27 valid expert questionnaires were selected. The INRM in Figure 2 clearly expresses the influence network relationship structure of the interaction in the improved model. INRM indicates that the impact priority of these three dimensions can be sorted in the following order: Social Interaction ($\boldsymbol{D}_3$) > Connect with Nature ($\boldsymbol{D}_2$) > Outdoor Activities ($\boldsymbol{D}_1$), and Outdoor Activities ($\boldsymbol{D}_1$) were influenced by Social Interaction ($\boldsymbol{D}_3$) and Connect with Nature ($\boldsymbol{D}_2$), whereas Connect with Nature ($\boldsymbol{D}_2$) was influenced by Social Interaction ($\boldsymbol{D}_3$). This result shows that when considering improving the dimensions of Outdoor Activities ($\boldsymbol{D}_1$), the Social Interaction ($\boldsymbol{D}_3$) and Connect with Nature ($\boldsymbol{D}_2$) dimensions should be emphasized. In addition, in order to improve the Connect with Nature ($\boldsymbol{D}_2$) dimension of the green space, improving the standards of the Social Interaction ($\boldsymbol{D}_3$) dimension should be highlighted. Similarly, from the perspective of INRM, the most influential criterion for Outdoor Activities ($\boldsymbol{D}_1$) is Suitable lighting ($\boldsymbol{C}_4$), and the order of influence of the remaining criteria is $\boldsymbol{C}_9$, $\boldsymbol{C}_3$, $\boldsymbol{C}_6$, $\boldsymbol{C}_2$, $\boldsymbol{C}_8$, $\boldsymbol{C}_1$, $\boldsymbol{C}_7$, and $\boldsymbol{C}_5$. Natural water characteristics ($\boldsymbol{C}_{10}$) was the most powerful criterion in the Connect with Nature ($\boldsymbol{D}_2$) dimension, and the order of influence of the remaining criteria is $\boldsymbol{C}_{15}$, $\boldsymbol{C}_{11}$, $\boldsymbol{C}_{16}$, $\boldsymbol{C}_{13}$, and $\boldsymbol{C}_{12}$. Public health service station ($\boldsymbol{C}_{23}$) was the most powerful criterion in the Social Interaction ($\boldsymbol{D}_3$) dimension, and the order of influence of the remaining criteria is $\boldsymbol{C}_{22}$, $\boldsymbol{C}_{17}$, $\boldsymbol{C}_{18}$, $\boldsymbol{C}_{20}$, and $\boldsymbol{C}_{19}$.

4.3. Empirical Cases. This section introduces the green open space cases of three communities among old downtown residential communities in a high-density city represented by Guangzhou, as well as the data collection and analysis process. In addition, sustainable and systematic improvement strategies for each real case are discussed on the basis of INRM and current performance evaluation results, taking INRM and IWs as the standards and expecting to enhance the health and well-being of the aged.

4.3.1. Old Nanhai County Community on Liurong Street. Under the jurisdiction of the Office of Liurong Street, Yuexiu District, Guangzhou City, Old Nanhai County Community has a total land area of about 4.28 hectares, with about 1,840 resident households and 5,155 residents. The aged residents account for about 38.79% of the population. The plan is shown in Figure 3. In the dimension of outdoor activities, most of the slow walking paths in the community are provided next to buildings (Part A of Figure 3). Having a long history, the community borders Qing General's Office site on Liurong Road in the east, Zhongshan Road in the south, Old Nanhai County Street in the west, and Liurong Temple on Fuquan Lane 1 in the north. The surrounding public transportation is well connected and accessible. After a micro reconstruction, the community is now equipped with relatively complete lighting and guidance systems (Part B of Figure 3) as well as infrastructures such as garbage-sorting stations and public toilets (Part C and Part D of Figure 3). A community square is provided with a space for fitness activity. The sewage system is well maintained. However, under the epidemic, there is lack of related outdoor ventilation machines and air cleaning devices. In the dimension of connecting with nature, the reconstruction did not take full consideration of creating natural water landscapes, soundscapes, plant arrangements, and the like due to the orientation toward cultural inheritance (Part E of Figure 3). However, there are complete interactive landscapes and outdoor landscapes. No obvious source of noise can be found in the community. In the dimension of social interaction, the community is a cultural propaganda demonstration community, with complete safety management facilities (such as monitoring systems) (Part F of Figure 3) as well as an activity square for residents to relax and socialize, multifunctional open spaces, and public health service stations (Part G, Part H, Part I, and Part J of Figure 3). Unfortunately, no spaces for children to play were found in the field investigation. The cultural heritage preservation

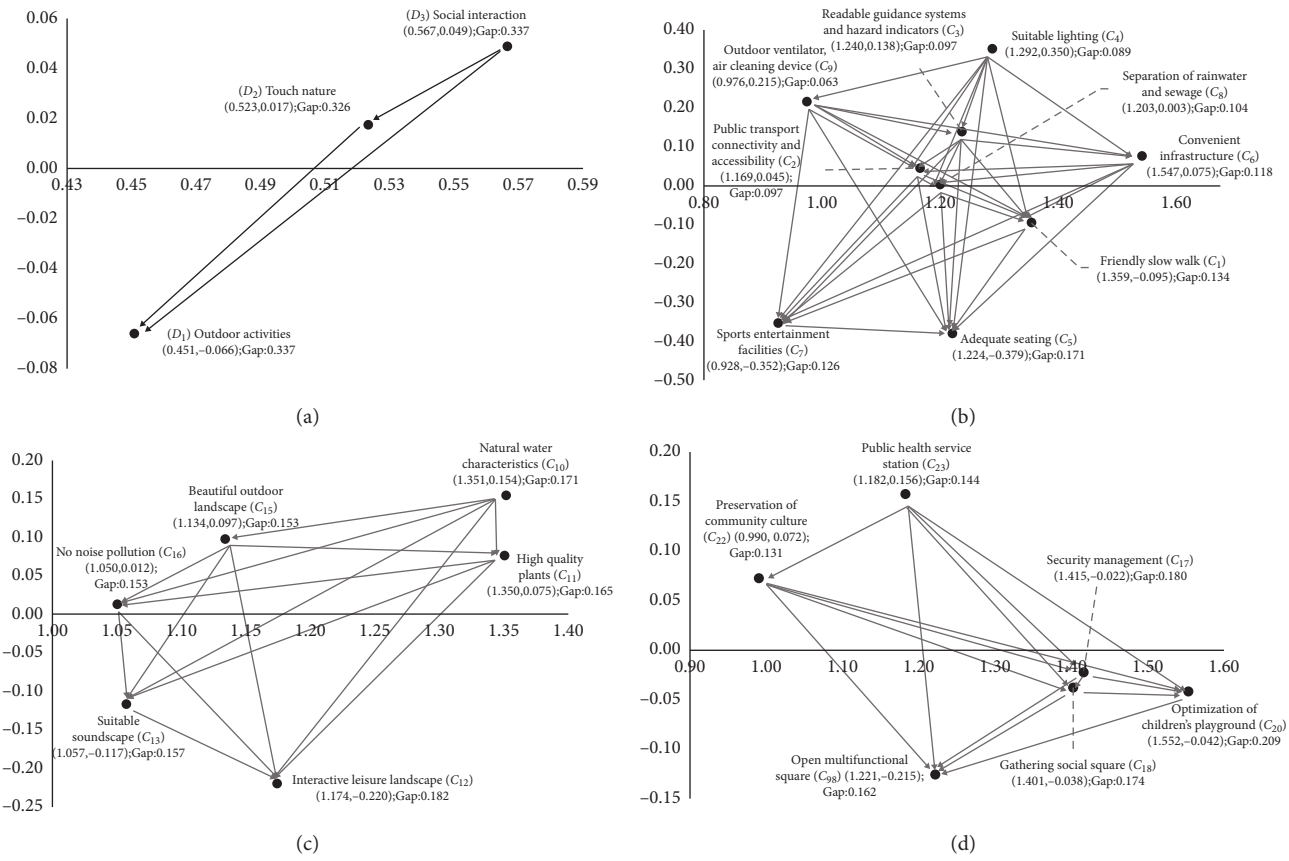


Figure 2: The INRM (influential network relation map) of the total influence relationships.

Figure 3: Plan of the Old Nanhai County Community on Liurong street.

design is worthy of reference for reconstruction of other communities (Part K and Part L of Figure 3).

4.3.2. Panqiu Community on Jinhua Street. Located on Jinhua Street, Liwan District, Guangzhou City, Guangdong Province and spanning from Anlongli, Xihua Road in the east and Ludi Street, North Guangfu Road in the south, to Panqiu Street, Xihua Road in the west and Xihua Road in the north, the Panqiu Community has a total area of about 0.089 square kilometers, with 1,478 permanent resident households, and 4,418 residents in total. The plan is shown in Figure 4. Jinhua Street Nursing Home and Jinhua Cultural Plaza falls within the area and most of the people active in the community are aged. Hence, well-being facilities and public infrastructures can be guaranteed (Part A, Part B, and Part C of Figure 4). In the community, there are complete tourism, culture, health care, education, and shopping facilities, and a good guidance system, instructions, traffic connection, and accessibility (Part D of Figure 4). A cultural square is provided with sports and fitness facilities and recreational areas (Part E and Part F of Figure 4). However, there is no children's playground. In the old downtown area of Guangzhou, micro-reconstruction activities have been carried out in recent years. Subsequently, the underground pipeline facilities of the community have been repaired, the existing lighting system is relatively complete, and the roads (including slow walking path) are well maintained (Part G of Figure 4). But with the outbreak, outdoor ventilation machines and air cleaning devices are still lacking. As a gathering area for the aged, the community was mainly reconstructed along humanity orientation lines. Safety management facilities (such as monitoring systems), public health service stations, and other facilities are in relatively good condition. Elevators are being added to old buildings in the community. So, noise mainly comes from the construction (Part H of Figure 4). However, the design considerations of plant arrangement, natural water landscapes, soundscapes, and other aspects remain to be improved (Part I of Figure 4). It was this community's focus on public infrastructure for the aged that provided the idea for conducting this research.

4.3.3. Sanfu Community on Jinhua Street. Located east of Jinhua Street, Liwan District, Guangzhou and bordering North Renmin Road in the east, North Guangfu Road in the west, 7th Zhongshan Road in the south, and Xihua Road in the north (the plan is shown in Figure 5), the Sanfu Community has a total area of about 0.01 square kilometers, with 1,520 resident households, 4,626 permanent residents, and 1097 immigrant residents. A total of 303 households are not registered in the community. The total population has reached up to 5,723 persons. The neighborhood committee currently involves 28 relief households and 14 low-income households. In addition, 72 persons with disabilities are registered. The current situation of public open spaces in the community is as follows. First of all, in the dimension of outdoor activities, most of the walking paths in the community are provided around residential buildings, and are shared by people and vehicles. However, many of these places do not provide barrier-free facilities (Part A and Part B of Figure 5). Public transportation around the community is well connected and accessible, and there are multiple bus lines and subways passing through. Rest seats and lighting existing in the community are old and somewhat damaged. No specific landscape lights are provided in the public open space (Part C of Figure 5). But in terms of guidance system (15-minute living circle in the community), there is relatively complete infrastructure such as a garbage sorting station (Part D of Figure 5). Public open spaces in the community are equipped with fitness equipment donated by the society and government (Part E and Part F of Figure 5). Under the outbreak, no outdoor ventilation machines and air cleaning devices have been installed; but measures such as regular manual spraying for disinfection are taken. In the dimension of connecting with nature, due to the impact of limited land for use, there is no large lawn area but merely a few small trees scattered and irregularly distributed along the street and in public open space. Plant arrangement, natural water landscapes, soundscapes, and the like are not reflected in the public open spaces of the community (Part G, Part H, Part I, and Part J of Figure 5). No obvious noise source is found. In the dimension of social interaction, safety management facilities (such as monitoring systems) are relatively complete. A paved square is the main form of public open space in the community and can be used to develop various activities and leisure and social interaction among residents (Part C of Figure 5). No space for children to play was found. However, field investigation found that most children play in the fitness activity area (Part K of Figure 5). The old city wall and the characteristic buildings of Xi Guan make up the landscape and cultural features of the community, but they are not embodied in the public open space.

4.3.4. Using VIKOR to Make Performance Evaluation and Improvement. On the basis of INRM and IWs discovered by DANP, we can ensure performance and improvement by using the VIKOR technique and further formulate sustainable and systematic strategies for improving green open space in real cases so as to promote healthy aging. Table 3 lists the performance and gap values of the dimensions/criteria in the three real cases. We assume that when the gap values are improved, the performance of each dimension/criterion will increase accordingly, and the overall case will tend to reach the desired level. Traditional discussions on evaluation and improvement often focus on the most significant practical problems (the largest gap), and the priority of improvement is also determined on this basis. However, the DANP-V model uses a systematic method to clarify the influence relation between the criteria and discover the influential source of the practical problems.

In the Old Nanhai County Community, the gap values of the three dimensions are ranked from large to small as follows: connecting with nature ($\mathbf{D}_2$), outdoor activities ($\mathbf{D}_1$), and social interaction ($\mathbf{D}_3$). As shown in Table 3, natural water feature ($\mathbf{C}_{10}$) has the largest gap value (0.5545) in $\mathbf{D}_2$ and also has strong IW. In comparison, this

Figure 4: Plan of the Panqiu Community on Jinhua street.

community currently has good, friendly slow walking paths, and has also obtained high satisfaction in terms of safety management and cultural preservation of the community. More importantly, INRM shows that $\mathbf{C}_{10}$ is the most dominant criterion in the specific dimension. Therefore, for improvement of the case, decision-makers should allocate more resources to natural water features ($\mathbf{C}_{10}$) to enhance the performance of the case in other criteria. On the other hand, the influence of $\mathbf{D}_3$ on $\mathbf{D}_2$ reminds decision-makers that they should pay more attention to some of the problems presented in the dimension of social interaction. For example, Table 3 shows that the community has poor performance in $\mathbf{C}_{18}$ and $\mathbf{C}_{20}$. This implies that decision-makers can think about how to meet people's demands for social interaction and children's entertainment based on the hydrophilicity of people. This could not only improve dimension $\mathbf{D}_3$ but also drive improvement of $\mathbf{D}_2$.

In the Sanfu Community, the gap values of the three dimensions are ranked from large to small as follows: connecting with nature ($\mathbf{D}_2$), outdoor activities ($\mathbf{D}_1$) and social interaction ($\mathbf{D}_3$). Except for $\mathbf{C}_{10}$, the community performs poorly in superior vegetation ($\mathbf{C}_{11}$) under the second dimension, and performs the worst in outdoor ventilation and air cleaning ($\mathbf{C}_9$) under $\mathbf{D}_1$. However, currently, the community has good traffic conditions and accessibility, and good performance in safety management and noise control. Figure 2 shows that in addition to incorporating a natural water element into the community, decision-makers should also pay attention to the aesthetics of the physical environment, including selection of plant species in the field, consideration of configuration and decoration, color matching, form arrangements, etc. Similarly, decision-makers should also start promoting social communication and guaranteeing healthy rights and interests of the residents, providing more convenient service facilities and open squares or pocket parks with good atmospheric conditions. Ventilation conditions of the structures and associated facilities should also be taken into consideration in the planning and design.

As can be seen from the performance results, the Panqiu Community also has the worst performance in $\mathbf{D}_2$ (0.5602), and slightly better performance in $\mathbf{D}_3$ (0.3476) than in $\mathbf{D}_1$. In addition to natural water features ($\mathbf{C}_{10}$), the community also performs poorly in noise pollution conditions under $\mathbf{D}_2$. At the same time, the community currently does not have high-quality plant arrangement and cannot form an interactive landscape for residents to relax. In $\mathbf{D}_1$, air quality, convenient basic service facilities, and treatment of rainwater and sewage in the community cannot be well guaranteed to the residents. However, the community currently has good conditions and accessibility and can provide the residents with good safety management and health services. Through field investigation, it was revealed that the current noise pollution in the community is due to the government-supported

Figure 5: Plan of Sanfu community on Jinhua street.

project for installing elevators in old residential buildings. The ongoing reconstruction project does not involve the overall consideration of high-quality plants, and it is even foreseeable that the project will occupy green area and public activity space.

The performance results show that among the three real cases, Sanfu Community is least suited to improving health and well-being construction for the aged, while the current development of better community is Old Nanhai County. However, the three real cases show high degrees of

Table 3: Performance Evaluation of the Case Study using VIKOR.

Dimensions/ Criteria	Influential Weights (IWs)	Old Nanhaixian Community		Sanfu Community		Panqiu Community	
		Performance f1j	Gap Ratio	Performance f2j	Gap Ratio	Performance f3j	Gap Ratio
Outdoor Activities($D1$)	**0.337**	**3.287**	**0.343**	**2.703**	**0.460**	**3.161**	**0.368**
Friendly slow walking ($C1$)	0.134	4.136	0.173	2.700	0.460	3.350	0.330
Public transport connectivity and accessibility ($C2$)	0.097	3.818	0.236	3.350	0.330	3.600	0.280
Readable guidance systems and hazard indicators ($C3$)	0.097	3.591	0.282	2.900	0.420	3.250	0.350
Suitable lighting ($C4$)	0.089	3.364	0.327	2.850	0.430	3.150	0.370
Adequate seating ($C5$)	0.171	3.091	0.382	2.700	0.460	3.350	0.330
Convenient infrastructure ($C6$)	0.118	2.818	0.436	2.350	0.530	2.750	0.450
Sports entertainment facilities ($C7$)	0.126	2.909	0.418	2.700	0.460	3.200	0.360
Separation of rainwater and sewage ($C8$)	0.104	2.909	0.418	2.500	0.500	2.850	0.430
Outdoor ventilator/air cleaning device ($C9$)	0.063	2.864	0.427	2.200	0.560	2.650	0.470
Touch Nature($D2$)	**0.326**	**3.002**	**0.400**	**2.323**	**0.535**	**2.199**	**0.560**
Natural water characteristics ($C10$)	0.171	2.227	0.555	1.800	0.640	1.750	0.650
High quality plants ($C11$)	0.165	2.864	0.427	2.100	0.580	2.200	0.560
Interactive Leisure Landscape ($C12$)	0.182	2.636	0.473	2.250	0.550	2.250	0.550
Suitable Soundscape ($C13$)	0.157	3.046	0.391	2.400	0.520	2.400	0.520
Beautiful outdoor landscape ($C15$)	0.153	3.546	0.291	2.400	0.520	2.600	0.480
No noise pollution ($C16$)	0.171	3.773	0.246	3.000	0.400	2.050	0.590
Social interaction($D3$)	**0.337**	**3.570**	**0.286**	**2.765**	**0.447**	**3.262**	**0.348**
Security management ($C17$)	0.180	4.091	0.182	3.150	0.370	3.550	0.290
Gathering social square ($C18$)	0.174	3.227	0.355	2.850	0.430	3.400	0.320
Total Performance		**3.289**		**2.600**		**2.881**	
Total Gap			**0.342**		**0.480**		**0.424**

Note: gap ratio $r_{kj} = (|f_j^{\text{aspired}} - f_{kj}|)/(|f_j^{\text{aspired}} - f_j^{\text{worst}}|) = (5 - f_{kj})/(5 - 0)$, $s_k = \sum_{j=1}^{n} w_j r_{kj}$, where f_{kj} denotes j criterion f of alternative, i.e., $f = 1(ONHX)$, $f = 2(SF)$ and $f = 3(PQ)$.

consistency in the three dimensions of performance; the notable problems are concentrated in $\mathbf{D}_2$. Combining practical experience, it can be deduced that green open spaces in the old downtown residential communities of Guangzhou generally lack resource input and planning and design considerations in the aspect of connecting with nature. However, Douglas et al. [18] and Azadeh Lak et al. [19] have expressed that connecting with nature plays a fundamental role in the physical and mental health of the aged and accomplishes the internal link between humans and nature. Thereby, it is believed that in follow-up planning of old downtown reconstruction, decision-makers should try to improve the opportunities and quality of connecting with nature for the aged in old communities in order to improve their health and well-being. Furthermore, we propose incorporating natural blue and green elements into the design of public living spaces in communities from the perspective of social interaction and care based on old people's preference for natural water environments and ornamental plants and the creation of interactive landscapes. To sum up, a bottom-up observation of the behavioral habits, perception needs, and psychological preferences of the aged in old communities is the key foundation for creating good outdoor activities and natural landscapes in green open spaces of communities. Especially after the outbreak of COVID-19, decision-makers should strengthen safety guarantees for social interaction of the aged in communities, and update the previous practices and contents of social activities on the basis of meeting the needs of the aged to practically create healthy, safe, and well-used outdoor activity spaces and natural landscapes with high ornamental value and interactivity.

5. Conclusions

This study aims to provide theoretical research and practical improvement strategies for improving the health and well-being of the aged in old downtown residential communities. First of all, 23 elements of green open spaces that can affect the health of the aged were extracted from previous literature and divided into three dimensions, namely, outdoor activities ($\boldsymbol{D}_1$), connecting with nature ($\boldsymbol{D}_2$), and social interaction ($\boldsymbol{D}_3$). Second, applicability and effectiveness of the 23 criteria were judged on the basis of expert knowledge. On this basis, an evaluation framework for evaluating green open spaces in old downtown residential communities was constructed. Combined with public health and urban design, the evaluation framework was further used to view the status quo of community environments and provide decision-making strategies for improving the health and well-being of

the aged. Third, the influence relation and priority between the dimensions and criteria of the evaluation framework were clarified. In an effort to loosen the independence between the criteria, an Influential Network Relation Map (INRM) was drawn, which made it possible to not only recognize the current dominant problems by situation evaluation analysis but also to identify the influential sources of the problems by clarifying the dominant influential relation. On this basis, improvement strategy was formulated, making sure to avoid the disadvantages of "taking stopgap measures" in traditional evaluation research and averting waste to the maximum extent possible on the premise of limited resources. Finally, taking Old Nanhai County Community, Sanfu Community, and Panqiu Community as representative real cases, performance evaluation was made to explore the focus issues existing in reality, and discussion of sustainable and systematic improvement strategies applicable to each of the communities was done based on INRM.

The limitations of this research should be acknowledged, namely, based on performance evaluation analysis, the improvement strategies explored in this research are only formulated for the three residential communities in the old downtown area of Guangzhou. Therefore, it is not recommended to apply these improvement strategies in relevant cases in other regions. However, the process of developing improvement strategies and the evaluation framework provided in this research can be used as a reference for investigating improvement priorities of residential communities in many areas. In addition, the modified VIKOR applied in this research is an additive performance evaluation method, while nonadditive situations are often the case in reality. Hence, in follow-up research, nonadditive performance evaluation methods can be taken to get closer to the real situation. On the other hand, if objective data can be collected by investigating a large number of old downtown residential communities, it would be possible to further make clear the core of the evaluation framework constructed in this research by using data exploration techniques such as Rough Set Approach (RSA) and clarifying the decision rule (if-then) in the investigated data based on the research results. Moreover, in combination with expert knowledge, the DEMATEL technique could be used to explore "decision-making path maps" designed for green open spaces in old downtown residential communities to improve the health and well-being of the aged.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare no conflicts of interest.

Authors' Contributions

Yu-Feng Luo and Si-Jie Li contributed equally in this study.

Acknowledgments

The authors are grateful to all the landscape designers who took part in this study. The authors would like to thank Ye Tao, the research assistant, for all the hard work and effort put into making the questionnaire survey a success. The authors are also immeasurably grateful to the alumni association of environmental design at the School of Architecture, Harbin Institute of Technology (HIT), for support during the research. The authors are grateful to the Graduate Education Innovation Project of Guangdong Province (Grant No. 2019JGXM36) and the Specialized Subsidy Scheme for Macao Higher Education Institutions in the Area of Research in Humanities and Social Sciences (and Specialized Subsidy Scheme for Prevention and Response to Major Infectious Diseases) (Grant No. HSS-MUST-2020-9) for their financial support. In addition, this study was also funded by the art projects supported by the National Social Science Foundation, project no. 18BG08, and the Graduate Teaching Reform Project of Guangzhou Academy of Fine Arts (Grant No. 6040119012).

References

[1] World Health Organization, *Development of a Proposal for a Decade of Healthy Ageing 2020–2030 Report by the Director-General*, World Health Organization, Geneva, Switzerland, 2019, https://apps.who.int/gb/ebwha/pdf_files/EB146/B146_23-en.pdf.

[2] World Health Organization, *Multisectoral Action for a Life Course Approach to Healthy Ageing: Draft Global Strategy and Plan of Action on Ageing and Health*, World Health Organization, Geneva, Switzerland, 2016, https://apps.who.int/gb/ebwha/pdf_files/WHA69/A69_17-en.pdf?ua=1.

[3] United Nations General Assembly, *World Population Prospects (Revised Report). Quito:*, United Nations General Assembly, New York,, NY, USA, 2017, https://www.un.org/en/sections/issues-depth/ageing/.

[4] C. W. Thompson, "Urban open space in the 21st century," *Landscape and Urban Planning*, vol. 60, no. 2, pp. 59–72, 2002.

[5] A. Haleem, M. Javaid, R. Vaishya, and S. G. Deshmukh, "Areas of academic research with the impact of COVID-19," *The American Journal of Emergency Medicine*, vol. 38, no. 7, pp. 1524–1526, 2020.

[6] N. A. Megahed and E. M. Ghoneim, "Antivirus-built environment: lessons learned from covid-19 pandemic," *Sustainable Cities and Society*, vol. 61, Article ID 102350, 2020.

[7] National Health Commission of the People's Republic of China, *Notice on Prevention and Control of New Coronary Virus Infection Pneumonia in the Elderly*, National Health Commission of the People's Republic of China, Beijing, China, 2020.

[8] F. Xu and Y. Zhou, "Social ex-post evaluation of old city zone renovation projects based on mathematical attribute measurement model," *Journal of Urban Planning & Development*, vol. 143, no. 1, Article ID 04016025, 2016.

[9] X. Sun, L. Wang, F. Wang, and S. Soltani, "Behaviors of seniors and impact of spatial form in small-scale public spaces in Chinese old city zones," *Cities*, vol. 107, Article ID 102894, 2020.

[10] H. Yang and J. Ma, "How an epidemic outbreak impacts happiness: factors that worsen (vs. Protect) emotional well-

being during the coronavirus pandemic," *Psychiatry Research*, vol. 289, Article ID 113045, 2020.

[11] D. E. V. Hoof, "Lockdown is the world's biggest psychological experiment," 2020, https://www.weforum.org/agenda/2020/04/this-isthe-psychological-side-of-the-covid-19-pandemic-that-were-ignoring/Available online:.

[12] N. K. Fofana, F. Latif, S. Sarfraz, M. F. Bashir, and B. Komal, "Fear and agony of the pandemic leading to stress and mental illness: an emerging crisis in the novel coronavirus (covid-19) outbreak," *Psychiatry Research*, vol. 291, Article ID 113230, 2020.

[13] M. Ekström, "Elderly people's experiences of housing renewal and forced relocation: social theories and contextual analysis in explanations of emotional experiences," *Housing Studies*, vol. 9, no. 3, pp. 369–391, 1994.

[14] A. Gaikwad and K. Shinde, "Use of parks by older persons and perceived health benefits: a developing country context," *Cities*, vol. 84, pp. 134–142, 2019.

[15] B.-W. Zhu, J.-R. Zhang, G.-H. Tzeng, S.-L. Huang, and L. Xiong, "Public open space development for elderly people by using the DANP-V model to establish continuous improvement strategies towards a sustainable and healthy aging society," *Sustainability*, vol. 9, no. 3, p. 420, 2017.

[16] Z. Allam and D. S. Jones, "On the coronavirus (COVID-19) outbreak and the smart city network: universal data sharing standards coupled with artificial intelligence (AI) to benefit urban health monitoring and management," *Healthcare Multidisciplinary Digital Publishing Institute*, vol. 8, no. 1, p. 46, 2020.

[17] J. H. Spencer, S. Saksena, and J. Fox, "New findings on links between urban expansion and viral disease in vietnam offer lessons for COVID-19," 2020, http://128.171.57.22/bitstream/10125/67578/1/20200401-New%20Findings%20on%20Links%20between%20Urban%20Expansion%20and%20Viral%20Disease%20in%20Vietnam.pdf.

[18] O. Douglas, M. Lennon, and M. Scott, "Green space benefits for health and well-being: a life-course approach for urban planning, design and management," *Cities*, vol. 66, pp. 53–62, 2017.

[19] A. Lak, R. Aghamolaei, H. R. Baradaran, and P. K. Myint, "A framework for elder-friendly public open spaces from the Iranian older adults' perspectives: a mixed-method study," *Urban Forestry & Urban Greening*, vol. 56, Article ID 126857, 2020.

[20] J.-J. Chen, Z.-X. Zhang, and Y. Long, "Strategies for improving the quality of public health-oriented street space —— from the perspective of spatial disorder," *City Planning Review*, vol. 44, 2020.

[21] G.-H. Tzeng and K.-Y. Shen, "New concepts and trends of hybrid," in *Multiple Criteria Decision Making*CRC Press; Taylor and Francis Group, Boca Raton, FL, USA, 2017.

[22] G. B. Qu, T. Y. Zhao, B. W. Zhu, G. H. Tzeng, and S. L. Huang, "Use of a modified danp-mv model to improve quality of life in rural residents: the empirical case of xingshisi village, China," *International Journal of Environmental Research and Public Health*, vol. 16, no. 1, 2019.

[23] L. Xiong, C. L. Teng, Y. Q. Li, Y. Z. Lee, B. W. Zhu, and K. Liu, "A qualitative-quantitative evaluation model for systematical improving the creativity of students' design scheme," *Sustainability*, vol. 11, 2019.

[24] L. Levy-Storms, L. Chen, and A. Loukaitou-Sideris, "Older adults' needs and preferences for open space and physical activity in and near parks: a systematic review," *Journal of Aging and Physical Activity*, vol. 26, no. 4, pp. 682–696, 2018.

[25] H. K. Esther, K. O. Winky, and H. W. Edwin, "Elderly satisfaction with planning and design of public parks in high density old districts: an ordered logit model," *Landscape and Urban Planning*, vol. 165, pp. 39–53, 2017.

[26] E. Cerin, C. H. P. Sit, A. Barnett, M.-C. Cheung, and W.-M. Chan, "Walking for recreation and perceptions of the neighborhood environment in older Chinese urban dwellers," *Journal of Urban Health*, vol. 90, no. 1, pp. 56–66, 2012.

[27] M. Moran, J. Van Cauwenberg, R. Hercky-Linnewiel, E. Cerin, B. Deforche, and P. Plaut, "Understanding the relationships between the physical environment and physical activity in older adults: a systematic review of qualitative studies," *International Journal of Behavioral Nutrition and Physical Activity*, vol. 11, no. 1, 2014.

[28] C. Ward Thompson and P. A. Aspinall, "Natural environments and their impact on activity, health, and quality of life," *Applied Psychology: Health and Well-Being*, vol. 3, no. 3, pp. 230–260, 2011.

[29] K. Villanueva, H. Badland, P. Hooper et al., "Developing indicators of public open space to promote health and wellbeing in communities," *Applied Geography*, vol. 57, pp. 112–119, 2015.

[30] T. Schmidt, J. Kerr, and J. Schipperijn, "Associations between neighborhood open space features and walking and social interaction in older adults-A mixed methods study," *Geriatrics*, vol. 4, no. 3, p. 41, 2019.

[31] M. Artmann, X. Chen, C. Iojă et al., "The role of urban green spaces in care facilities for elderly people across European cities," *Urban Forestry & Urban Greening*, vol. 27, pp. 203–213, 2017.

[32] P. A. Aspinall, C. Ward Thompson, S. Alves, T. Sugiyama, R. Brice, and A. Vickers, "Preference and relative importance for environmental attributes of neighbourhood open space in older people," *Environment and Planning B: Planning and Design*, vol. 37, no. 6, pp. 1022–1039, 2010.

[33] A. Kemperman and H. Timmermans, "Green spaces in the direct living environment and social contacts of the aging population," *Landscape and Urban Planning*, vol. 129, pp. 44–54, 2014.

[34] K. L. Cain, R. A. Millstein, J. F. Sallis et al., "Contribution of streetscape audits to explanation of physical activity in four age groups based on the Microscale Audit of Pedestrian Streetscapes (MAPS)," *Social Science & Medicine*, vol. 116, pp. 82–92, 2014.

[35] J. Sallis, K. Cain, T. Conway et al., "Is your neighborhood designed to support physical activity? A brief streetscape audit tool," *Preventing Chronic Disease*, vol. 12, 2015.

[36] J. Wang, "Vision of China's future urban construction reform: in the perspective of comprehensive prevention and control for multi disasters," *Sustainable Cities and Society*, vol. 64, Article ID 102511, 2021.

[37] J. Leng, Q. Wang, and K. Liu, "Sustainable design of courtyard environment: from the perspectives of airborne diseases control and human health," *Sustainable Cities and Society*, vol. 62, Article ID 102405, 2020.

[38] A. M. Sheela, A. Ghermandi, P. Vineetha, R. V. Sheeja, J. Justus, and K. Ajayakrishna, "Assessment of relation of land use characteristics with vector-borne diseases in tropical areas," *Land Use Policy*, vol. 63, pp. 369–380, 2017.

[39] G. Buonanno, L. Stabile, and L. Morawska, "Estimation of airborne viral emission: quanta emission rate of SARS-CoV-2 for infection risk assessment," *Environment International*, vol. 141, Article ID 105794, 2020.

[40] C. Sun and Z. Zhai, "The efficacy of social distance and ventilation effectiveness in preventing COVID-19 transmission," *Sustainable Cities and Society*, vol. 62, Article ID 102390, 2020.

[41] F. A. Berlanga, M. R. De Adana, I. Olmedo, J. M. Villafruela, J. F. San José, and F. Castro, "Experimental evaluation of thermal comfort, ventilation performance indices and exposure to airborne contaminant in an airborne infection isolation room equipped with a displacement air distribution system," *Energy and Buildings*, vol. 158, pp. 209–221, 2018.

[42] B. Munyan, S. Neer, D. Beidel, and F. Jentsch, "Olfactory stimuli increase presence in virtual environments," *PLoS One*, vol. 11, no. 6, Article ID e0157568, 2016.

[43] R. Leaver and T. Wiseman, "Garden visiting as a meaningful occupation for people in later life," *British Journal of Occupational Therapy*, vol. 79, no. 12, pp. 768–775, 2016.

[44] J. Finlay, T. Franke, H. McKay, and J. Sims-Gould, "Therapeutic landscapes and wellbeing in later life: impacts of blue and green spaces for older adults," *Health & Place*, vol. 34, pp. 97–106, 2015.

[45] J. J. Alvarsson, S. Wiens, and M. E. Nilsson, "Stress recovery during exposure to nature sound and environmental noise," *International Journal of Environmental Research and Public Health*, vol. 7, no. 3, pp. 1036–1046, 2010.

[46] E. Ratcliffe, B. Gatersleben, and P. T. Sowden, "Bird sounds and their contributions to perceived attention restoration and stress recovery," *Journal of Environmental Psychology*, vol. 36, pp. 221–228, 2013.

[47] Ryan, C. 14 Patterns of Biophilic Design: Mystery.

[48] G. B. Diette, N. Lechtzin, E. Haponik, A. Devrotes, and H. R. Rubin, "Distraction therapy with nature sights and sounds reduces pain during flexible bronchoscopya," *Chest*, vol. 123, no. 3, pp. 941–948, 2003.

[49] M. T. Gonzalez and M. Kirkevold, "Benefits of sensory garden and horticultural activities in dementia care: a modified scoping review," *Journal of Clinical Nursing*, vol. 23, no. 19, pp. 2698–2715, 2014.

[50] S.-H. Park and R. H. Mattson, "Effects of flowering and foliage plants in hospital rooms on patients recovering from abdominal surgery," *HortTechnology*, vol. 18, no. 4, pp. 563–568, 2008.

[51] E. H. K. Yung, S. Conejos, and E. H. W. Chan, "Social needs of the elderly and active aging in public open spaces in urban renewal," *Cities*, vol. 52, pp. 114–122, 2016.

[52] W. Sun, L. Tang, and B. L. Allman, "Environmental noise affects auditory temporal processing development and NMDA-2B receptor expression in auditory cortex," *Behavioural Brain Research*, vol. 218, no. 1, pp. 15–20, 2011.

[53] C. R. Kight and J. P. Swaddle, "How and why environmental noise impacts animals: an integrative, mechanistic review," *Ecology Letters*, vol. 14, no. 10, pp. 1052–1061, 2011.

[54] K. Hung and J. L. Crompton, "Benefits and constraints associated with the use of an urban park reported by a sample of elderly in Hong Kong," *Leisure Studies*, vol. 25, no. 3, pp. 291–311, 2006.

[55] S. A. Stansfeld and M. P. Matheson, "Noise pollution: non-auditory effects on health," *British Medical Bulletin*, vol. 68, no. 1, pp. 243–257, 2003.

[56] V. Ceccato and R. Bamzar, "Elderly victimization and fear of crime in public spaces," *International Criminal Justice Review*, vol. 26, no. 2, pp. 115–133, 2016.

[57] C. Y. Jim and X. Shan, "Socioeconomic effect on perception of urban green spaces in Guangzhou, China," *Cities*, vol. 31, pp. 123–131, 2013.

[58] F. Bastani, H. Haghani, and P. Goodarzi, "Age-friendly cities features from the elderly's perspectives underscoring "community support and health services"" *Journal of Client-Centered Nursing Care*, vol. 2, no. 2, pp. 103–114, 2016.

[59] T. Sugiyama and C. W. Thompson, "Outdoor environments, activity and the well-being of older people: conceptualising environmental support," *Environment and Planning A: Economy and Space*, vol. 39, no. 8, pp. 1943–1960, 2007.

[60] A. D. A. M. Kemperman and H. J. P. Timmermans, "Heterogeneity in urban park use of aging visitors: a latent class analysis," *Leisure Sciences*, vol. 28, no. 1, pp. 57–71, 2006.

[61] N. R. Gee and M. K. Mueller, "A systematic review of research on pet ownership and animal interactions among older adults," *Anthrozoös*, vol. 32, no. 2, pp. 183–207, 2019.

[62] D. Miller, S. Staats, and C. Partlo, "Discriminating positive and negative aspects of pet interaction: sex differences in the older population," *Social Indicators Research*, vol. 27, no. 4, pp. 363–374, 1992.

[63] X. Chen, H. Zhu, and Z. Yuan, "Contested memory amidst rapid urban transition: the cultural politics of urban regeneration in Guangzhou, China," *Cities*, vol. 102, Article ID 102755, 2020.

[64] L. K.-S. Wong and P. H.-F. Yu, "Urban renewal and social sustainability: an exploratory study of the urban life of Wan Chai," *Springerplus*, vol. 4, no. 2, p. P5, 2015.

[65] S. L. Bell, C. Phoenix, R. Lovell, and B. W. Wheeler, "Green space, health and wellbeing: making space for individual agency," *Health & Place*, vol. 30, pp. 287–292, 2014.

[66] C. W. Thompson, "Linking landscape and health: the recurring theme," *Landscape and Urban Planning*, vol. 99, no. 3-4, pp. 187–195, 2011.

[67] S. R. Kellert and E. O. Wilson, *The Biophilia Hypothesis*, Island Press, Washington, D.C., USA, 1993.

[68] S. Habibi Ghahfarrokhi, H. Reza Khankeh, A. A. Akbari Kamrani, and R. Fadayevatan, "What is a key step in the falling process in older people? A qualitative study in an Iranian context," *Educational Gerontology*, vol. 42, no. 4, pp. 265–275, 2016.

[69] P. M. Dall, S. L. H. Ellis, B. M. Ellis et al., "The influence of dog ownership on objective measures of free-living physical activity and sedentary behaviour in community-dwelling older adults: a longitudinal case-controlled study," *BMC Public Health*, vol. 17, no. 1, pp. 1–9, 2017.

[70] G. B. Dalkey, *The Delphi Method: An Experimental Study of Group Opinion*, Rand Corp, Santa Monica, CL, USA, 1969.

[71] A. Ishikawa, M. Amagasa, T. Shiga, G. Tomizawa, R. Tatsuta, and H. Mieno, "The max-min Delphi method and fuzzy Delphi method via fuzzy integration," *Fuzzy Sets and Systems*, vol. 55, no. 3, pp. 241–253, 1993.

[72] T. J. Murray, L. L. Pipino, and J. P. van Gigch, "A pilot study of fuzzy set modification of Delphi," *Human Systems Management*, vol. 5, no. 1, pp. 76–80, 1985.

[73] C. B. A Zheng, "Fuzzy evaluation model for the maturity of software organizations to enhance personnel ability," M. A. Thesis, National Taiwan University of Science and technique, Taipei, Taiwan, 1996.

Hindawi
Journal of Healthcare Engineering
Volume 2021, Article ID 5556172, 11 pages
https://doi.org/10.1155/2021/5556172

Research Article

Application of Hybrid Multiple Attribute Decision-Making Model to Explore the Design Strategies of Children's Facilities in Neighborhood Open Spaces Based on Sensory Integration Theory

Ge Wang, Ruo-Ying Wang, Ting-Lan Liu, Ying Zhuo, and Kang Shen

School of Architecture and Applied Art, Guangzhou Academy of Fine Arts, Guangzhou 511400, China

Correspondence should be addressed to Kang Shen; shk@gzarts.edu.cn

Received 5 February 2021; Revised 12 April 2021; Accepted 19 May 2021; Published 30 May 2021

Academic Editor: Hsiao-Ting Tseng

In recent years, the prevalence of sensory integration disorders in children in urban areas has increased. Most existing sensory integration treatments are located in hospital-based sensory integration units; however, medical resources are extremely limited, making it difficult to guarantee the appropriate treatment time and intervention results for many children. The concept of sensory integration therapy must be disseminated widely and correctly to meet these children's needs. Although most urban communities have a high number of children's spaces, these spaces require improvement. This study proposes the incorporation of the concept of sensory integration therapy into neighborhood open spaces for children to positively impact children's sensory development. The purpose of this study is to determine the effective facility factors of an occupational therapy room, translate them into a community facility design, clarify the categories and relative importance of each design attribute, and explore the design strategies of the children's facilities in neighborhood open spaces based on the sensory integration theory. This study investigates the importance of the sensory integration treatment level. The facilities in neighborhood open spaces for children can be considered systemic structures consisting of five partitioned units with different levels of importance among the synergistic components within each unit. These structures will enable children to experience sensory stimulation during daily outdoor play and will serve as preventive and therapeutic tools.

1. Introduction

Sensory integration refers to the integration of sensory information input by the brain from various parts of the human body that enables the body to respond to internal and external perceptions. Only through sensory integration, can different parts of the nervous system coordinate to produce a smooth interaction with the environment. A sensory integration disorder (SID) is diagnosed when the sensory system does not work properly [1]. According to Saeki et al. [2], SID can be divided into five categories: body movement coordination disorders, form and space perception disorders, body balance disorders, auditory-language disorders, and tactile defensiveness disorder. Sensory integration therapy provides sensory input from the vestibular system, muscles, joints, and skin, so that children can integrate these senses and generate adaptive responses [1]. Based on Ayres' theory, previous research has used appropriate motor responses combined with vestibular, proprioceptive, and tactile stimuli to improve children's academic performance, motor coordination, and language functions [3, 4]. SID occur mostly in children aged 3–12 years. The preventive stage of SID is before the age of three, the complementary stage of intervention is between the ages of three and six, and the optimal treatment stage is before the age of six [5].

At present, most sensory integration therapy is achieved through occupational therapy (OT), as proposed by Barton [6], and includes a rehabilitation process of applying purposeful and selected operational activities to evaluate, treat, and train patients who have lost the ability to take care of themselves and work in varying degrees due to physical, mental, and developmental dysfunctions or disabilities. The

rate of SID among Chinese children aged 3–12 years ranges from 10% to 30% [7, 8], indicating a large demand for OT among Chinese children. Easy and frequent access to OT is an important factor in maintaining treatment outcomes and improving the well-being of patients with SID [9]. At present, hospitals in many areas of China have specialized OT rooms for sensory integration treatment for children; however, these hospitals are located in big cities and are very limited in number. For example, only four hospitals in Guangzhou have licensed OT rooms. In addition, the construction of an OT room costs more than 3 million yuan [10]. In addition, the concentration of medical resources makes it difficult for children to receive treatments conveniently and frequently, which often delays and interrupts treatments.

The main way for children to recognize the world is through play, and children's facilities in neighborhood open spaces are important to support play. In these facilities, children imitate reality with specific game behaviors, gain initial personal experiences, learn to recognize the environment, and develop additional abilities. Children's facilities in neighborhood open spaces can help parents and children establish a good parent-child relationship, provide parents with opportunities to communicate with their children, and help solve children's psychological problems in a healthy and effective way. Proper children's outdoor recreation facilities may prevent isolation, improve the psychological health of children, and meet children's cognition and willpower training [11]. Previous studies have focused on the planning and design of such facilities in neighborhood open spaces from the perspective of public health or human factors engineering [12]. However, children's facilities in most urban communities in China only meet the children's minimal need for outdoor activities due to efforts to control costs, and the number and types of children's facilities are insufficient in some communities where the design of the facilities is simple or in need of maintenance or updating. Furthermore, the area devoted to children's facilities is often insufficient and not interesting to children [13]. If the important features of sensory integration therapy facilities in OT rooms can be used in the design of children's facilities in neighborhood open spaces, these spaces will not only meet children's daily needs of outdoor activity but also provide preventive and therapeutic sensory activities. However, few studies have integrated the designs used in OT rooms based on sensory integration theories when planning children's facilities. It is necessary to determine what factors currently used in OT rooms are effective, identify the potential organizational structure relationships of these effective factors, determine what specific design features can be used in the design of children's facilities in neighborhood open spaces, and clarify which design features should be the core components of the design.

The purpose of this study is to determine the factors of OT rooms that are effective and interpret them into features that can be included in children's facilities in neighborhood open spaces. This study also clarifies the category and relative importance of each design feature attribute and explores the design strategy of children's facilities in neighborhood open spaces based on sensory integration theory. A hybrid multiple attribute decision-making (MADM) model was used. The fuzzy Delphi method (FDM) was used to extract the valid facility factors of OT rooms using a questionnaire based on expert experience, followed by the exploratory factor analysis (EFA) technique, applied to determine the validity and categories of each design feature. A design framework of the sensory integration therapy factors interpreted as design features of children's facilities in neighborhood open spaces was constructed. The DEMATEL-based ANP (DANP) technique was used to determine the most dominant influence points of design of children's facilities in neighborhood open spaces. Based on the results of the FDM-EFA-DANP analysis, the design strategies and basic guidelines for integrating OT into children's facilities in neighborhood open spaces were explored. The main contribution of this study is to combine sensory integration therapy with children's facilities in neighborhood open spaces, which provides a new perspective and specific design strategies for their design. The rest of the paper is organized as follows: Section 2 provides a review of previous researches, Section 3 makes a description of the research methods and steps applied in this study, Section 4 provides a graphical description of the results and discusses the design strategies, and Section 5 summarizes the contribution of the full text and clarifies the research limitations and future research.

2. Literature Review

2.1. The Application of the Sensory Integration Theory. American psychologist A. Jean Ayres first proposed the theory of sensory integration. At the beginning of the 20th century, studies regarding various fields of sensory integration were conducted in China. As theories of sensory integration emerged, the practices of related fields were enriched. An increasing number of scholars have recognized the necessity of sensory integration in children's daily lives and have undergone training related to therapies of the sensory integration theory. Sensory integration training has also been adopted by occupational therapists [14], providing an important knowledge and skill base for practitioners worldwide. The sensory integration theory is one of the earliest theories of operational therapy and guides clinicians when treating sensory disorders that affect performance while providing much evidence to validate its structure.

Ayres' early publications have contributed to the development of this theory since the 1950s, and it is one of the most cited and used theories in operational therapy [15]. In the 1970s, many Western countries established children's sensory integration training centers to providers regarding children's sensory integration abilities, especially those of preschool children. Studies have shown that over 90% of kindergartens in Western developed countries include children's sensory integration training [16]. In the 1980s, sensory integration training was incorporated into kindergartens in Asia. Japan and South Korea were the first countries to offer sensory integration training courses in

kindergarten programs, followed by Hong Kong and Taiwan of China.

2.2. Related Research Regarding Sensory Integration. In recent years, there has been a growing interest in children's SID and sensory integration training leading to sensory integration rooms for children being included in medical institutions and kindergartens. However, some educators and parents of young children are underinformed regarding SID, and the number of therapists is limited, leading to numerous problems in the treatment of SID in some areas. Parents, schools, and medical institutions are not fully aware of the importance of sensory integration for children, especially in remote areas where medical conditions are not well developed. Due to the lack of professional guidance and of a unified, multidisciplinary approach, the number of treatment personnel is insufficient, and available treatments are incomplete. In addition, the price of sensory integration training rooms is too high for most areas, resulting in interrupted and ineffective treatments for children with fewer resources. Most research regarding sensory integration has been conducted in cities, with the inclusion of almost no rural children. Also, very few healthy children are included in studies regarding SID, as most studies only include patients with distinctive disorders. It is important to consider the requirements for sensory integration in healthy children when developing new facilities.

A review of previous research shows that sensory integration training is mainly conducted in sensory integration training rooms in rehabilitation schools, early education institutions, and hospitals. These settings are typically used for children with overt features of SID and families who take the initiative to treat their children with serious disorders. Sensory integration theory is being integrated into preschool education [17], and some parents and kindergarten teachers are conducting sensory integration therapy training [18].

Significant research has been conducted regarding the testing and refinement of the sensory integration assessment system [19], the influence of the sensory integration theory on sensory regulation and sleep habits in children with autism [20], the influence of the sensory integration theory on children with aphasia [21], the influence of the sensory integration theory on children with attention deficit hyperactivity disorder or learning disorders [22], and the relationship of the sensory integration theory with other medical fields such as neuromechanics [23]. The idea of sensory integration has been recognized in China for several years, though the development of sensory integration equipment has not been performed. Sensory integration activities have been included in an overall interactive game environment to enhance the effect of cultivating children's sensory integration abilities in Hong Kong, Macao, and Western countries, indicating that the research regarding sensory integration has made great progress, though many shortcomings remain.

2.3. Application of the Sensory Integration Theory in Related Fields. The sensory integration theory has been applied effectively in multiple related medical fields. The development of multisensory activities based on the theory of sensory integration has occurred for several years in some regions. These activities are mostly used in the fields of early childhood development and rehabilitation training of children, patients with reduced mobility, and elderly adults. Sensory integration activities enhance or soothe the sensory system using visual, auditory, tactile, and olfactory equipment. The new generation of sensory integration is based on the design of the sensory integration activity environment commonly used overseas and by the Hong Kong Occupational Therapists [24], whose concept differs from the sensory integration training environment used in China. In recent years, sensory integration activities have been applied to the training of children with dysthymia and in special physical education courses. The sensory integration theory can be helpful when designing children's educational products and be used to treat children with ADHD and other learning disorders. In addition, the theory of sensory integration can also be applied to the design of an interactive environment that can connect several sensory interactions into one activity, resulting in simultaneous multisensory and multisensational signals and feedback training.

3. Methods

Three quantitative analysis techniques from the field of operations research were used to form a hybrid MADM analysis model: the FDM, EFA, and DANP. The FDM was used to extract the effective facility factors that have an important influence in the OT room. This technology has been widely used in planning and evaluating research in related fields such as regional governance, community management, and landscape architecture [25, 26]. Compared with the traditional Delphi method, the advantages of the FDM include the reduction of the number of surveys, the complete expression of experts' opinions, rational experts that are solution-oriented, and an economical solution in terms of both time and cost. In this study, FDM technology was used to confirm the effectiveness of facility factors based on expert experience to establish a stable design framework [27]. EFA was used to determine the essential structure of multivariate observational variables and perform dimensionality reduction. In this study, the technology is used to classify the effective facility factors and summarize the relevant factors into the same dimension. Thirdly, in view of the correlation among the evaluation elements under the dimension, this study will apply DANP technology to clarify the influence network relationship and relative weight among the relevant elements. The application of this analysis technique relies on expert domain knowledge, through pairwise comparison between design criteria, to clarify the stakeholders' considerations of the relative importance of the criteria.

3.1. FDM Technique. In this study, the FDM was used to integrate expert opinions using two triangular fuzzy numbers [28] and to determine whether expert cognition shows a

consistent convergence effect using the gray zone verification method. The following three-step process was used:

Step ($F1$): The most conservative cognitive opinion and the most optimistic cognitive opinion provided by all experts to each factor i were statistically analyzed, and extreme values beyond two standard deviations were eliminated. Then, the minimum value $\mathbf{C}_L^i$, geometric mean value $\mathbf{C}_M^i$, and maximum value $\mathbf{C}_U^i$ in the remaining most conservative cognitive value and the minimum value $\mathbf{O}_L^i$, geometric mean value $\mathbf{O}_M^i$, and maximum value $\mathbf{O}_U^i$ in the most optimistic cognitive value were calculated.

Step ($F2$): Based on the calculation results of step ($F1$), the three-angle fuzzy number $\mathbf{C}^i = (\mathbf{C}_L^i, \mathbf{C}_M^i, \mathbf{C}_U^i)$ of the most conservative cognition and the three-angle fuzzy number $\mathbf{O}^i = (\mathbf{O}_L^i, \mathbf{O}_M^i, \mathbf{O}_U^i)$ of the most optimistic cognition for each factor i were calculated (Figure 1).

Step ($F3$): Determining whether the experts' opinions present a consistent convergence effect can be accomplished in the following ways:

(a) When no overlap between the two triangular fuzzy numbers is present ($\mathbf{C}_U^i \le \mathbf{O}_L^i$), the opinion interval value of each expert includes a consensus section that the opinion tends to be within; the consensus value $\mathbf{G}_U^i$ of this facility factor i is calculated usingfd1

$$\mathbf{G}_U^i = \frac{\mathbf{C}_M^i + \mathbf{O}_M^i}{2}. \tag{1}$$

(b) If there is an overlap between the two triangular fuzzy numbers ($\mathbf{C}_U^i > \mathbf{O}_L^i$) and the gray area ($\mathbf{Z}^i = \mathbf{C}_U^i - \mathbf{O}_L^i$) of the fuzzy relationship is smaller than the range $\mathbf{M}^i = \mathbf{O}_M^i - \mathbf{C}_M^i$ between the geometric mean of optimistic cognition and the geometric mean of conservative cognition for the facility factor by the expert, there is no consensus section for each expert's opinion interval value; however, the two with extreme opinions (the most conservative opinion of the optimistic cognition and the most optimistic opinion of the conservative cognition) were not significantly different from other opinions, leading to divergent opinions. Therefore, the consensus value of this facility factor i is equal to the fuzzy set obtained by the intersection (minimum) operation of the fuzzy relation of two triangular fuzzy numbers. The quantization score of the fuzzy set with the maximum membership value can be obtained using equations (2) and (3).

$$\mathbf{F}^i\left(x_j\right) = \left\{\int_x \left\{\min\left[\mathbf{C}^i\left(x_j\right), \mathbf{O}^i\left(x_j\right)\right]\right\} \mathrm{d}x\right\}, \tag{2}$$

$$\mathbf{G}^i = \left\{x_j \mid \max \mu_{F^i}\left(x_j\right)\right\}. \tag{3}$$

(c) If there is an overlap between the two triangular fuzzy numbers, ($\mathbf{C}_U^i > \mathbf{O}_L^i$), and the gray area ($\mathbf{Z}^i = \mathbf{C}_U^i - \mathbf{O}_L^i$) of the fuzzy relationship is larger than the range $\mathbf{M}^i = \mathbf{O}_M^i - \mathbf{C}_M^i$ between the geometric mean of optimistic cognition and the geometric mean of conservative cognition, there is no consensus section for each expert's opinion interval value, and the two extreme opinions differed too much from other opinions, leading to divergent opinions. Therefore, the questionnaires and FDM should be repeated until all the evaluation items reached convergence and a corresponding consensus value is obtained.

3.2. EFA Technique. The EFA is a dimension-reducing method of multivariate statistics that is used to identify the latent variables from manifest variables. A five-step process was used to complete the main procedure of the principal component analysis using EFA:

Step ($E1$): The correlation matrix $\mathbf{R}$ or variance-covariance matrix for the objects to be assessed was found.

Step ($E2$): The eigenvalues λ_k, $k = 1, 2, \ldots, m$ and eigenvectors $\beta_k = [\beta_{1k}, \ldots, \beta_{ik}, \ldots, \beta_{pk}]$ were identified for the assessment of the factor loading $a_{ik} = \sqrt{\lambda_k}\beta_{ik}$ and the number of factors m.

Step ($E3$): The eigenvalue ordering $\lambda_1 > \cdots > \lambda_k > \cdots > \lambda_m$, where $\lambda_m > 1$, was used to determine the number of common factors to be extracted by a predetermined criterion.

Step ($E4$): According to Kaiser [29], the varimax element was used to find the rotated factor loading matrix, which provides additional insights for the rotation of the factor axis.

Step ($E5$): The factor referring to the combination of manifest variables was named.

3.3. DANP Technique. The DANP technique was used to divide the relation matrix of the total influence from the first phase into influential weights (IWs, also termed global weights) of the criteria (vector w^g) using the basic concept ANP to calculate the local weights of each criterion (vector w_C^l) and dimensions (vector w_D^1) [30]. Recently, this analysis technology has been widely used in the field of public health and urban planning. It can effectively integrate the experts' domain knowledge from stakeholders, clarify the impact network relationship and relative weight between evaluation criteria, and provide decision-makers with dynamic impact viewpoint [31–33].

Step ($D1$): A direct influence relation matrix $\mathbf{A}$ was established based on questionnaire results. The assessment scores in the questionnaire included 0 (no influence), 1, 2, 3, and 4 (highly influential).

Step ($D2$): The average direct influence relation matrix $\mathbf{E}$ was established. The average scores of the H experts were determined using $e_{ij} = (1/H)\sum_{h=1}^{H} a_{i,j}^h$.

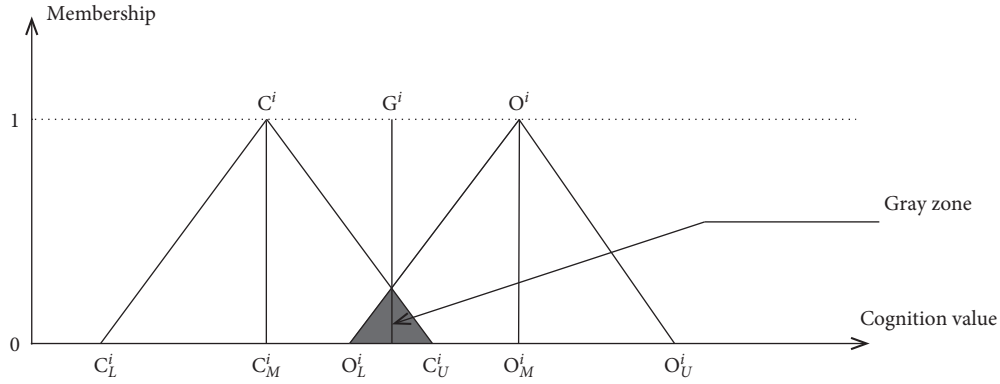


Figure 1: Triangular fuzzy numbers formed in the fuzzy Delphi method (FDM).

Step ($D3$): The consensus was examined using a threshold of the average gap ratio of 5%. Values less than 5% indicate a confidence level above 95%, which represents a stable system. When an unstable system is obtained, the first phase is implemented again to verify if the data were collected correctly and whether the number of experts was sufficient.

Step ($D4$): The normalized average direct influence relation matrix $\mathbf{D}$ is formulated to remove the units and orientations of values and to convert the values of the numbers to between 0 and 1.

Step ($D5$): The total influence relation matrix $\mathbf{T}$ was constructed using $\mathbf{T} = \mathbf{D}(\mathbf{I} - \mathbf{D})^{-1}$. This matrix served as the foundation for drawing the influential network relation map (INRM) in this study.

Step ($D6$): The total influence relation matrix of criteria $\mathbf{T}_C$ and dimensions $\mathbf{T}_D$ were constructed.

Step ($D7$): The unweighted supermatrix $\mathbf{W}^a$ was calculated.

Step ($D8$): The weighted supermatrix $\mathbf{W}$ was calculated.

Step ($D9$): The limit supermatrix was determined by multiplying the weighted supermatrix $\mathbf{W}$ by itself to achieve a convergent and stable matrix $\mathbf{W}^g$. The values in each column of the weighted supermatrix $\mathbf{W}$ represent the IWs.

4. Results and Discussion

4.1. Determining Which Factors in OT Rooms Are the Most Effective Using Expert Opinions. In this study, the patient room for surgery for four months was observed and recorded so as to collect the relevant information of patient-room facilities. Then, the qualitative data were obtained for this study from the collected photos, text memos, pictures, and videos in combination with relevant literature. The field investigation was carried out in the sensory integration treatment room of Guangzhou Rehabilitation Hospital. The relevant professional therapists were interviewed; we learnt about and experienced the common equipment and facilities in the training room. Finally, 27 facility factors were extracted from the treatment room (see supplementary file) based on the general inductive analysis method. Although continuous coding comparison has been carried out in the whole process of qualitative data induction and analysis, the effectiveness of these 27 convenience factors still needs to be tested by expert opinions. As described above, this study uses the FDM to integrate the domain knowledge of experts and selects the effective facility factors from the OT room on the basis of the consensus of experts.

The fuzzy Delphi questionnaire was completed by professional doctors and therapists who perform sensory integration therapy. All interviewees had at least five years of working experience, bachelor's degrees, and professional therapist qualification certificates. A total of 12 experts completed the questionnaire. After excluding invalid questionnaires, a total of 10 valid questionnaires were included in the analysis, including eight from female experts. One participating expert is a registered operative therapist, is master of operational therapy in Hong Kong, and has more than ten years of relevant practical and scientific research experience in this field. Based on the semantic discrimination of the five-point Likert scale, the threshold value of expert opinion was set at 6.667. When the validity of the facility factor was rated as less than 6.667, the facility factor was determined to be ineffective for the treatment of SID (Table 1).

4.2. Clarifying the Potential Hierarchical Relationship of Facility Factors in Sensory Integration Therapy. EFA was used to understand the content validity and potential hierarchical structure of the effective facility factors of the OT room, providing a basic reference for the construction of a framework for the design of children's facilities with sensory integration therapy in communities. From the perspective of community-based children's design, children's activity facilities should be viewed as a systematic whole, with well-organized subunits and synergistic interconnections among the components.

A questionnaire including items regarding the 20 facility factors was designed using a seven-point Likert-type scale and distributed to doctors and therapists involved in sensory integration therapy. Any unanswered questions were set as missing values. A total of 245 questionnaires were distributed, and 231 valid responses were received. The principal component analysis and the maximum variation method were used during data processing. Items with eigenvalues

Table 1: Validity screening of sensory integration therapy facility factors.

Facility factor	Conservative value $\mathbf{C}^i$		Optimistic value $\mathbf{O}^i$		Single value $\mathbf{a}^i$		Geometric mean $\mathbf{M}$		Calibration value		Consensus value
	Min	Max	Min	Max	Min	Max	$\mathbf{C}^i$	$\mathbf{O}^i$	$\mathbf{a}^i$	$\mathbf{M}_i$–$\mathbf{Z}_i$	$\mathbf{G}^i$
Soft building block house	3	7	6	10	4	9	5.161	7.191	6.295	1.030	**6.393**
Hot dog cushion	3	8	5	9	5	8	4.752	7.846	6.406	0.094	**6.401**
Sponge mat	3	9	7	9	4	9	5.568	8.395	6.227	0.827	7.578
Soft building blocks	2	10	7	10	6	10	5.725	9.068	7.306	0.342	7.978
Patterned soft cubes	1	7	5	7	2	7	3.322	6.450	4.648	1.128	**5.566**
Beanbag chair	1	8	6	8	4	8	3.651	6.959	6.007	1.308	**6.361**
Rotary drum	2	8	6	8	5	8	5.453	7.533	6.586	0.080	6.751
Folding ball pool	3	8	5	10	4	8	4.637	7.705	5.685	0.069	**6.337**
Wooden pull-out bed magic board	3	8	7	10	5	9	5.709	8.347	7.163	1.638	7.370
Back magic board wooden ladder	2	8	6	10	4	9	5.646	7.804	7.186	0.159	6.868
Back magic board rope net	3	8	7	10	5	8	5.473	8.243	7.075	1.769	7.330
Climbing wall	3	8	6	9	5	8	5.119	8.075	6.488	0.956	6.838
Balancing swing	2	8	6	10	5	8	5.241	7.566	6.342	0.325	6.724
Pumpkin swing	3	8	5	10	5	9	4.643	7.714	6.624	0.071	**6.341**
Square board swing	3	9	7	10	4	10	5.735	8.763	7.447	1.028	7.701
Sliding swing	3	9	7	9	4	10	6.209	8.415	7.468	0.206	7.673
Spinning top	3	8	6	10	4	9	5.301	8.198	6.533	0.897	6.898
Hanger rope	3	9	7	10	4	10	6.389	8.423	7.359	0.034	7.706
Padded-side bouncing bed	3	8	6	9	5	9	5.404	7.708	6.853	0.305	6.794
Padded slipway	3	8	6	9	5	8	4.743	7.708	6.722	0.966	6.688
Skateboard	4	9	7	9	5	9	6.106	8.463	6.822	0.357	7.672
Ball chair	4	8	7	10	7	9	5.876	8.347	7.683	1.471	7.388
Challenging facilities	5	9	7	9	6	10	6.418	8.469	7.390	0.050	7.725
Fruit-featured beanbag	4	8	7	10	6	10	6.439	8.241	7.590	0.802	7.443
Light-sensing game board	4	9	7	10	6	9	5.640	8.204	7.499	0.564	7.527
Spinning mirror ball	5	7	6	9	6	9	6.084	7.516	7.661	0.431	**6.623**
Virtual game	4	7	7	9	6	8	5.596	8.139	6.971	$\mathbf{C}_U^i \le \mathbf{O}_L^i$	6.868
Number of valid facility factors selected				**21**			Consensus threshold value				**6.667**

greater than 1 were selected, and those with a factor load less than 0.5 were rejected. The number of common factors was not limited. The 231 responses were found to be suitable for factor analysis (KMO = 0.980), with an internal consistency (Cronbach's α) of 0.977 among the 20 questionnaire items.

The 20 facility factors were categorized into five dimensions with a total cumulative explained variation of 78.723% (Table 2). As the factor loads of the square board swing and padded slipway were less than 0.5, these factors were excluded from the framework of sensory integration therapy for children. In this study, the following five dimensions were used: mobile experience series ($\mathbf{D}_1$), goal guidance series ($\mathbf{D}_2$), protection perception series ($\mathbf{D}_3$), limb coordination series ($\mathbf{D}_4$), and climbing series ($\mathbf{D}_5$). $\mathbf{D}_1$ includes five facility factors with an internal consistency of 0.918. $\mathbf{D}_2$ includes six facility factors with an internal consistency of 0.932. $\mathbf{D}_3$ includes three facility factors with an internal consistency of 0.864. $\mathbf{D}_4$ includes two facility factors with an internal consistency of 0.830. $\mathbf{D}_5$ includes two facility factors with an internal consistency of 0.827.

4.3. Clarifying the Influential Network Relation and IWs of the Design Criteria. After the clarification of the content validity and potential hierarchical structure of each facility factor, a focus group was formed including nine professional designers with similar educational backgrounds and experience in the design of open spaces in residential areas. A semistructured meeting was conducted regarding the 18 facility factors determined to be effective in sensory integration therapy, aiming to summarize and refine the design characteristics corresponding to each facility factor. This meeting resulted in a design framework for children's facilities that effectively integrates sensory integration therapy factors (Table 3). The design framework includes five facility dimensions, 18 facility factors, and corresponding facility design criteria. The construction of this hierarchical structure framework provides designers with a key design decision-making basis in practical engineering. The construction of this facility design framework enhances the designers' ability to control the overall morphological design of children's facilities in the community and guides the designers in the organization of the subunits of the children's facility system at the level of sensory integration therapy to stimulate creativity and enhance the synergy among the components of each unit.

In this study, DANP questionnaires were designed based on the construction of the design framework, which were used for the stakeholders from the circles of production,

Table 2: Factor analysis of facility factors in sensory integration therapy.

($N = 231$) Facility factors	Facility dimension				
	Mobile experience ($\mathbf{D}_1$)	Goal guidance ($\mathbf{D}_2$)	Protection perception ($\mathbf{D}_3$)	Limb coordination ($\mathbf{D}_4$)	Climbing ($\mathbf{D}_5$)
Balancing swing	0.698				
Soft building blocks	0.624				
Sliding swing	0.582				
Ball chair	0.581				
Virtual game	0.542				
Rotary drum		0.690			
Spinning top		0.617			
Challenging facility		0.594			
Skateboard		0.592			
Wooden pull-out bed magic board		0.567			
Climbing wall		0.541			
Sponge mat			0.687		
Hanger rope			0.676		
Light-sensing game board			0.507		
Square board swing					
Fruit-featured beanbag				0.703	
Padded-side bounce bed				0.548	
Padded slipway					
Back magic board wooden ladder					0.668
Back magic board rope net					0.541
Cronbach α	**0.918**	**0.932**	**0.864**	**0.830**	**0.827**
Cumulative Variance	**78.723%**				

Table 3: Design framework of children's facilities based on sensory integration theory.

Facility dimension	Facility factor	Design criteria
Mobile experience ($\mathbf{D}_1$)	Balancing swing	Suspended cushion with narrow and soft texture ($\mathbf{C}_{11}$)
	Soft building blocks	Soft building blocks with different color and shape ($\mathbf{C}_{12}$)
	Sliding swing	For children to sit on and slide when grasping a rope ($\mathbf{C}_{13}$)
	Ball chair	Gliding seat with axle wheel ($\mathbf{C}_{14}$)
	Virtual game	Video game with virtual scene ($\mathbf{C}_{15}$)
Goal guidance ($\mathbf{D}_2$)	Rotary drum	Hollow cylinder for children's multifaceted physical interaction ($\mathbf{C}_{21}$)
	Spinning top	Hard object for children to sit on or lie in and rotate ($\mathbf{C}_{22}$)
	Challenging facility	Information-guided interactive game ($\mathbf{C}_{23}$)
	Skateboard	Soft cushion with axle wheels for sliding ($\mathbf{C}_{24}$)
	Wooden pull-out bed magic board	Suspended cradle for children to lie down ($\mathbf{C}_{25}$)
	Climbing wall	Wall to guide the children for climbing ($\mathbf{C}_{26}$)
Protection perception ($\mathbf{D}_3$)	Sponge mat	Soft cushion to give children protection on the ground ($\mathbf{C}_{31}$)
	Hanger rope	Suspended hammock that can be moderately swung ($\mathbf{C}_{32}$)
	Light-sensing game board	Different colors and shapes of visual light-emitting components ($\mathbf{C}_{33}$)
Limb coordination ($\mathbf{D}_4$)	Fruit-featured beanbag	Graspable fruit-featured small soft object ($\mathbf{C}_{41}$)
	Padded-side bounce bed	Elastic net with soft padded enclosure for children to bounce vertically ($\mathbf{C}_{42}$)
Climbing ($\mathbf{D}_5$)	Back magic board wooden ladder	Ladder-like structure designed to fix the climbing rhythm and direction ($\mathbf{C}_{51}$)
	Back magic board rope net	Rope net structure designed with picking targets ($\mathbf{C}_{52}$)

government, and academics. Most of the experts who received the questionnaires had participated in the design of children's activity facilities in the open space of urban neighborhood. Eventually, a total of 15 expert questionnaires were collected, of which 13 were valid and passed the consistency test by the expert opinions. The results of data analysis are shown in Figure 2. Among the five design dimensions, goal guidance ($\mathbf{D}_2$) has the highest influence weight, followed by $\mathbf{D}_1$, $\mathbf{D}_3$, $\mathbf{D}_5$, and $\mathbf{D}_4$ in turn. In this design framework system, $\mathbf{D}_2$ is the most influential design dimension. In $\mathbf{D}_2$ dimension, information-guided interactive game ($\mathbf{C}_{23}$) is the design criterion which has the most influence weight, but soft discussion with axle wheels for sliding ($\mathbf{C}_{24}$) is the most influential one. The influence weight

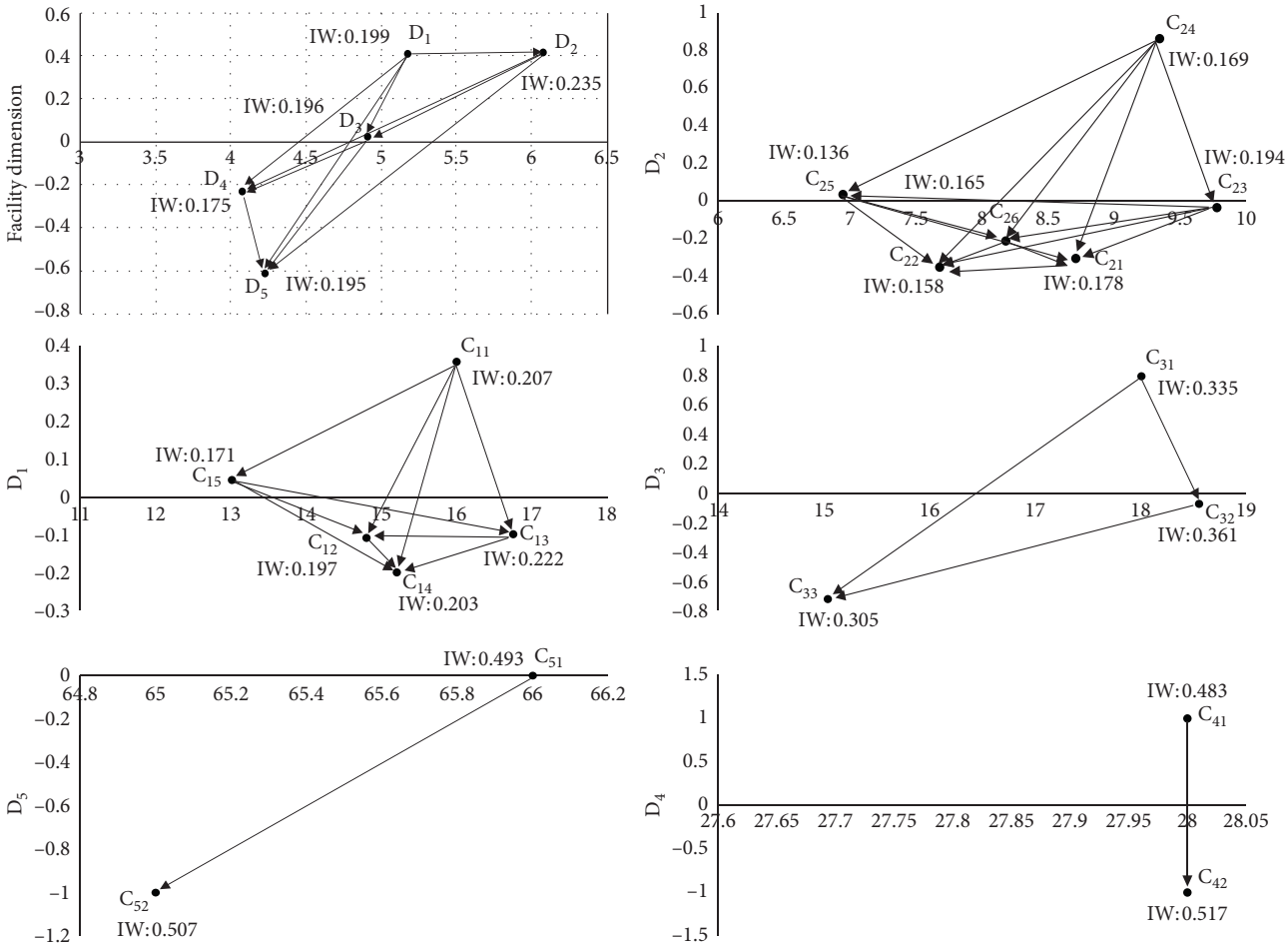


Figure 2: The influential network relation map (INRM) of total influence relationships.

and dominant influence of $\mathbf{D}_1$ are second only to $\mathbf{D}_2$. In the perspective of mobile experience ($\mathbf{D}_1$), children with dyskinesia were trained in physical motor coordination disorder, and the relative importance of $\mathbf{C}_1$ was the highest. $\mathbf{C}_1$ criterion refers to the children's activity facilities in the neighborhood open space, which can be used for children to grasp and slide, and suspended discussion with narrow and soft texture ($\mathbf{C}_{11}$) has the most dominant influence. $\mathbf{D}_5$ and $\mathbf{D}_4$ are the two design dimensions with the weakest influence weight and relatively weak dominant influence in this design framework.

In this study, the effective facility factors in the OT room were determined and the potential structural relationships of these facilities factors were explored. The main characteristics of facility design and the relative weight of design criteria were clarified, providing a key framework for the design of children's facilities that incorporate sensory integration therapy factors in neighborhood open spaces. Table 2 shows that the children's facilities in neighborhood open spaces are systemic structures that can be composed of five subunits. Each subunit serves a different purpose of sensory integration therapy, and the components of the facility within each unit remain synergistically related to each other. Based on the analysis results in Figure 2, this study integrates the influence weight and influence network relationship between design dimensions/criteria and explores the design strategy of children's facilities in open neighborhood space based on sensory integration theory. This design strategy can also be understood as a design guideline for children's facilities based on sensory integration theory or a decision-making proposal for launch of related design resources. In order to enhance the extensive sensory integration training, this study suggests that the design of children's facilities in neighborhood open space should focus on three aspects of goal guidance, mobile experience, and protection perception, so that the training facilities can be planned at different levels and the training contents of multiple mutual supporting can be output based on the law of sensory development of children with sensory disorders [34]. First of all, the most important design criterion in the goal guidance unit is to guide children to make corresponding behavior activities in the interactive game by receiving the output of information through hearing and identifying it through brain. In the process of voice perception, children's ability to integrate auditory and visual information is trained to improve their ability of voice understanding. It is very important in the early language development of children with SID [35]. In addition, this study suggests that designers should also consider children's sitting, lying, sliding, and hanging

facilities as the basis of the system based on the design of children's mobile experience in neighborhood open space. With the help of effective guidance, children can make many corresponding behavioral activities including sliding, climbing, and rotation.

Secondly, designers should attach importance to the application of virtual reality technology and the influence of video games in the facility system. Virtual reality technology can be regarded as an ideal tool for children to practice their behavior in the situation of role playing, and it also provides safe environment for rule learning and task repetition [36]. Dana et al. [37] have found that virtual reality game intervention lasting for 12 weeks could effectively improve the dynamic balance of children with developmental coordination disorder. In addition, in terms of protection perception facilities, designers should provide luminous game props of different colors and shapes and focus on the design of cushion facilities to protect children's safety when they are on the ground or hanging in the air. The environment of training space should cater to the SID children with various habits and movement tendencies. For the sake of safety, walls and floors need to be covered with soft materials, and the surface in the space is composed of soft objects, such as large cushions and flexible units of cushions of various shapes [38]. Children's activity space and all its elements should provide a variety of sensory inputs (vision, touch, and hearing) to provide children with the opportunities to explore their body consciousness and balance as much as possible. On the whole, designers should try to make children understand the narration of movement space, involving the single movement of children and the operation of triggering subsequent movement in the movement. In terms of body coordination and climbing training, designers need to make the supporting surface for climbing and the lower surface for landing clear and focus on supporting the body in different directions to promote children's sense of balance and functional training of vestibular system.

5. Conclusion

In this study, the effective facility factors of an OT room were identified, categorized, and incorporated into a design framework for children's facilities that include sensory integration therapy in neighborhood open spaces. Through a general inductive analysis, 27 facility factors were extracted and the FDM was used to select 20 factors of OT rooms that were thought to be effective by experts. To further confirm the content validity and potential structural relationship of these effective facility factors, EFA was used to determine five categories of facility factors, and 18 facility factors were extracted. Finally, the design characteristics of each facility factor were summarized to construct a framework of children's facility design with sensory integration therapy. Finally, in view of the correlation among the evaluation elements under the dimension, this study will apply DANP technology to clarify the influence network relationship and relative weight among the relevant elements.

The analysis results of the above research stages are the suggestions in the segmentation and supplement of the design criteria on the basis of the relevant current provisions of game facilities in *Code for Planning and Design of Urban Residential Areas* and *Code for Design of Parks* in China. A total of 5 facility dimensions and 18 facility factors will be taken as important design criteria and will be used by designers in the design strategy of existing children's facility transformation and new product development in a priority way based on the weight and actual implementation conditions. Based on such modes as literature review, questionnaire survey, expert interview, and field research, this study proves that it is feasible to apply indoor sensory integration training equipment and facilities to outdoor children's space, and it not only reduces the burden of medical staff and accompanying family members to a large extent but also helps children to receive sensory stimulation and improve their skills in daily life activities. The purpose of the study is to further solve the difficulties of children in the process of sensory integration therapy. Additionally, it is the first attempt to combine theory with practical medical device design in children's space facilities in community in China, which is undoubtedly the convenience for the majority of children. In the future study, the real-time feedback and evaluation will be followed up, and it should be noted that the design criteria of facilities with the same dimensions should be more closely combined or placed together to facilitate the use of the facilities and the enhancement of effectiveness. Furthermore, it should be noted that the influence of outdoor environmental factors should be considered and that how to use more weather resistant facility materials and form relationship while ensuring that the design criteria will not be affected by the outdoor environment factors will become one of the key points of further research in this paper. In addition, based on the results of this study, a number of different design schemes can be made in the future according to the specific design strategies of children's facilities in the display situation. The performance of the design scheme before construction can be analyzed from the perspective of sensory integration therapy by using the appropriate evaluation model.

Data Availability

The data used to support the findings of this study are included in the article.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

The authors are grateful to all the medical staff and designers who took part in this study. The authors would like to thank Lei Xiong, the research assistant, for all the hard work and efforts put in to make the questionnaire survey a success. The authors are also immeasurably grateful to the Alumni Association of the Environmental Art Design, School of Architecture and Applied Art, Guangzhou Academy of Fine Arts, for the support during the research. This study was funded by the art projects supported by National Social

Science Foundation, Project no. 18BG08. This work was also supported by the Graduate Teaching Reform Project of Guangzhou Academy of Fine Arts (Grant no. 6040119012).

Supplementary Materials

This is an introduction to the 27 facility factors mentioned in the paper: (1) soft building blockhouse (available to stimulate the body sense organ of the child and actively control the message accepted by the body sense organ), (2) hot dog cushion (the child lies down in the middle of the soft mat and experiences the feeling of being wrapped up with stimulation provided to its vestibule), (3) sponge mat (offer sufficient protection to the child during the training period), (4) soft building blocks (the game player combines the props in different forms based on the color and shapes; it is available for multiple players to play the game together, thereby guiding the players in the training activities for large muscle and strengthening the ability of the sense of balance), (5) patterned soft cubes (used for early childhood education; by breaking the traditional teaching theory for woodblocks, the soft blocks can protect the child during activity; moreover, the number on the blocks are available for the training of cognitive abilities of the child), (6) beanbag chair (the game player sits on the bean bag chair, thus stimulating the tactile sensory system of the player), (7) rotary drum (it is a balance game in which the player can grovel or stand on the roller, thereby promoting the overall development of the vestibular function and improving the inherent feeling and tactile stimulus of the vestibule in the meantime), (8) folding ball pool (it consists of heaped plastic small balls in different colors with moderate hardness, which is available to improve tactile sensitiveness or insufficiency: strengthening vestibular balance, physical coordination ability, and brain cadres function), (9) wooden pull-out bed magic board (experience the feeling of lying in the cradle in the infantile period, thus improving vestibular sensation and releasing muscle overstrain), (10) back magic board wooden ladder (available to provide diversified and challenging kinesthesia training for the child with sensory integration dysfunction of different level), (11) back magic board rope net (available to provide diversified and challenging kinesthesia training for the child with sensory integration dysfunction of different level), (12) climbing wall (it is a guided target remotely controlled by the tutor, which is available to provide diversified and challenging kinesthesia training for the child with sensory integration dysfunction of different level), (13) balancing swing (let the child sit on the lathy swing to strengthen the coordinating and balancing ability of the body), (14) pumpkin swing (improve the activity of the inherent sensory system of the vestibule and strengthen the tactile system; improve the body coordination of the child through acquiring a large amount of vestibule information, which is helpful for the inherent sensation input integration of the vestibule; it is also beneficial for the development of limb strength of the child), (15) square board swing (it is a coaxial rotary swing, on which the player can pose differently, including lying on its back or stomach, sitting or standing up, and so on; the tactile mat is available to provide more comfort tactile relaxation feeling to some players who suffer from tactile sensory disorder), (16) sliding swing (it works together with the dual-layer game platform and the ball pool and forms an interesting sensation integration game, thereby facilitating the learning of vestibular sensation, vision, proprioception, and tactile sensation), (17) spinning top (in sensation integration training, body large gyroscope, rotary cylinder, carousel-carrousel, rotary chair, and so on are all used to train vestibular balance), (18) hanger rope (it promotes the normalization of the vestibular system of the child by inputting appropriate sensation of rocking to the child), (19) padded-side bouncing bed (the child observes the surrounding environment and implements goal-response action when bouncing horizontally), (20) padded slipway (the child glides downward lying on its stomach with the skidding board placed on the skidding platform), (21) skateboard (the child glides lying on its stomach on the skidding board, thereby stimulating the vestibular system and the proprioceptive sense and promoting the double-side coordination of the body), (22) ball chair (it helps the child to build its vestibular sensory function, control the the sense of gravity, and develop the balance ability of the child; it has the best effect on the child with hyperactivity or poor coordination ability; in the meantime, it is also beneficial for the exercising of the child's waist and abdomen strength), (23) challenging facilities (improve the body coordination of the child through acquiring a large amount of vestibule information, which is helpful for the inherent sensation input integration of the vestibule; it is also beneficial for the development of limb strength of the child), (24) fruit-featured beanbag (improve the visual discrimination ability, fine operation ability, and reactivity through instructions), (25) light-sensing game board (train the abilities for cognition, language understanding, expression, and eye-hand coordination through visual sensation feelings of light), (26) spinning mirror ball (the light and shadow reflected by the rotating of the mirror ball are available to motivate the player with a weak sense of sight and help them to increase the attention to the surrounding environment, thereby improving the ability for spatial judgment and communication), and (27) virtual game (implement 3D interaction with different gaming methods and restore the real situation in a combination of AR technology augmented reality; make improvements in different aspects, including social communication, emotion, and feeling). (*Supplementary Materials*)

References

[1] A. J. Ayres, "Improving academic scores through sensory integration," *Journal of Learning Disabilities*, vol. 5, no. 6, pp. 338–343, 1972.

[2] K. Saeki, F. A. Clark, and S. P. Azen, "Performance of Japanese and Japanese-American children on the motor accuracy-revised and design copying tests of the southern California sensory integration tests," *American Journal of Occupational Therapy*, vol. 39, no. 2, pp. 103–109, 1985.

[3] A. J. Ayres, "Shirley stockmeyer: a physical therapist in an academic setting," in *The Body Senses and Perceptual Deficit: Proceedings of the Occupational Therapy Symposium on Somatosensory Aspects of Perceptual Deficit*Boston University, Boston, MA, USA, 1973.

[4] A. J. Ayres, "Learning disabilities and the vestibular system," *Journal of Learning Disabilities*, vol. 11, no. 1, pp. 30–41, 1978.

[5] R. Y. Li and H. L. Gong, "A experimentation study on constitution of pre -school autistic children of training in sensory integration," *Journal of Beijing Sport University*, vol. 28, no. 10, pp. 1344–1346, 2005.

[6] G. E. Barton, "Occupational therapy," *The Trained Nurse and Hospital Review*, vol. 54, pp. 138–140, 1915.

[7] H. P. Wang, *Sensory Integration Training for Special Children*, Peking University Press, Beijing, China, 2011.

[8] N. Wang, Q. X. Du, M. Y. Yin, and Y. L. Chang, "Investigation of sensory integration dysfunction of preschool hearing-impaired children," *Chinese Scientific Journal of Hearing and Speech Rehabilitation*, vol. 17, no. 5, pp. 363–365, 2019.

[9] American Occupational Therapy Association, "The scope of occupational therapy services for individuals with an autism spectrum disorder across the life course," *American Journal of Occupational Therapy*, vol. 64, pp. S125–S136, 2010.

[10] Y. Chen, C. Pan, and D. Long, "Training and practice of teachers for occupation therapy," *Chinese Journal of Rehabilitation Theory and Practice*, no. 8, pp. 791–793, 2013.

[11] M. B. Staempfli, "Reintroducing adventure into children's outdoor play environments," *Environment and Behavior*, vol. 41, no. 2, pp. 268–280, 2009.

[12] Y. A. Petrova and O. A. Sysoeva, "Design trends of children's playgrounds in modern urban environment," *IOP Conference Series: Materials Science and Engineering*, vol. 463, no. 3, Article ID 032075, 2018.

[13] Y. Lin, "Sensitive zone and sensitive period of children's growth—study on children's activity venues in the sensitive period of urban community growth," *Theory Monthly*, vol. 2012, no. 5, pp. 156–160, 2012.

[14] S. S. Roley, Z. Mailloux, H. Miller-Kuhaneck, and T. Glennon, "Understanding Ayres' sensory integration," *OT Practice*, vol. 12, no. 17, 2007.

[15] S. Mulligan, "Advances in sensory integration research," *Sensory Integration: Theory and Practice*, pp. 397–411, F. A. Davis, Philadelphia, PA, USA, 2002.

[16] M. H. Wang, L. h. Yang, and P. F. Wen, *The Research Status of Sensory Disorder and Training of Arch Sensory Integration*, School of Physical Education and Health, Zhaoqing University, Zhaoqing, China, 2017.

[17] C. M. Holland, E. I. Blanche, and B. L. Thompson, "Quantifying therapists' activities during sensory integration treatment for young children with autism," *Physical & Occupational Therapy in Pediatrics*, vol. 41, no. 3, pp. 284–299, 2020.

[18] R. Novianti and D. Satria, "Pelatihan terapi sensori integrasi bagi orang tua dan guru PAUD," *CARADDE: Jurnal Pengabdian Kepada Masyarakat*, vol. 3, no. 2, pp. 181–188, 2020.

[19] T. A. May-Benson, S. A. Schoen, A. Teasdale, and J. Koomar, "Inter-rater reliability of goal attainment scaling with children with sensory processing disorder," *The Open Journal of Occupational Therapy*, vol. 9, no. 1, pp. 1–13, 2021.

[20] R. Akarsu, B. Ozturk, and C. Karatekin, "Investigation of the effect of sensory integration therapy and foot reflexology applications on sensory modulation and sleep in a case with autism," *International Journal of Basic and Clinical Studies*, vol. 9, no. 2, pp. 114–121, 2021.

[21] T. Jegadeesan and P. Nagalakshmi, "Effect of sensory integration approach on children with dyspraxia," *Executive Editor*, vol. 11, no. 12, p. 88, 2020.

[22] S. Shahriarpour and N. Mirzakhani, "Sensory processing disorder and its effect on children skill and development in different neurodevelopmental disorders," *Journal of Clinical Physiotherapy Research*, vol. 6, no. 1, 2021.

[23] S. Honarvar, C. Kim, Y. Diaz-Mercado et al., "Unveiling the neuromechanical mechanisms underlying the synergistic interactions in human sensorimotor system," *Scientific Reports*, vol. 11, no. 1, pp. 1–16, 2021.

[24] S. C. Bodison and L. D. Parham, "Specific sensory techniques and sensory environmental modifications for children and youth with sensory integration difficulties: a systematic review," *American Journal of Occupational Therapy*, vol. 72, no. 1, pp. 1–11, 2018.

[25] Y. Wang and G.-T. Yeo, "Intermodal route selection for cargo transportation from Korea to central asia by adopting fuzzy Delphi and fuzzy ELECTRE I methods," *Maritime Policy & Management*, vol. 45, no. 1, pp. 3–18, 2018.

[26] Y. Liu and L. Li, "Mountainous city featured landscape planning based on GIS-AHP analytical method," *ISPRS International Journal of Geo-Information*, vol. 9, no. 4, p. 211, 2020.

[27] B.-W. Zhu, J.-R. Zhang, G.-H. Tzeng, S.-L. Huang, and L. Xiong, "Public open space development for elderly people by using the DANP-V model to establish continuous improvement strategies towards a sustainable and healthy aging society," *Sustainability*, vol. 9, no. 3, p. 420, 2017.

[28] T. B. Jeng, "Fuzzy assessment model for maturity of software organization in improving its staff's capability," Department of Information Management, School of Management, National Taiwan University of Science and Technology, Taipei, Taiwan, Master degree, 2001.

[29] H. F. Kaiser, "The varimax criterion for analytic rotation in factor analysis," *Psychometrika*, vol. 23, no. 3, pp. 187–200, 1958.

[30] Y.-P. O. Yang, H.-M. Shieh, and G.-H. Tzeng, "A VIKOR technique based on DEMATEL and ANP for information security risk control assessment," *Information Sciences*, vol. 232, pp. 482–500, 2013.

[31] I.-M. Feng, J.-H. Chen, B.-W. Zhu, and L. Xiong, "Assessment of and improvement strategies for the housing of healthy elderly: improving quality of life," *Sustainability*, vol. 10, no. 3, p. 722, 2018.

[32] Y. Liu, Y. Yang, Y. Liu, and G.-H. Tzeng, "Improving sustainable mobile health care promotion: a novel hybrid MCDM method," *Sustainability*, vol. 11, no. 3, p. 752, 2019.

[33] C.-H. Liu, "Strategies of managing coach driver job stress for sustainable coach tourism industry-the use of DANP-V model," *Sustainability*, vol. 12, no. 9, p. 3690, 2020.

[34] I. Obrusnikova and A. R. Cavalier, "Perceived barriers and facilitators of participation in after-school physical activity by children with autism spectrum disorders," *Journal of Developmental and Physical Disabilities*, vol. 23, no. 3, pp. 195–211, 2011.

[35] E. G. Smith and L. Bennetto, "Audiovisual speech integration and lipreading in autism," *Journal of Child Psychology and Psychiatry*, vol. 48, no. 8, pp. 813–821, 2007.

[36] S. Parsons and P. Mitchell, "The potential of virtual reality in social skills training for people with autistic spectrum disorders," *Journal of Intellectual Disability Research*, vol. 46, no. 5, pp. 430–443, 2002.

[37] A. Dana, A. Hamzeh Sabzi, and E. Christodoulides, "The effect of virtual reality exercises on dynamic balance of children with developmental coordination disorder," *Journal of Humanities Insights*, vol. 3, no. 03, pp. 123–128, 2019.

[38] Y. A. Yatmo, P. Atmodiwirjo, and K. D. Paramita, "Topological reading of movement connectivity in sensory integration space for autistic children," *Space and Culture*, vol. 20, no. 1, pp. 24–41, 2017.

Hindawi
Journal of Healthcare Engineering
Volume 2021, Article ID 9943402, 22 pages
https://doi.org/10.1155/2021/9943402

Research Article

Blockchain-Based Reversible Data Hiding for Securing Medical Images

Ji-Hwei Horng,[1] **Ching-Chun Chang**,[2] **Guan-Long Li**,[3] **Wai-Kong Lee**,[4] **and Seong Oun Hwang**[4]

[1]*Department of Electronic Engineering, National Quemoy University, Kinmen 89250, Taiwan*
[2]*Department of Electronic Engineering, Tsinghua University, Beijing 100084, China*
[3]*Department of Information Engineering and Computer Science, Feng Chia University, Taichung 40724, Taiwan*
[4]*Department of Computer Engineering, Gachon University, Republic of Korea*

Correspondence should be addressed to Seong Oun Hwang; bardic@naver.com

Received 23 March 2021; Revised 7 April 2021; Accepted 24 April 2021; Published 8 May 2021

Academic Editor: Chi-Hua Chen

Medical images carry a lot of important information for making a medical diagnosis. Since the medical images need to be communicated frequently to allow timely and accurate diagnosis, it has become a target for malicious attacks. Hence, medical images are protected through encryption algorithms. Recently, reversible data hiding on the encrypted images (RDHEI) schemes are employed to embed private information into the medical images. This allows effective and secure communication, wherein the privately embedded information (e.g., medical records and personal information) is very useful to the medical diagnosis. However, existing RDHEI schemes still suffer from low embedding capacity, which limits their applicability. Besides, such solution still lacks a good mechanism to ensure its integrity and traceability. To resolve these issues, a novel approach based on image block-wise encryption and histogram shifting is proposed to provide more embedding capacity in the encrypted images. The embedding rate is over 0.8 bpp for typical medical images. On top of that, a blockchain-based system for RDHEI is proposed to resolve the traceability. The private information is stored on the blockchain together with the hash value of the original medical image. This allows traceability of all the medical images communicated over the proposed blockchain network.

1. Introduction

The medical industry has been moving toward the digitized era, wherein a large amount of medical information is stored in a digital form and communicated digitally [1]. This helps in streamlining the acquisition, processing, and management of medical information and, at the same time, improving the efficiency in the medical industry. From all medical information, medical health record (MHR) is the most vital part, as it keeps all the important and private information regarding the patients and their diagnosis. MHR usually includes the patients' information like personal data, medical history, medical images, diagnosis reports, etc. Due to the booming of telemedicine technology, the exchange of medical images is becoming an important trend in the medical industry [2]. Medical image is one of the critical pieces of information in MHR that reveals a lot of sensitive information, which needs to be protected against malicious intrusion.

Traditional cryptographic algorithms can be used in protecting MHR (including medical images) effectively. For instance, Alam et al. [3] had proposed a framework for provisioning healthcare data, wherein Elliptic Curve Cryptography (ECC) and Advanced Encryption Standard (AES) are being used to encrypt the medical data. Recently, there is an increasing trend in employing reversible data hiding (RDH) techniques to embed sensitive information into medical images. RHD schemes have found some applications to the medical images, which are reported by Yang et al. [4, 5]. This shows that RDH is a promising candidate in securing medical images with the additional ability to embed sensitive information, which is not found in traditional cryptographic algorithms.

RDH is a technique that allows perfect recovery of the original plain image and the embedded data. The embedded data is usually some important or sensitive data to be hidden within the plain image. For instance, in the context of medical images, this can be the patient's personal data, diagnosis report, and summary of past medical records related to the medical images. The study [6] presented a flexible RDH scheme based on quad-tree and pixel value ordering (PVO) to exploit the similarity between neighboring pixels to hide more data. The paper [7] proposed a method that does not take the difference of the neighboring pixels in an image. Instead, they rearrange the columns/rows of the image in a way that improves the smooth regions, resulting in an increase in embedding capacity.

Recently, more attention turned toward RDH on the encrypted images (RDHEI). This is to ensure that the security and privacy of the transmitted image are being protected. Although these proposed RDHEI schemes are advanced and able to securely communicate the images and hidden data, they are still vulnerable to certain malicious activities. In particular, one can still modify the pixel values in an encrypted medical image for a malicious purpose, which should be detected. In other words, the integrity of RDHEI schemes needs to be checked, and this is still an open research problem to date. Moreover, the records of medical data communication are not properly protected, which makes the tracing and tracking of such communication a challenging issue to be resolved.

Blockchain is an emerging technology that aims to replace or compensate for the traditional centralized systems. It can be regarded as a Distributed Ledger Technology (DLT), wherein the transaction and storage of data are performed in a distributed manner. In such a way, even though there is no trust among all communicating parties, they can still trust the blockchain network. One of the key features of the blockchain is that all the data stored in the chain are connected through a cryptographic hash, which is very costly (or almost impossible) to tamper with.

Blockchain was recently applied to the healthcare system to improve security. For instance, [8] proposed Guardhealth, a decentralized blockchain system for privacy preserving and data sharing in medical industry. In this paper, our aim is to apply blockchain to improve integrity and traceability of the RDHEI scheme for protecting medical images.

An RDHEI scheme that vacates room after encryption of the cover image is proposed. The cover image is first divided into small blocks and then permuted by a permutation key. Then, each block is stream ciphered by an encryption key. The data hider vacates the embedding room and hides secret data using the histogram shifting technique. Our scheme can achieve a high embedding capacity compared to the state-of-the-art scheme by Zhang et al. [9]. A blockchain-based system is proposed to provide additional features to the proposed RDHEI scheme, which can be very useful in securing medical images. The contribution of this paper is summarized as follows:

(1) This paper proposed an RDHEI scheme to embed private information into the medical images. The proposed scheme employed stream cipher to vacate more space for data embedding capacity compared to the state-of-the-art scheme proposed by Zhang et al. [9].

(2) A novel blockchain-based RDHEI system was proposed, wherein the hash value is generated from the output of RDHEI and stored on the blockchain. This ensures that any attempt to tamper with the medical images can be detected easily.

The proposed system allows the user to exchange the ciphered steganography medical image (CSMI) securely with other members within the blockchain network. This is an important contribution to the medical industry as important medical information can be communicated frequently without worrying about security issues.

2. Background

2.1. Overview of Blockchain Technology. Blockchain is a distributed ledger system that is developed to work in an environment wherein the participating parties do not trust each other. In contrast to the traditional server-centric model, blockchain requires each participating node to store a copy of the ledger which records all the transactional details. Since the ledger is kept locally by all participating nodes, they can perform an audit on the transactions locally. With this feature, even though there is no trust among the participating nodes, one can still trust the consensus achieved through blockchain.

Another unique feature offered by blockchain is the introduction of a cryptographic hash to link up all the transaction records. Referring to Figure 1, all valid transactions are grouped into a block within a fixed time interval. A new hash value is generated based on these transaction details together with the hash of the previous block. Followed by this is the consensus process (e.g., Proof of Work (PoW)) to approve the transaction. During the consensus process, only the node that successfully solved the given difficult puzzle can add this block into the existing blockchain. The generation of hash in each transaction block is linked with the hash of the previous block. To modify one of the transaction records in the blockchain, one must generate a lot of valid blocks through the consensus process and overwrite the subsequent blocks. Since the PoW is a time-consuming process, it is very difficult to generate a lot of new blocks in a short time; this makes blockchain an immutable solution to many applications.

The consensus process through the PoW process in the blockchain is time consuming, which makes the transaction slow (e.g., 10 minutes for Bitcoin). Another alternative is to employ consortium blockchain, wherein a list of trusted members is predefined. In such case, a lightweight consensus process like Practical Byzantine Fault Tolerance (PBZT) [10] can be used. This allows timelier communication between the nodes within the same consortium blockchain, without sacrificing the key security features (i.e., auditability, traceability, and integrity) in the blockchain.

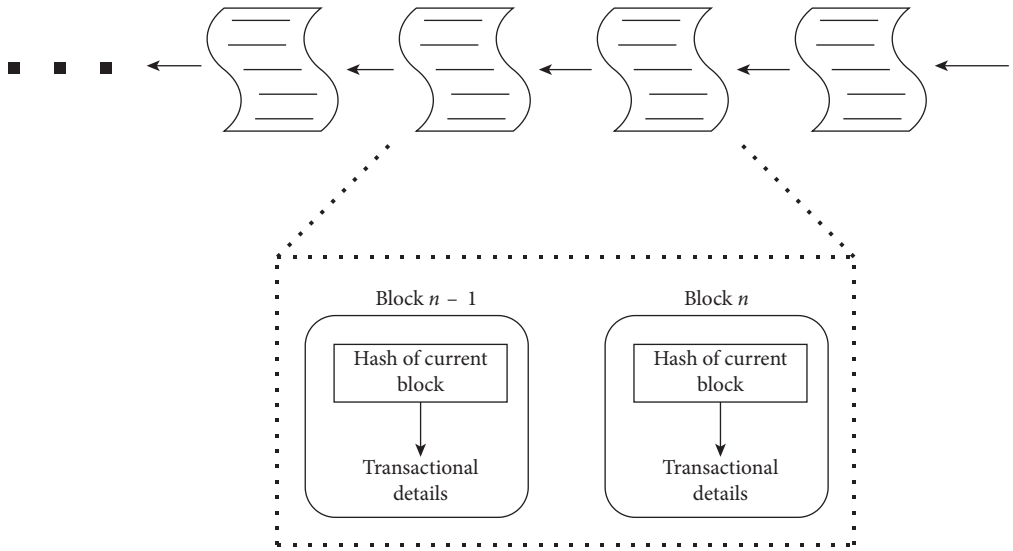


Figure 1: Blockchain data structure.

2.2. Overview of RDHEI. Many image steganography techniques have been proposed in the past in order to hide secret data in a way that the stego image appears with no difference from the cover image. To achieve this goal, only a slight modification of the pixel value is allowed, which severely limits the hiding capacity of a cover image. If we further constrain that the cover image should be perfectly recoverable [11] after the extraction of secret data, the embedding capacity is reduced to a very low level. Many reversible data hiding methods have been proposed to solve this challenging problem, which can be classified into three categories including lossless compression, histogram shifting, and pixel value differencing.

The lossless compression [12, 13] exploits redundancy of the cover image to embed data, in which the smooth images can hide more data than the complex images. Histogram shifting [14, 15] was proposed later on, wherein the data is embedded by shifting one side of a peak value in the histogram outward, leveraging the expanded extra space. Pixel value differencing is a more advanced technique [16, 17] that computes the differences of the pixel value in an image block and hides data by expanding the difference, in a way similar to the histogram shifting method. All these methods can embed secret data and recover the cover image after data extraction. The emergence of cloud services provides another possible framework, in which the data hiding operation is performed by a third party. RDHEI scheme can be used under this framework, wherein the content owner encrypts the cover image and sends it to the data hider to embed the secret data and forward it to the receiver. At the receiver end, the secret data can be extracted using the embedding key and the cover image can be reconstructed using the encryption key. In case the receiver is not authorized to acquire the secret data, he can still obtain the slightly distorted cover image using the encryption key only.

RDHEI schemes proposed in the past can be classified into two categories: vacating room after encryption (VRAE) and vacating room before encryption (VRBE). VRAE technique vacates room for embedding after the encryption. In 2008, Puech et al. [18] proposed an RDHEI scheme that applies bit substitution to embed secret data into an encrypted image. Later on, Zhang [19] proposed a separable RDH scheme, in which the receiver can extract secret data by using an encryption key and recover the cover image using a data hiding key. When the receiver has both keys, he can obtain both secret data and recover the image. VRBE is first proposed by Ma et al. [20] in 2013. By vacating embedding room before encryption, they claim that their scheme can achieve real reversibility; i.e., the data extraction and image recovery are free of any error. In 2016, Cao et al. [21] proposed a VRBE scheme based on sparse coding, in which the leading residual errors and learned dictionaries, together with the secret data, are embedded into the encrypted image. Malik et al. proposed an RDHEI scheme [22] that creates spare space using the prediction-error estimation method. The data is embedded into the most significant bits of the encrypted image. However, an additional location map is required to mark the nonembeddable pixels. RDHEI schemes can also be used to embed sensitive information into medical images. For instance, [4] had proposed a scheme to embed information into medical images through RDH and homomorphic encryption. Later on, [5] improved the scheme by proposing a novel paradigm (encrypt-then-RDH) to enhance the security of RDH over medical images.

Recently, Zhang et al. [9] proposed an interesting work along this research direction, in which they claim to achieve high embedding capacity. Referring to Figure 2, Zhang et al. [9] use additive homomorphic encryption and block permutation to encrypt the cover. The encryption and decryption formulas for homomorphic encryption are given by equations (1) and (2), respectively.

$$C = E(P, K) = (P + K) \bmod N, \tag{1}$$

$$P = D(C, K) = (C - K) \bmod N, \tag{2}$$

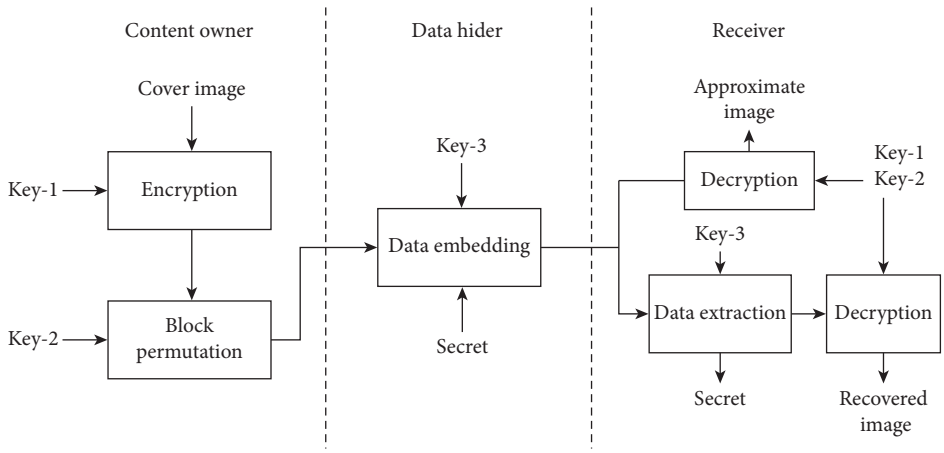


Figure 2: The RDHEI scheme proposed by Zhang et al. [9].

where $N = 256$, K is Key-1, and $E(\cdot)$ and $D(\cdot)$ are encryption and decryption algorithms, respectively. For each subblock, all pixels are encrypted using the same key to preserve the correlation between them. A block permutation method based on the chaotic algorithm is applied to strengthen security. An additional key sequence Key-2 is required to execute this operation. In the data embedding phase, the pixels in a subblock are classified into three sets according to their locations. For each set of pixels, a particular prediction rule is applied to measure prediction errors. Then, the histogram expansion and shifting process is used to embed secret data into the prediction errors. By controlling the embedding interval created by histogram expansion, the embedding capacity is adjustable. An embedding key Key-3 is required to store the information about embedding intervals.

This data hiding scheme suffers from some problems. First, homomorphic and block permutation encryptions of the cover image are not secure enough. The overall effect of image encryption is shifting of pixel gray level and changing of block spatial location. Secondly, the embedding key is not completely random, because it stores the parameters for data hiding, which is actually metadata for secret message extraction. The public key system could not help to share the encryption key under this scheme. The third and the most critical one is the problem of overflow. To hide the more secret data, the wider embedding interval should be expanded, and larger prediction errors exist. The encrypted pixel values may overflow the formal range of 0 to 255. To deal with the problems above, we propose a novel data hiding scheme in the following section.

3. The Proposed Reversible Data Hiding Scheme for Medical Images

In this section, we propose an RDHEI scheme based on stream ciphering and histogram shifting techniques. Our scheme includes three phases, namely, (1) the image encryption, (2) the data hiding, and (3) the data extraction and image decryption as shown in Figure 3. Three secret keys are involved in the proposed RDHEI scheme. *Key-I* and *Key-II* are the image encryption keys and *Key-III* is the data hiding key. The length of each key could be 64, 128, or 256 bits depending on the required security level. In the image encryption phase, the gray level cover image occupied by the content owner is firstly divided into mutually exclusive blocks of size 3×3. Block permutation with encryption key *Key-I* is applied to crumble the spatial relationship between image blocks. Then, stream ciphering with encryption key *Key-II* is leveraged to encrypt the image. In the data hiding phase, the data hider embeds secret data and metadata using the histogram shifting technique with *Key-III*. At the receiver end, the metadata is extracted first. With histogram rough restoration, the approximate cover image can be recovered using encryption keys *Key-I* and *Key-II*. With histogram fine restoration, the cover image can be perfectly recovered. For an authenticated receiver, a secret message can be deciphered from secret data using *Key-III*.

3.1. Cover Image Encryption Phase. The cover image encryption phase includes two steps: block permutation and stream ciphering. The cover image is divided into mutually exclusive blocks of size 3×3 first. Then, *Key-I* is applied to generate a random sequence of length equal to the number of total blocks in the image and all blocks are reordered according to the random sequence. An illustrative tiny image of 2×2 blocks is shown in Figure 4, where (a) is the original tiny image and (b) is the permuted image according to the random sequence 2; 3; 1; 0. After block permutation, we apply *Key-II* to generate a secret stream. Then, each image block is encrypted by

$$C = E(P, K) = K \oplus P, \tag{3}$$

where K is an 8-bit stream code truncated from the secret stream, P and C represent the input block and the cipher block, respectively, and the encryption function $E(\bullet)$ is an Exclusive-OR operation denoted by "$\oplus$." All blocks are sequentially encrypted with a distinct 8-bit code segment for each. Pixels within the same block are encrypted with the same code segment; therefore, their correlation is preserved.

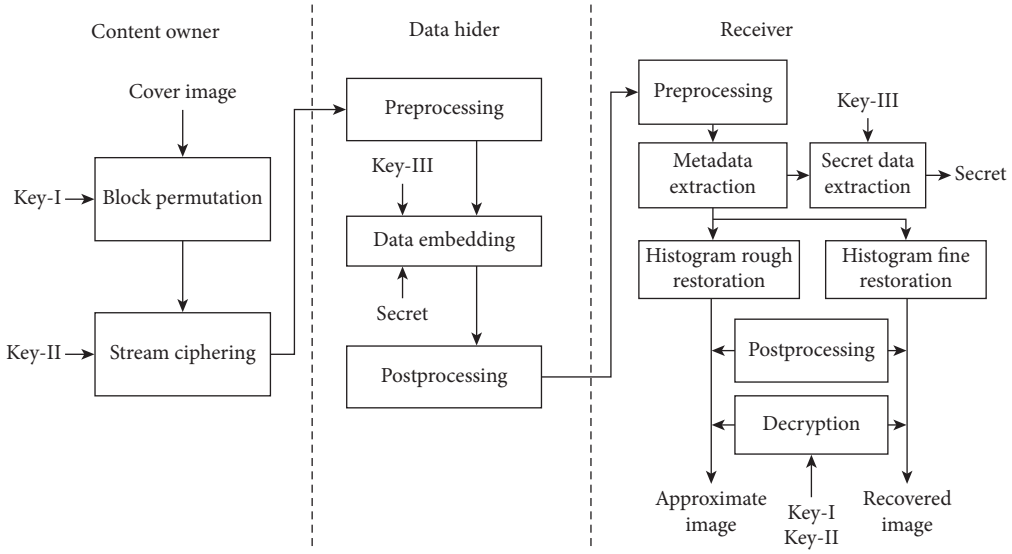


Figure 3: The proposed RDHEI scheme.

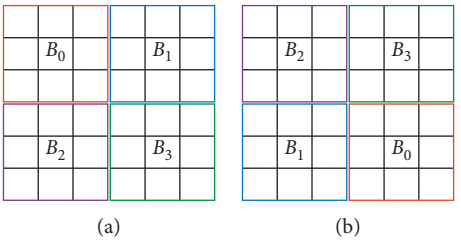


Figure 4: An illustrative example of block permutation.

This is an important feature for the following data embedding phase.

An illustrative example of the cover image encryption is shown in Figure 5. To simplify the representation, we use two blocks only to demonstrate the encryption operation. The two image blocks in (a) are permuted using the sequence generated by *Key-I* to obtain (b). Then, the two blocks are stream ciphered by two code segments generated by *Key-II*, individually. We take the first pixel of each block as an example. The binary codes of the first code segment and first-pixel value are $62 = (00111110)_2$ and $104 = (01101000)_2$, respectively. By Exclusive-OR operation, it results in $(01010110)_2 = 86$. The code segment and first-pixel value of the second block are $193 = (11000001)_2$ and $106 = (01101010)_2$, respectively. It results in $(10101011)_2 = 171$. Other pixels are calculated in the same way to get the result (c).

3.2. Data Embedding Phase. The data embedding process for the data hider is as shown in Figure 3. The preprocessing for each block in the encrypted image is an internal Exclusive-OR operation given by

$$g_i = c_i \oplus c_c, \quad i = 1, 2, \ldots, 8, \tag{4}$$

where c_c is the center pixel value of a block, c_i is a pixel value around the center pixel as shown in Figure 6(a), and g_i is its corresponding output pixel value as shown in Figure 6(b).

By simple derivation as given by equation (5), we can find that g_i is equal to $p_c \oplus p_i$. For a smooth block, all pixel values are close to each other. Therefore, there are a lot of zeros in the output image.

$$g_i = c_i \oplus c_c = (p_i \oplus K) \oplus (p_c \oplus K) = p_i \oplus p_c, \quad i = 1, 2, \ldots, 8. \tag{5}$$

After preprocessing, we apply the histogram shifting technique to embed secret data. According to the required embedding capacity, a threshold of gray level g_{th} is determined and the histogram is expanded by

$$g_i' = \begin{cases} g_i \times 2, & g_i < g_{th}, \\ g_i + g_{th}, & g_i \geq g_{th} \text{ and } g_i < 128, \\ g_i, & g_i \geq 128 + g_{th}. \end{cases} \tag{6}$$

A vacating band of width g_{th} after the gray level 128 is created. Then, the first g_{th} gray levels are shifted to even values without disrupting their order and the gray levels are shifted outward between g_{th} and 128. The threshold value g_{th}, recorded with six bits, and the information in the vacating band are encoded and stored as metadata.

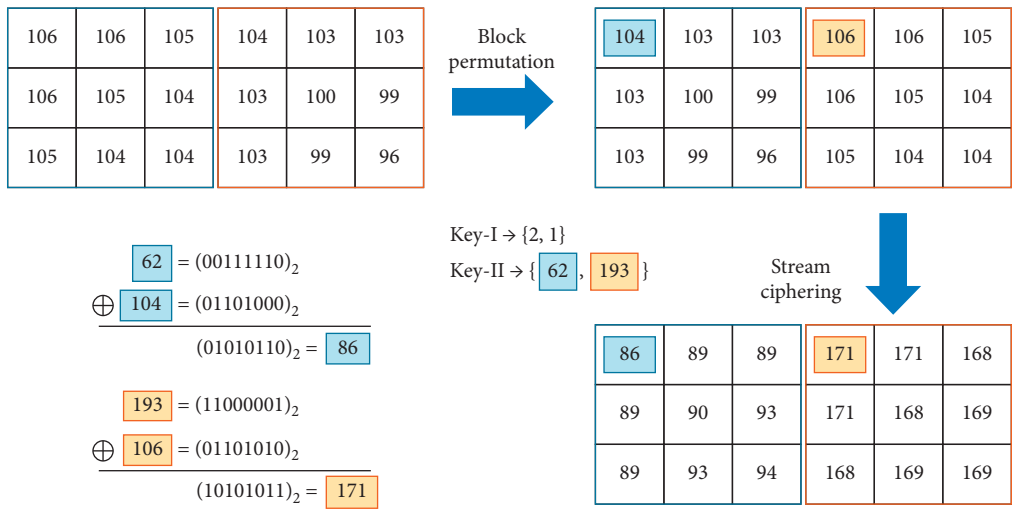


Figure 5: An illustrative example of block permutation and stream ciphering.

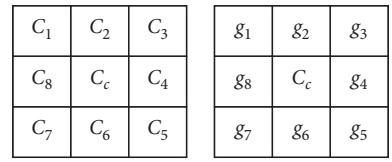


Figure 6: Pixel labeling of an encrypted block and its corresponding output after preprocessing. (a) Pixel labeling of a block and (b) output of preprocessing.

The binary secret stream is encrypted using *Key-III* with the conventional stream ciphering technique into an encrypted bitstream $S = \{b_1, b_2, b_3, \ldots, b_N\}$. For all embeddable pixels with $g_i' < 2 \times g_{th}$, we collect them in ascending order of gray level and in a raster scan order for pixels of the same gray level. Then, we consecutively embed secret bit by

$$\widehat{g}_i = g_i' + b_j. \tag{7}$$

After embedding, postprocessing with the same operation as preprocessing is executed again to reverse the effect. For each block,

$$\widehat{g}_i' = \widehat{g}_i \oplus \widehat{g}_c, \quad i = 1, 2, \ldots, 8, \tag{8}$$

where the definitions of $\widehat{g}_c$ and $\widehat{g}_i$ can refer to Figure 5.

To achieve reversibility of the proposed scheme, the histogram band in the range $128 \le g_i < 128 + g_{th}$ erased by the vacating process of histogram shifting should be recorded. Since only a very small amount of data is required to record, we design a simple and straightforward coding rule. Six bits are used to record g_{th}; twelve bits are used to record the total number of coded pixels; 3 bytes for each pixel are used to record its gray level g_i and coordinates (x, y) in the image. The image encryption and data embedding process is summarized as follows.

Input: cover image, binary secret stream, encryption key: *Key-I, Key-II,* data hiding key: *Key-III*

Output: encrypted image with data embedded

Content owner:

Step 1: encrypt the image by block permutation with *Key-I* and stream ciphering with *Key-II*

Data hider:

Step 2: preprocess the encrypted image according to equation (4)

Step 3: apply histogram shifting to vacate embedding area by equation (6)

Step 4: encode the erased histogram band into metadata

Step 5: collect the embeddable pixels in ascending order of gray level and raster scan order

Step 6: apply stream ciphering to the secret binary stream with *Key-III* into an encrypted data stream

Step 7: embed metadata first and then encrypted data stream by equation (7)

Step 8: postprocess the embedded image by equation (8) to produce the output image

The tiny image example given in Figure 5(c) is applied to illustrate the data embedding process as shown in Figure 7. The two encrypted image blocks are shown in Figure 7(a). The preprocessed blocks are shown in Figure 6(b), where we apply the threshold gray level $g_{th} = 4$ and the embeddable

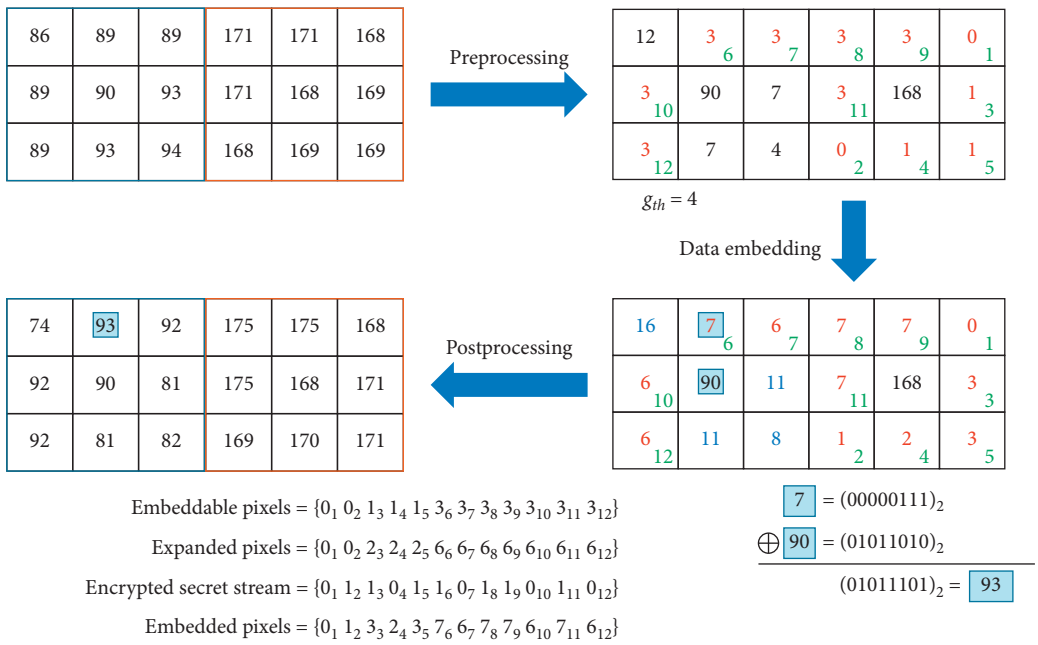


Figure 7: An illustrative example of the proposed image encryption and data embedding process.

pixels are denoted by red characters. The embedding order of all embeddable pixels is labeled with green Arabic numerals. Except for the center pixel of each block, all pixel values are expanded according to equation (6). Then, the encrypted secret stream is embedded in the predefined order by equation (7) to obtain Figure 7(c). Finally, the two blocks are postprocessed to get Figure 7(d). An example of pixel processing is also given in the figure.

3.3. Data Extraction and Image Recovery Phase. The data extraction and cover image recovery are executed in the reverse order of encryption and embedding. The overall process is summarized as follows (Figure 8).

The data extraction and image recovery process.

Input: encrypted image with data embedded, encryption key: *Key-I*, *Key-II*, data hiding key: *Key-III*.

Output: binary secret stream, cover image.

Receiver:

Step 1: preprocess each block of the input image by

$$\widehat{g}_i = \widehat{g}_i' \oplus \widehat{g}_c, \quad i = 1, 2, \ldots, 8, \tag{9}$$

where the notation definitions of $\widehat{g}_c$ and $\widehat{g}_i'$ are the same as equation (9).

Step 2: collect the first six pixels belonging to $\{0, 1\}$ in the raster scan order and convert them to get g_{th}.

Step 3: according to g_{th}, collect all embeddable pixels in the ascending order of $\widehat{g}_i/2$ and raster scan order for the pixels of the same value, where $\lfloor \cdot \rfloor$ is the floor function.

Step 4: consecutively extract the metadata and encrypted secret stream by even-odd decision.

Step 5: decipher the binary secret stream using *Key-III*.

Step 6: backward shift the pixel values and recover the vacating band according to the metadata. The backward shifting is given by

$$g_i = \begin{cases} \lfloor \dfrac{\widehat{g}_i}{2} \rfloor, & \widehat{g}_i < 2 \times g_{th}, \\ \widehat{g}_i - g_{th}, & \widehat{g}_i \geq 2 \times g_{th} \text{ and } \widehat{g}_i < 128 + g_{th}, \\ \widehat{g}_i, & \widehat{g}_i \geq 128 + g_{th}. \end{cases} \tag{10}$$

where $\lfloor \cdot \rfloor$ is the floor function.

Step 7: postprocess each block by

$$c_i = g_i \oplus c_c, \quad i = 1, 2, \ldots, 8, \tag{11}$$

where the notation definitions of c_c and g_i are the same as equation (4).

Step 8: decipher the histogram recovered image using *Key-II*.

Step 9: apply backward permutation to obtain a cover image using *Key-I*.

The data embedded tiny illustrative image in Figure 7(d) is applied as an input of the data extraction and image recovery process as shown in Figure 7(a). The two blocks are preprocessed to get Figure 7(b). Then, according to $g_{th} = 4$, the secret data stream can be extracted, and the histogram can be recovered as shown in Figure 7(c). Next, we postprocess the two blocks to get Figure 7(d). Finally, we use the encryption key to decipher the image blocks. Note that we do not show the extraction of metadata and recovery of vacated histogram band as in an actual application.

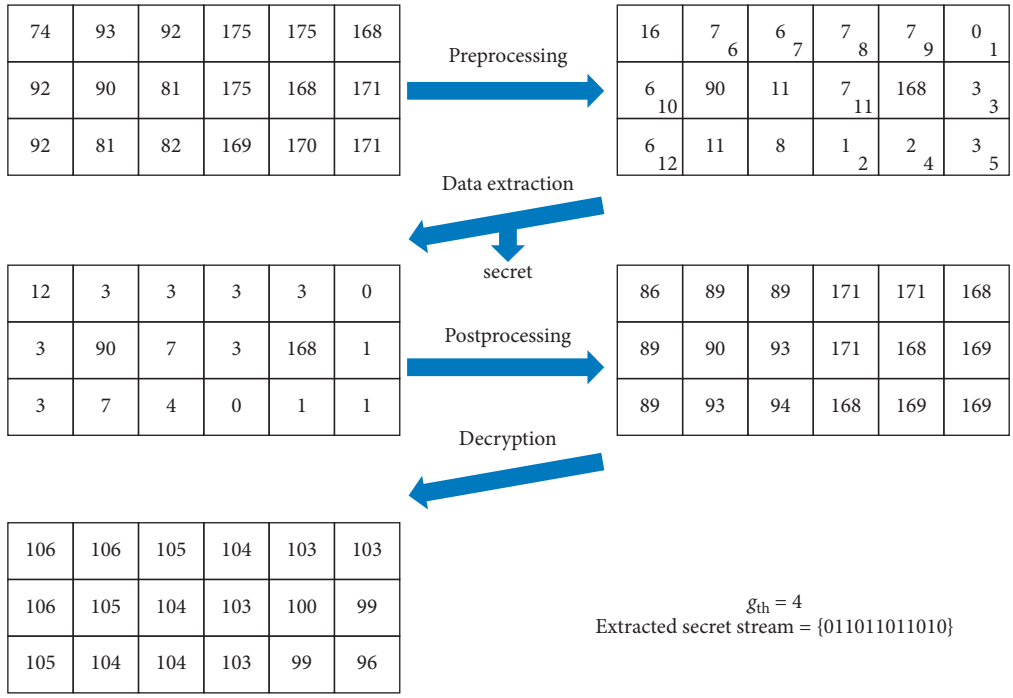


Figure 8: An illustrative example of the proposed data extraction and image recovery process.

4. Blockchain-Based Reversible Data Hiding System

The proposed blockchain system is presented in this section, with the aim of providing integrity and traceability in existing RDHEI. A consortium blockchain system is used in this paper, in which the participants in the blockchain network are trusted parties. For instance, the hospitals and health research institute (HRI) can form a consortium blockchain to share the medical images among themselves. In this way, an expensive consensus algorithm (e.g., PoW) is not required; it can be replaced with a lightweight algorithm like PBZT. Moreover, consortium blockchain is also a more appropriate choice as we do not expect any unauthorized person to join the blockchain network and gain access to the medical information, which is supposed to be private.

4.1. Architecture of the Proposed Blockchain-Based Reversible Data Hiding System. Figure 9 shows the architecture of the proposed blockchain system, which can be used to exchange medical images safely. Doctor A first encrypts the medical records of a patient and then generates the CSMI using the proposed RDHEI scheme, wherein the encrypted medical records are embedded into the patient's medical image. The encrypted medical records and CSMI are stored in the database of his hospital (Hospital X). A transaction block is generated and added to the blockchain.

With the proposed blockchain system, anyone who wants to share medical information can verify the integrity of transmitted data at the receiver end. For instance, Doctor A shares the CSMI with Doctor B and Medicate Institute Y. Upon receiving the CSMI, Doctor B first computes the hash value of received data and compares it against the blockchain to verify its integrity. He then extracts the hidden medical records with a legitimate key and then decrypts the medical records with another legitimate key. A similar process is performed by Medical Institute Y. The details on how to generate a new block and verify the integrity of CSMI are presented in the next section.

4.2. Process of Generating New Blocks and Verifying the Integrity of CSMI. Referring to Figure 10, the medical records of a patient are first encrypted by a symmetric key algorithm through the following equation:

$$C_{\mathrm{MR}} = \mathrm{Enc}(K_1, \mathrm{MR}), \tag{12}$$

where MR refers to the medical records, K_1 is the symmetric key, and C_{MR} is the resultant ciphertext. Next, we generate another key $K_2 = \textit{Key-I} \,||\, \textit{Key-II} \,||\, \textit{Key-III}$, where *Key-I*, *Key-II*, and *Key-III* are the keys used in the proposed RDHEI scheme. We embed the encrypted medical records (C_{MR}) into the medical image and generate a CSMI through the proposed RDHEI scheme. In Step 3, the generated CSMI and the hash value of the previous block in blockchain (H_{prev}) are concatenated. The hash value of the current block is generated through the following equation:

$$H_{\mathrm{curr}} = \mathrm{Hash}(H_{\mathrm{PREV}} || \mathrm{CSMI}), \tag{13}$$

where Hash can be any standardized cryptographic hash function (e.g., SHA-2 and SHA-3) and H_{curr} is the hash value of the current block. The newly generated block is transmitted to the blockchain network for the consensus process, wherein a lightweight algorithm can be used. Once the peers

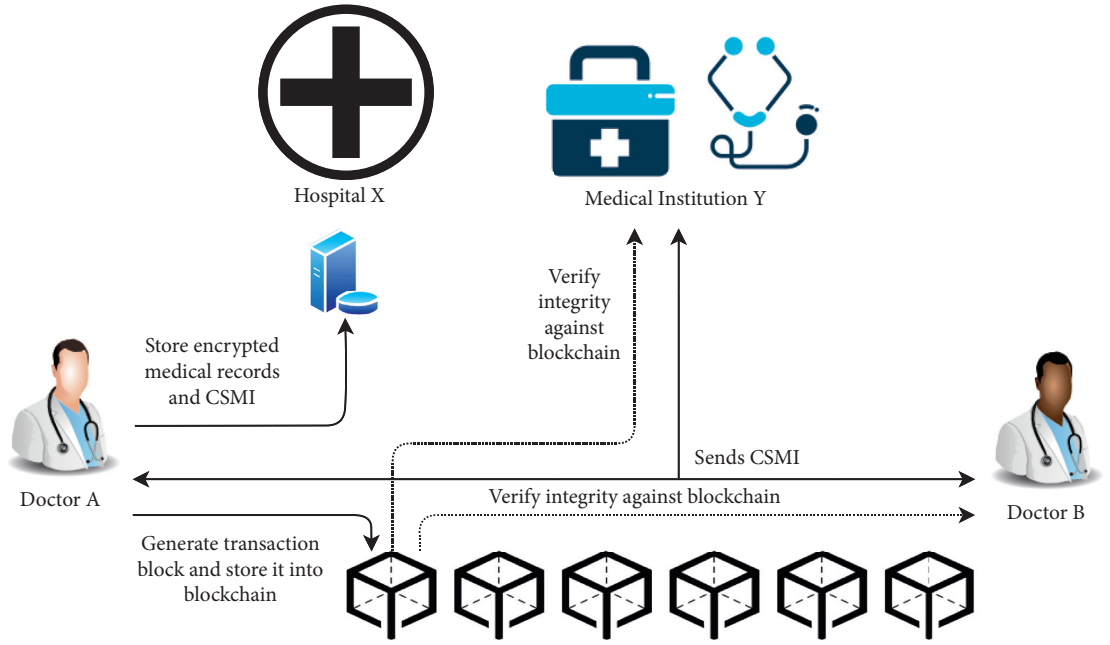


Figure 9: Architecture of the proposed blockchain system.

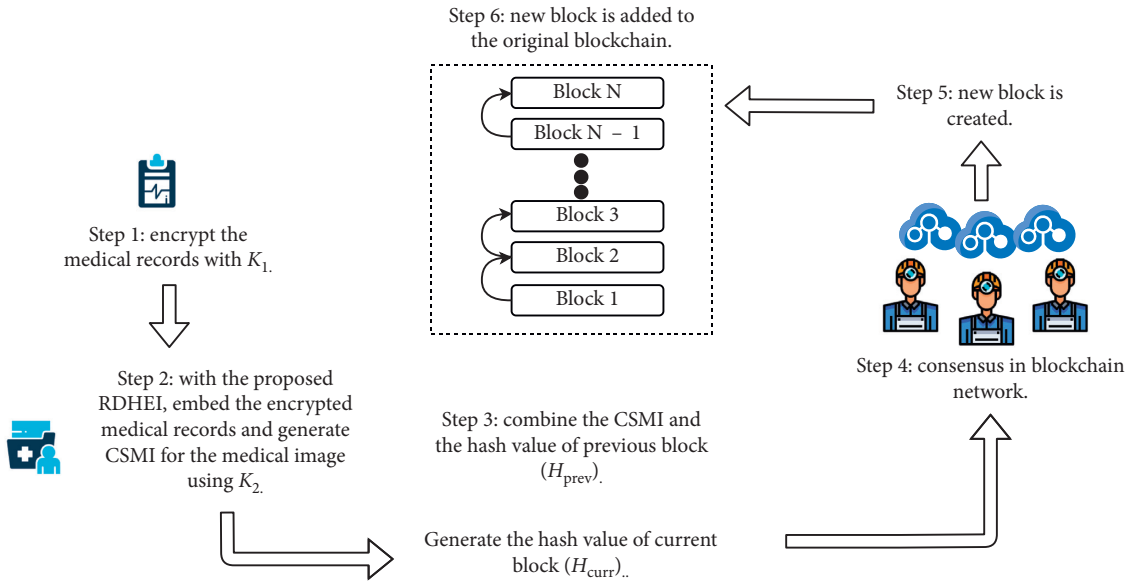


Figure 10: Transaction flow of the proposed blockchain system.

in the blockchain network approved the current block, a new block is generated (Step 5) and is added to the existing blockchain (Step 6). Assume that a person receives CSMI′ from a legitimate member in the blockchain network. To verify the integrity of the received CSMI, he first computes the hash value: $H_{curr}{}' = \text{Hash}(H_{PREV}||\text{CSMI})$ following the same hash function used in the block creation process. Then, he compares $H_{curr}{}'$ against the hash value H_{curr} stored in the blockchain. If both values are the same, then CSMI′ is untampered; otherwise, CSMI′ or C_{MR} has been modified.

The proposed blockchain system can be implemented with any popular blockchain framework, e.g., Hyperledger Fabric and Ethereum. Step 1 can employ an industry-grade block cipher (e.g., AES) to perform the encryption, while Step 2 can be carried out by using the techniques described in Section 3. Subsequently, Steps 3–6 are common operations found in a standard blockchain framework, which can be implemented easily.

4.3. Security Analysis. In this section, a security analysis on the proposed blockchain system is provided, with consideration of various attacking scenarios.

(1) Attackers cannot retrieve the steganography of medical images. In the proposed system, the CSMI is stored in the database of the hospital. Assuming that

the malicious attacker has access to the CSMI, he cannot recover the steganography of medical images (SMI). This is because he does not hold K2 to successfully recover the SMI. On top of that, he can never extract the patient's medical records embedded into CSMI

(2) The proposed system can protect the confidentiality of medical records. The patient's medical records are being encrypted by a symmetric key algorithm with key K_1, before embedding it into the medical image. Assume that the malicious attacker has access to the CSMI, and he had successfully extracted the SMI. In such a situation, he cannot successfully extract the medical records, because he does not have the K_1 to decrypt them. Hence, the confidentiality of the patient's medical records is protected

(3) The proposed system can achieve integrity. It is also possible that the malicious attacker is interested in creating fake CSMI instead of extracting information from it. For instance, one can generate fake keys K_1' to encrypt the legitimate medical records (C_{MR}') and then embed them to the other medical image to produce a fake CSMI′. This fake CSMI′ is being stored in the database as a new entry. However, the hospital can detect this fake CSMI′ easily by comparing it against the hash value stored in our blockchain system. Since CSMI′ is calculated from fake keys K_1' and K_2', the hash value generated (H_{curr}') is not the same as the one stored in the blockchain (H_{curr}). In order to create fake CSMI′ that everyone trusts, the attacker needs to modify the blockchain accordingly. However, referring to equation (13), the hash value in each new block depends on the previous block. To modify one record in blockchain, the attacker must also recalculate all the subsequent blocks, which is almost impossible. Hence, the proposed system is secure against malicious attacks that attempt to compromise its integrity

(4) The image encryption scheme is robust under chosen-plaintext analysis. The permutation process with *Key-I* has corrupted the block correspondence between the cover image and the encrypted image. Without the mapping information between the image blocks before and after encryption, chosen plaintext analysis is useless. Even if the block mapping is known, it is still too complicated for analysis since image blocks are encrypted with distinct keys generated by *Key-II*

5. Experimental Results and Discussion

In this section, we present the experimental results of the proposed RHDEI and its application to medical images. The hardware resources are Intel ® Core ™ i7-3770 CPU @ 3.40 GHz and 8 GB RAM PC. The application software is MATLAB R2017a running with Windows 10 Professional operating system. Six standard test images (see Figure 11) including "Airplane," "Baboon," "Boat," "Lena," "Peppers," and "Sailboat" are applied to demonstrate the effectiveness of our scheme. The encrypted images are shown in Figure 12. The preprocessed images and postprocessed images with data embedded are given in Figures 13 and 14, respectively.

The histograms of the test images and their corresponding encrypted images are given in Figures 15 and 16. Referring to Figure 16, the histogram of encrypted images is evenly distributed in the entire range of gray levels regardless of the different image features. This indicates a high-security level of the proposed image encryption scheme.

The histogram after the preprocess of the embedding phase is given in Figure 17. For smooth images, the histogram is more concentrated while it is more distributed for complex images such as "Baboon." Note that the histogram band of $[128, 192]$ gray levels is very "clear" by observation. To know the details, we further restrict the display range of the vertical axis to $[0, 200]$ pixels as shown in Figure 18. There are some pixels distributed in the band. To achieve the complete reversibility of the embedding scheme, the metadata includes three data segments. The first segment records the threshold g_{th} with six bits. The second segment records the number of pixels in the vacating band with twelve bits. The last segment contains all details of each pixel in the vacating band by one byte for its gray level and two bytes for its coordinates in the image. Therefore, the total length of metadata is $6 + 12 + (3 \times 8 \times N_p)$, where N_p is the number of pixels in the vacating band. After histogram shifting and data embedding, the histogram in Figure 17 changes to the distribution shown in Figure 19, where the applied threshold of vacating band is $g_{th} = 64$. Figure 20 shows the resulting histogram of the marked encrypted image after postprocessing. The histogram is still evenly distributed and preserves at a high-security level.

Since the total number of pixels in the vacating band is very small, we can recover the histogram in a more efficient way by retrieving the first two segments of metadata to get g_{th} and determine the remaining length of metadata to be discarded. Thus, the recovery of the vacating band can be skipped, and the process is proceeded directly to extract the secret data stream. To know the visual quality of the approximated image, we apply the peak-signal-to-noise ratio (PSNR) index given by

$$\mathrm{MSE} = \frac{1}{W \times H} \sum_{i=1}^{W} \sum_{j=1}^{H} \left(I_{i,j} - I_{i,j}'\right)^2, \tag{14}$$

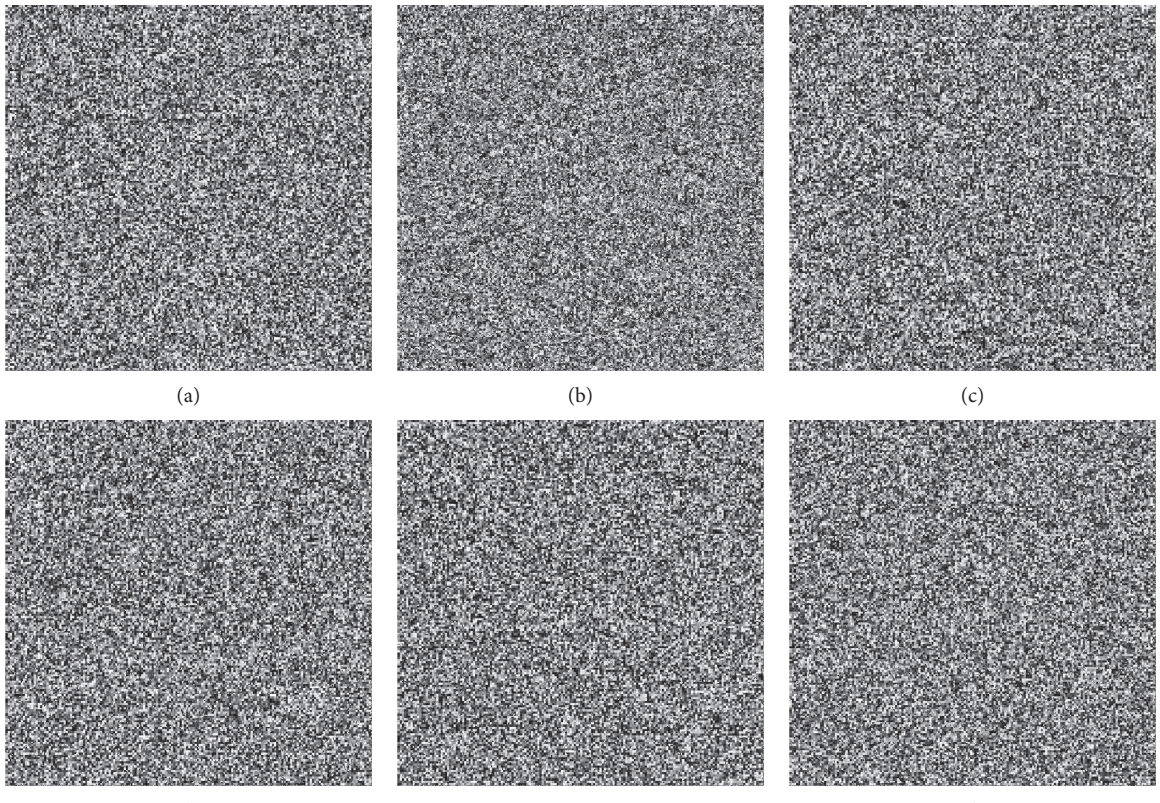
$$\mathrm{PSNR} = 10 \log_{10} \frac{255^2}{\mathrm{MSE}}, \tag{15}$$

where MSE is the mean square error between the cover image $I_{i,j}$ and the approximated image $I_{i,j}'$ and W and H are the width and height of the image. The second measure of similarity between them is the structural similarity (SSIM) defined by

(a) (b) (c)

(d) (e) (f)

Figure 11: Cover images of size 512 × 512. (a) Airplane. (b) Baboon. (c) Boats. (d) Lena. (e) Peppers. (f) Sailboat.



(a) (b) (c)

(d) (e) (f)

Figure 12: The encrypted images. (a) Airplane. (b) Baboon. (c) Boats. (d) Lena. (e) Peppers. (f) Sailboat.

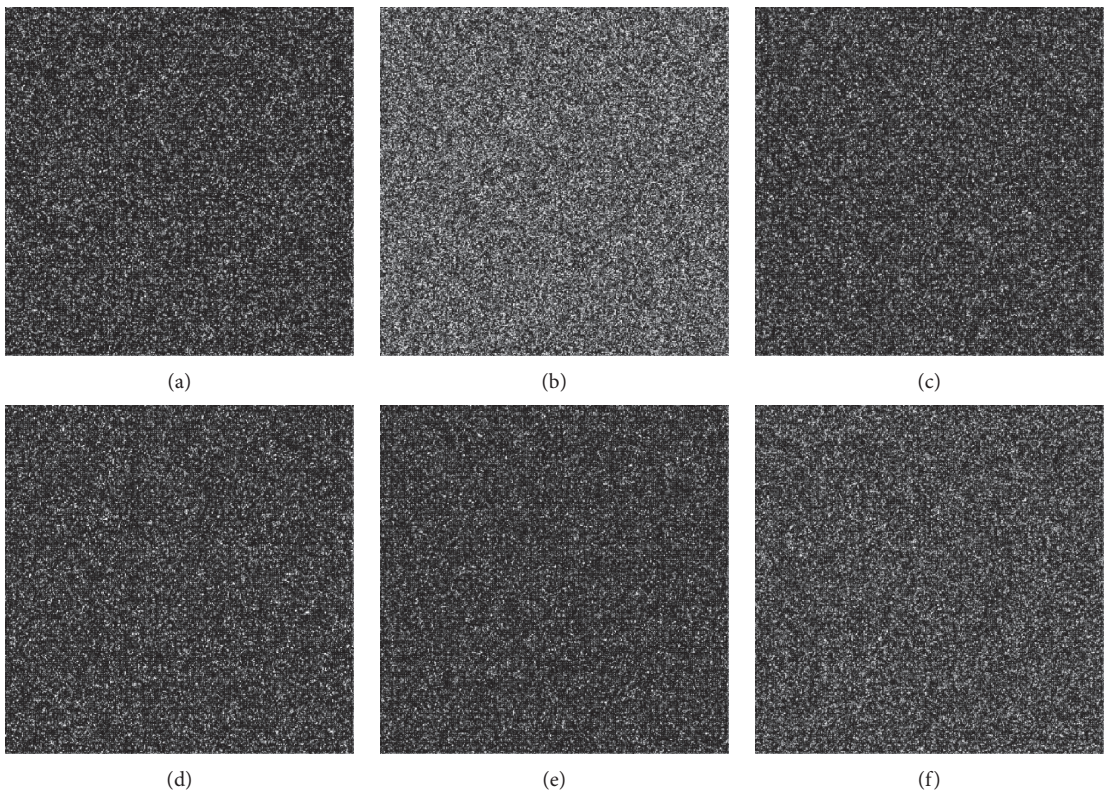


Figure 13: Preprocessed images. (a) Airplane. (b) Baboon. (c) Boat. (d) Lena. (e) Peppers. (f) Sailboat.

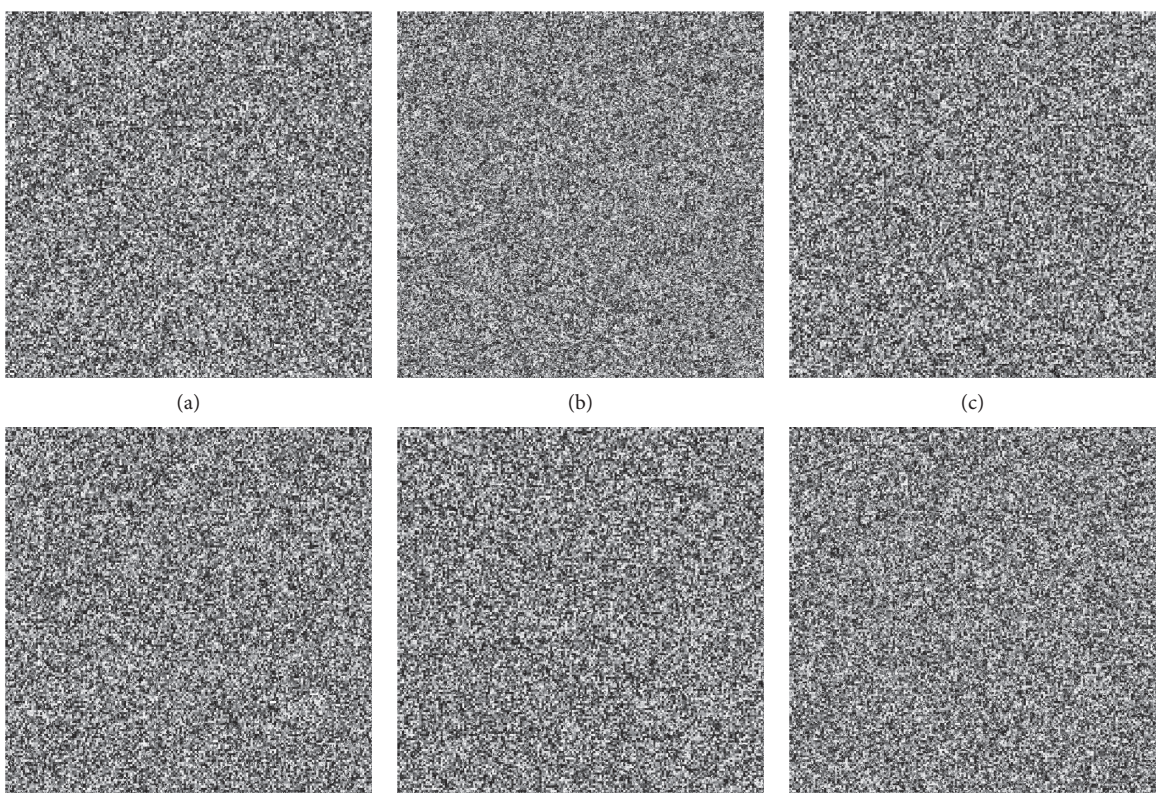


Figure 14: Postprocessed images with data embedded. (a) Airplane. (b) Baboon. (c) Boats. (d) Lena. (e) Peppers. (f) Sailboat.

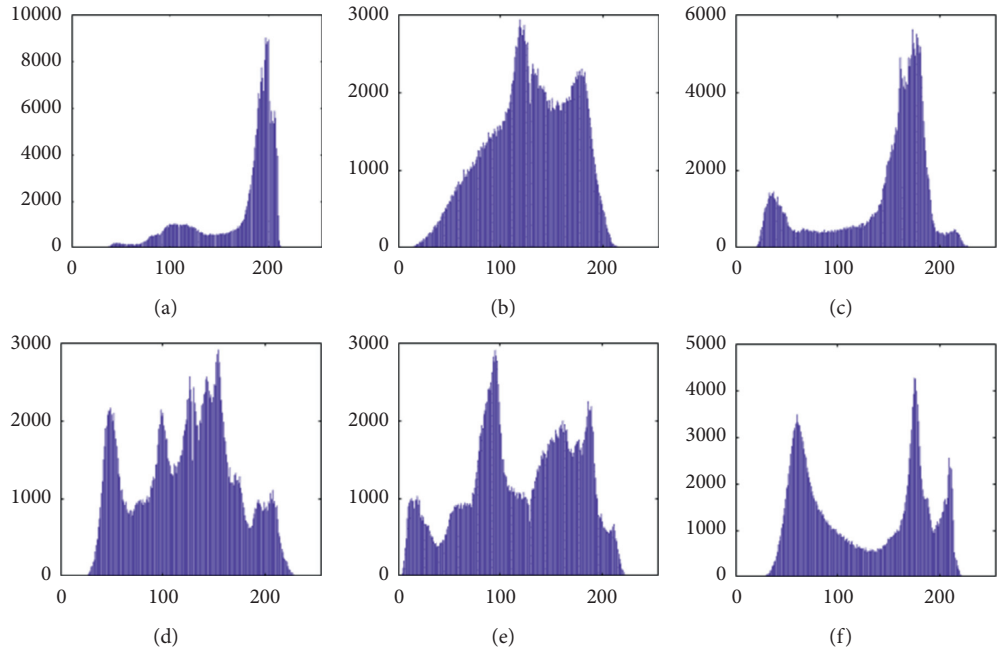


Figure 15: Histogram of cover images. (a) Airplane. (b) Baboon. (c) Boat. (d) Lena. (e) Peppers. (f) Sailboat.

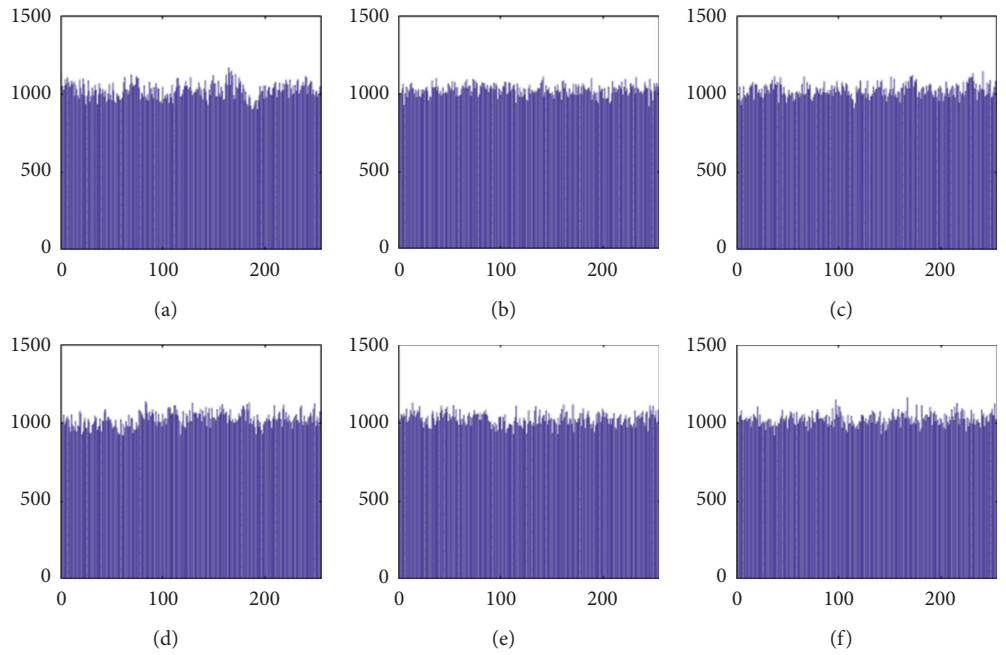


Figure 16: Histogram of encrypted images. (a) Airplane. (b) Baboon. (c) Boat. (d) Lena. (e) Peppers. (f) Sailboat.

$$\mathrm{SSIM}(I, I') = \frac{(2\mu_I\mu_{I'} + c_1)(2\sigma_{II'} + c_2)}{(\mu_I^2 + \mu_{I'}^2 + c_1)(\sigma_I^2 + \sigma_{I'}^2 + c_2)}, \quad (16)$$

where μ_I, $\mu_{I'}$, σ_I, and $\sigma_{I'}$ are the mean values and standard deviations and $\sigma_{II'}$ is the covariance of the two images.

The experimental results of the proposed RDHEI scheme with different thresholds are listed in Table 1. The "capacity" and "metadata" are measured in bits. The actual embedding capacity of secret data can be calculated by subtracting the amount of metadata from the "capacity" value. The "ER" is the embedding rate measured in bits per pixel (bpp). The PSNR and SSIM values listed in the table indicate that the approximated images under fast recovery have good visual quality.

Figure 21 shows the PSNR values with respect to the embedding rate (ER). The threshold value determines the

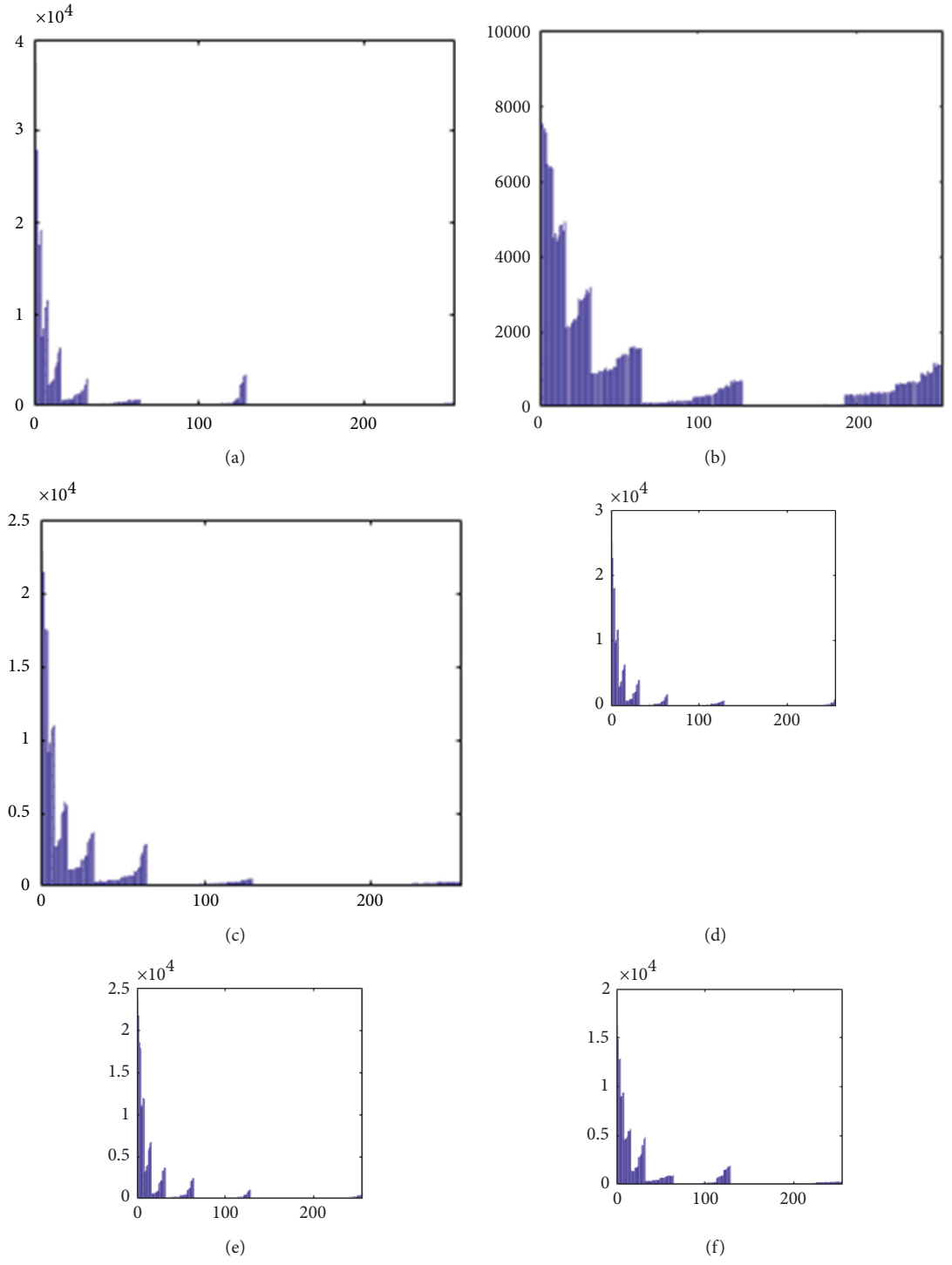


Figure 17: Histogram of preprocessed images. (a) Airplane. (b) Baboon. (c) Boat. (d) Lena. (e) Peppers. (f) Sailboat.

width of the vacating band and therefore degrades the visual quality. However, the ER also increases with the increased threshold. Another important point of observation is that the PSNR value is dependent on the image feature. The most complex image "Baboon" has the lowest PSNR level among all.

The second set of test images are two medical images downloaded from the website of MIDAS/National Alliance for Medical Image Computing (NAMIC) as shown in Figure 22. The image size 256 × 256 is relatively smaller than the standard test images in the first experiment. The encrypted images, preprocessed images, and postprocessed images are given in Figures 23–25, respectively. Histograms of the cover images are shown in Figure 26. Notice that the pixel values are concentrated to very low gray levels due to the inherent nature of MRI images. This phenomenon is

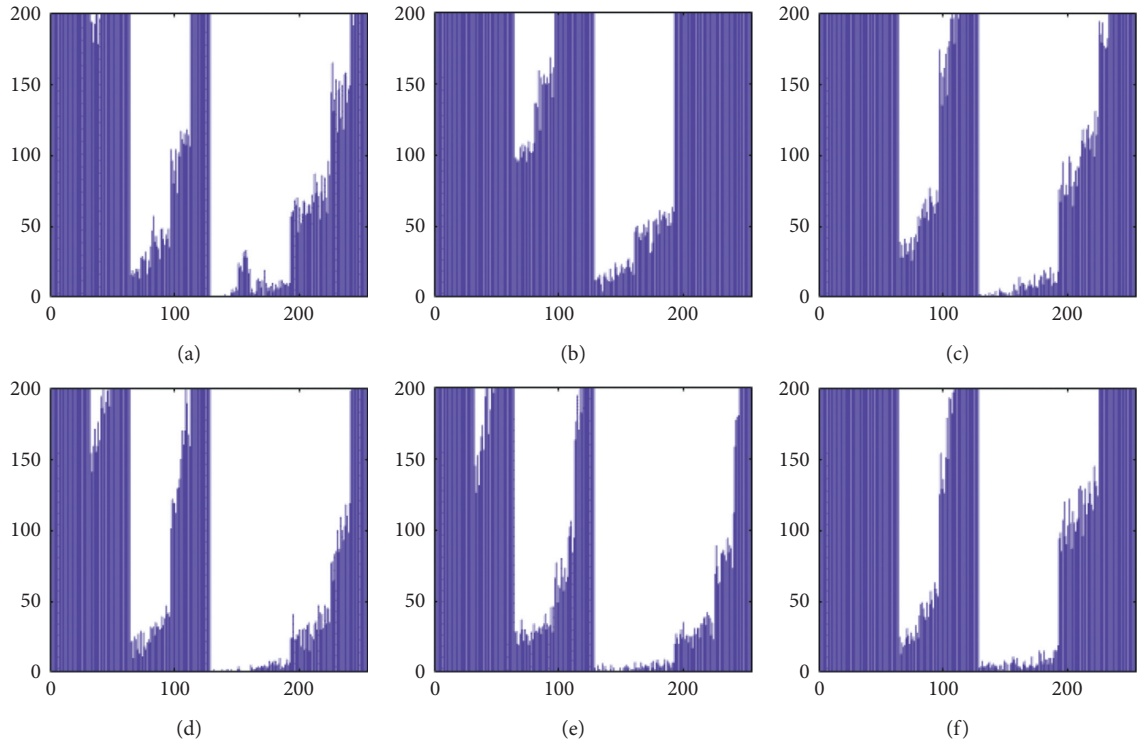


Figure 18: Histogram of preprocessed images at a finer scale. (a) Airplane. (b) Baboon. (c) Boat. (d) Lena. (e) Peppers. (f) Sailboat.

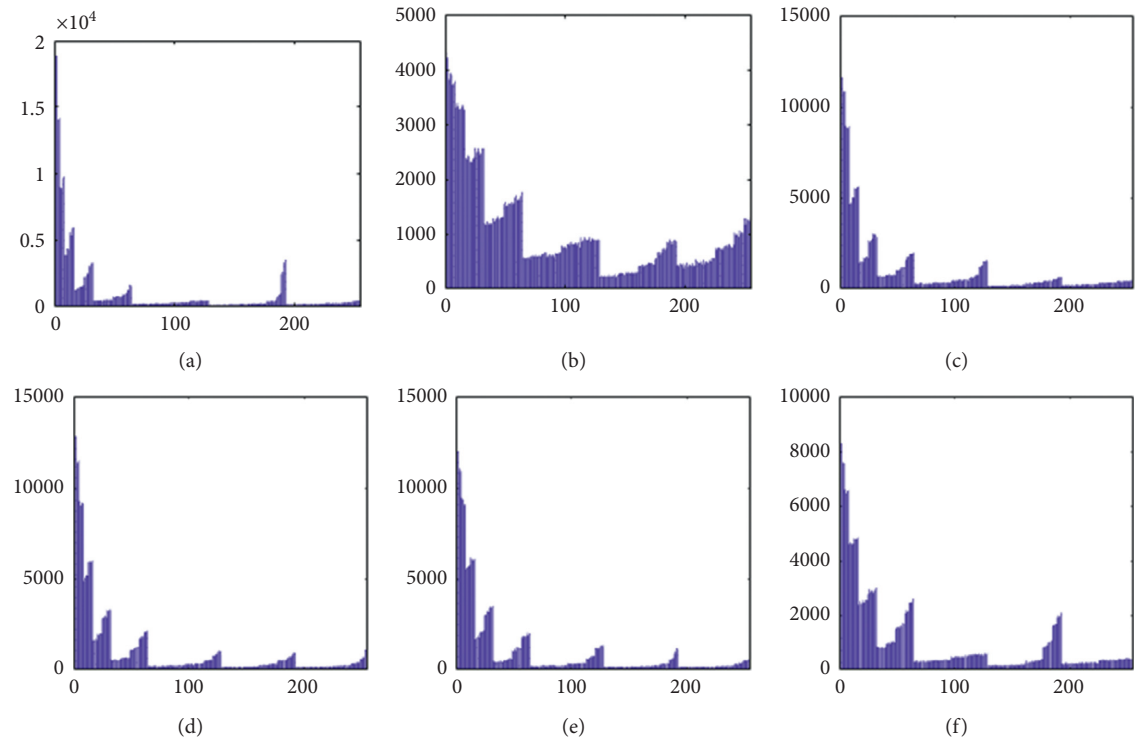


Figure 19: Histogram of data embedded images before postprocessing. (a) Airplane. (b) Baboon. (c) Boat. (d) Lena. (e) Peppers. (f) Sailboat.

especially beneficial to our data hiding scheme. The histogram distributions for different processing phases are given in Figures 27–31. The experimental values are listed in Table 2. The major difference from the first experiment is that the embedding rate reaches 0.6 bpp at the lowest threshold of $g_{th} = 4$, which is much higher than the capacity

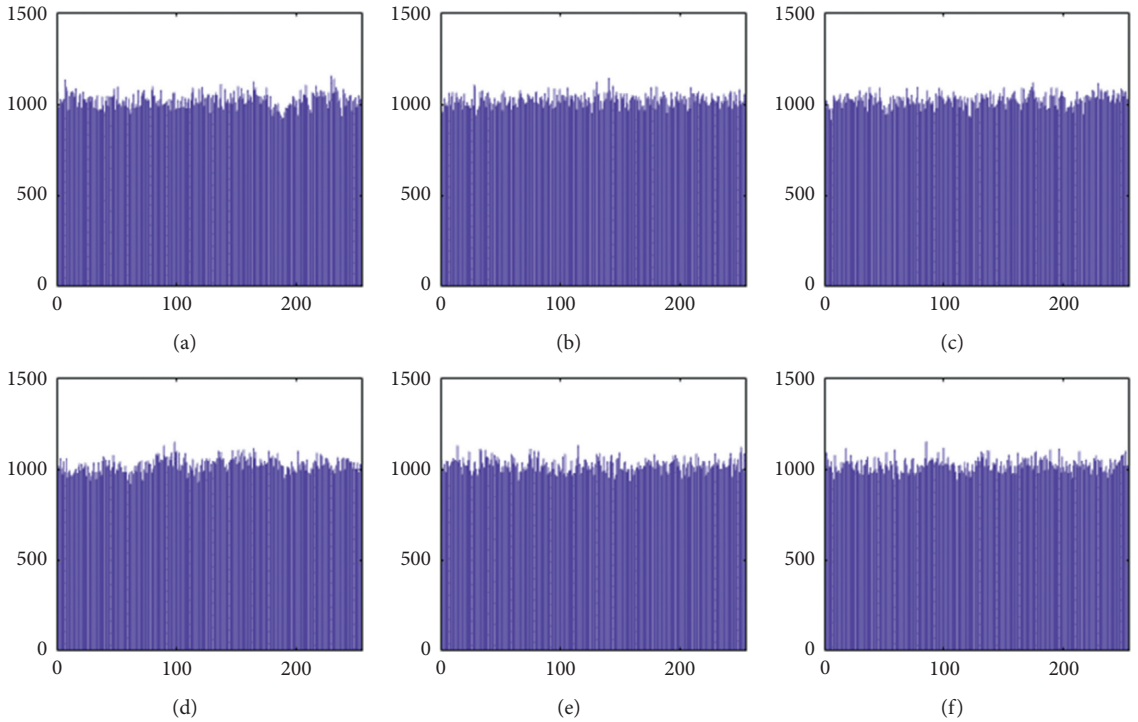


Figure 20: Histogram of postprocessed (final) images. (a) Airplane. (b) Baboon. (c) Boat. (d) Lena. (e) Peppers. (f) Sailboat.

Table 1: Experimental results of the proposed scheme with different thresholds.

Image	Metrics	$g_{th} = 4$	$g_{th} = 8$	$g_{th} = 16$	$g_{th} = 32$	$g_{th} = 64$
	Capacity	102125	140358	171658	190931	202715
	Metadata	42	66	162	6666	12858
Airplane	ER (bpp)	0.38	0.53	0.65	0.72	0.77
	PSNR	61.64	58.29	54.77	37.72	34.35
	SSIM	0.9999	0.9999	0.9999	0.995	0.9912
	Capacity	30506	56047	93320	134879	173447
	Metadata	1242	2298	4962	12810	51138
Baboon	ER (bpp)	0.11	0.21	0.35	0.51	0.66
	PSNR	42.61	40.32	36.99	32.79	26.12
	SSIM	0.999	0.9983	0.9961	0.9895	0.9516
	Capacity	79348	119982	153134	184078	209831
	Metadata	114	210	546	1866	10626
Boat	ER (bpp)	0.3	0.45	0.58	0.7	0.8
	PSNR	56.36	52.21	46.9	40.83	32.5
	SSIM	0.9999	0.9999	0.9996	0.9983	0.9879
	Capacity	83908	127070	164152	194295	209334
	Metadata	138	162	354	738	4506
Lena	ER (bpp)	0.32	0.48	0.62	0.74	0.79
	PSNR	53.11	52.07	48.24	45.27	36.04
	SSIM	0.9999	0.9999	0.9997	0.9993	0.9944
	Capacity	81892	127696	167267	195067	214845
	Metadata	258	498	954	1794	5586
Pepper	ER (bpp)	0.31	0.48	0.63	0.74	0.81
	PSNR	51.15	47.48	43.9	40.5	35.82
	SSIM	0.9998	0.9997	0.9993	0.9982	0.9944
	Capacity	56910	93681	134503	176246	195452
	Metadata	474	930	1794	3906	10122
Sailboat	ER (bpp)	0.21	0.35	0.51	0.67	0.74
	PSNR	45.23	41.76	39.11	35.87	31.43
	SSIM	0.9995	0.999	0.998	0.9958	0.9872

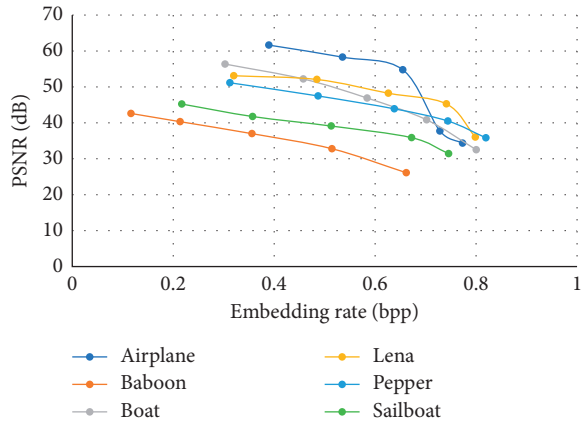
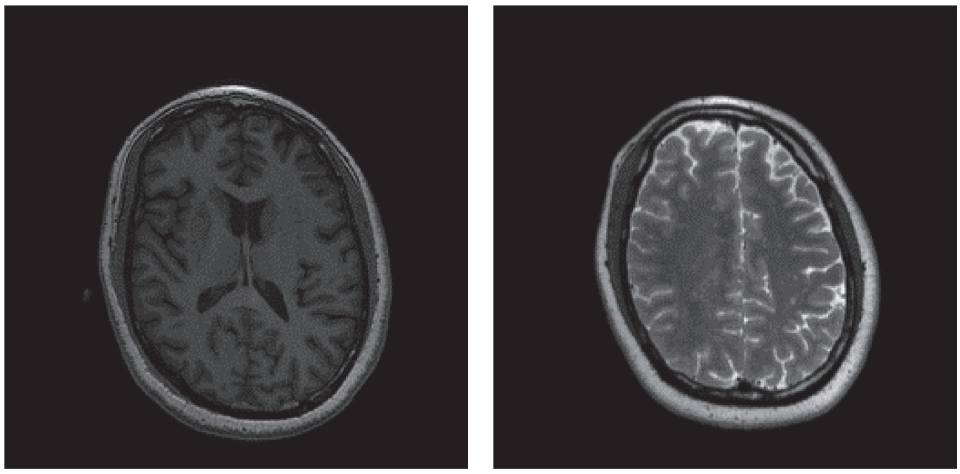
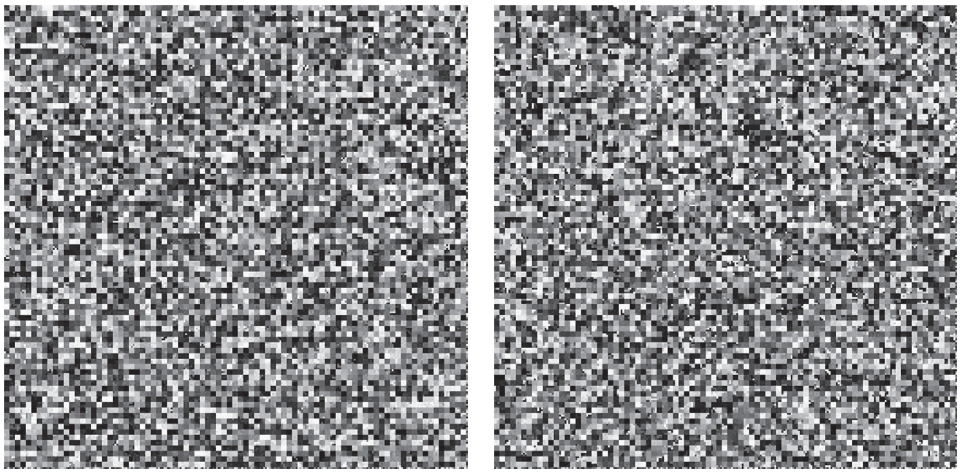


Figure 21: Performance comparisons of test images with the proposed scheme.



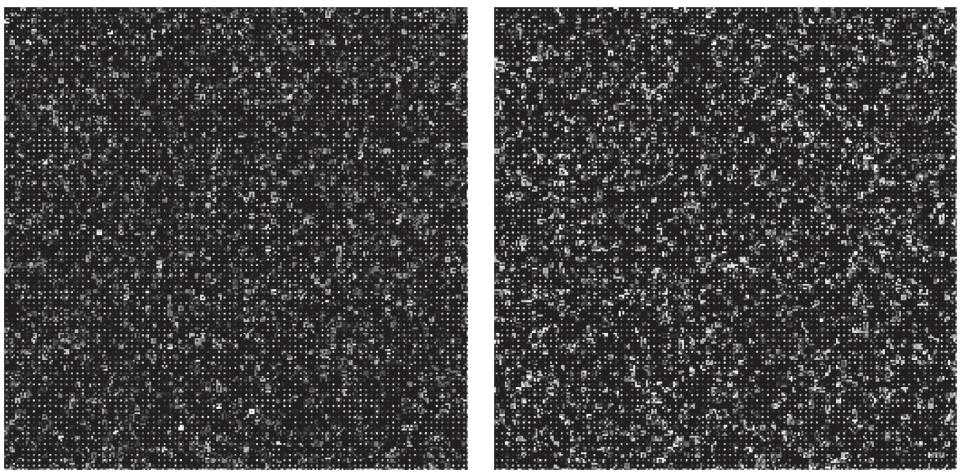
(a) (b)

Figure 22: Cover images of size 256×256. (a) Medical image A. (b) Medical image B.



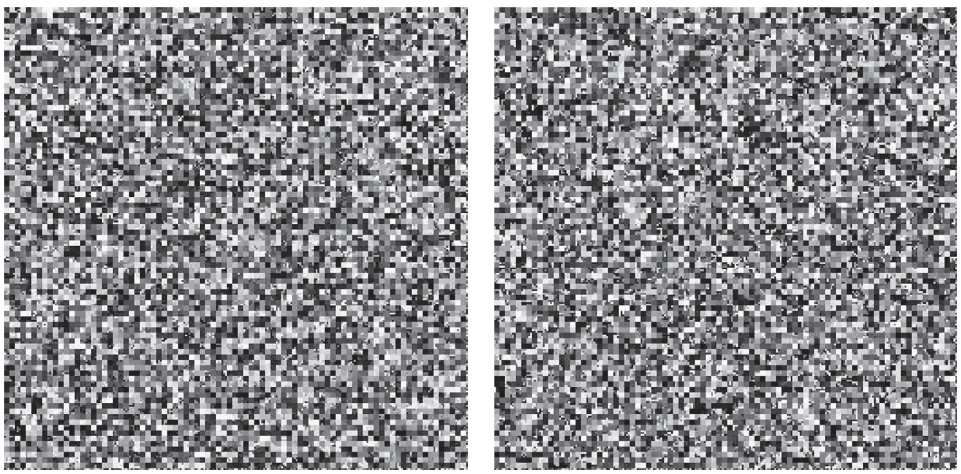
(a) (b)

Figure 23: The encrypted images. (a) Medical A. (b) Medical B.



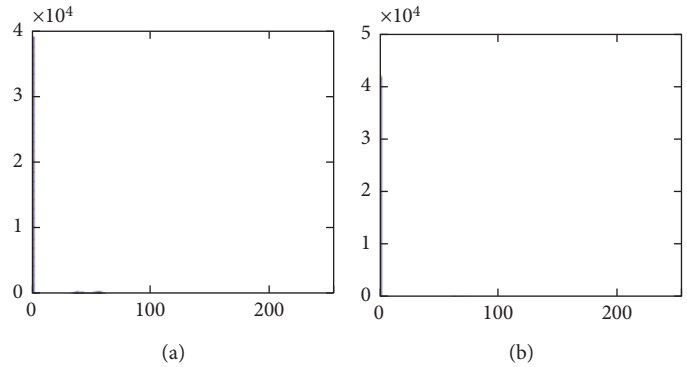
(a) (b)

Figure 24: Preprocessed images. (a) Medical A. (b) Medical B.



(a) (b)

Figure 25: Postprocessed images with data embedded. (a) Medical A. (b) Medical B.



(a) (b)

Figure 26: Histogram of cover images. (a) Medical A. (b) Medical B.

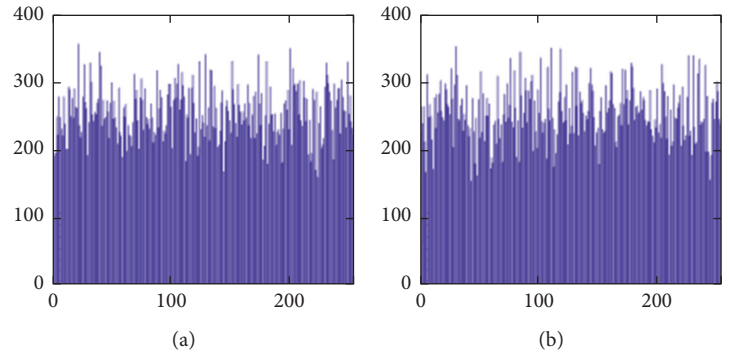


Figure 27: Histogram of encrypted images. (a) Medical A. (b) Medical B.

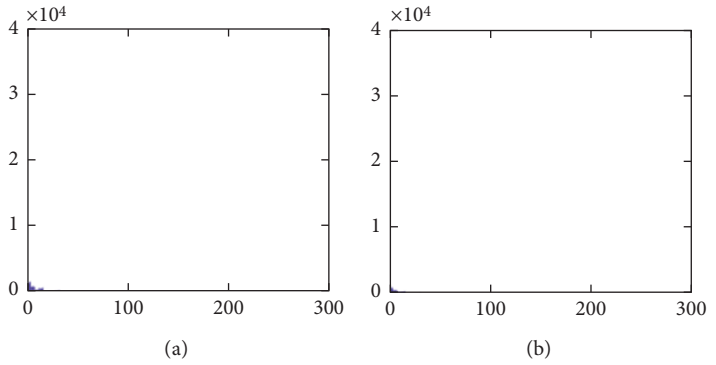


Figure 28: Histogram of preprocessed images. (a) Medical A. (b) Medical B.

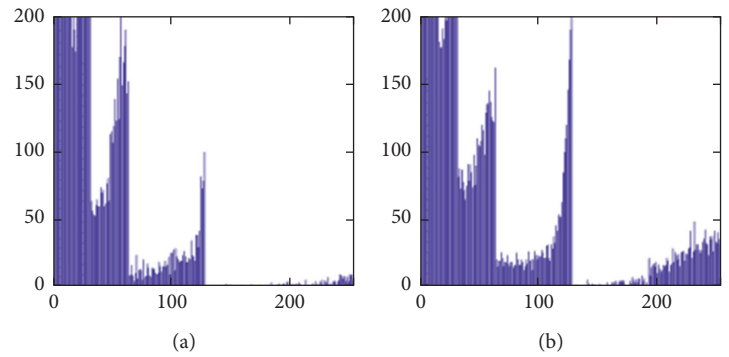


Figure 29: Histogram of preprocessed images at a finer scale. (a) Medical A. (b) Medical B.

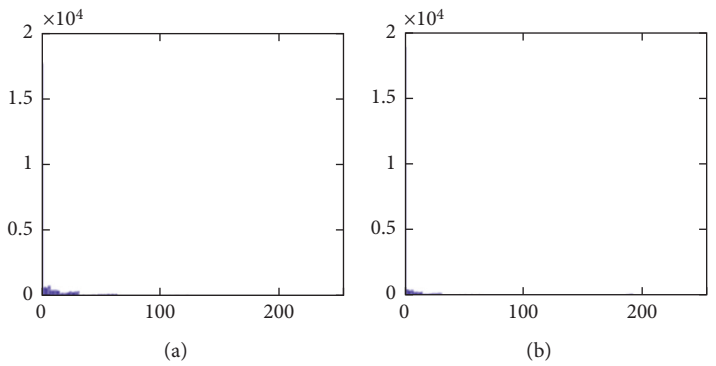


Figure 30: Histogram of data embedded images. (a) Medical A. (b) Medical B.

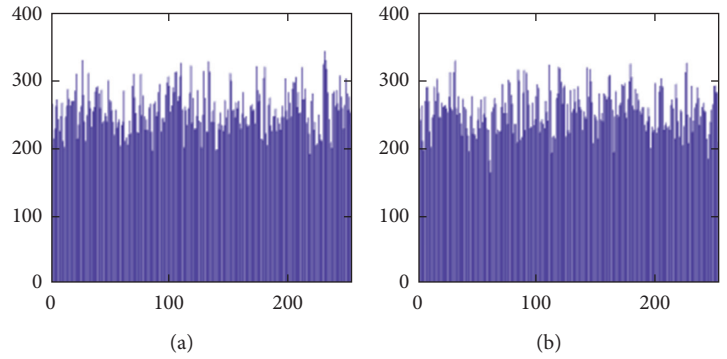


Figure 31: Histogram of postprocessed images. (a) Medical A. (b) Medical B.

Table 2: Experimental results of the MRI medical images with different thresholds.

Image	Metrics	$g_{th} = 4$	$g_{th} = 8$	$g_{th} = 16$	$g_{th} = 32$	$g_{th} = 64$
Medical A	Capacity	39671	43490	48507	52655	56038
	Metadata	66	114	234	738	1554
	ER (bpp)	0.6	0.66	0.74	0.8	0.85
	PSNR	70.1	57.02	42.68	37.44	35.34
	SSIM	0.9999	0.9999	0.9989	0.9971	0.9953
Medical B	Capacity	40608	43388	46851	50168	53448
	Metadata	138	258	546	1122	4338
	ER (bpp)	0.61	0.66	0.71	0.76	0.81
	PSNR	49.6	45.67	41.45	37.48	30.82
	SSIM	0.9999	0.9996	0.999	0.9976	0.9871

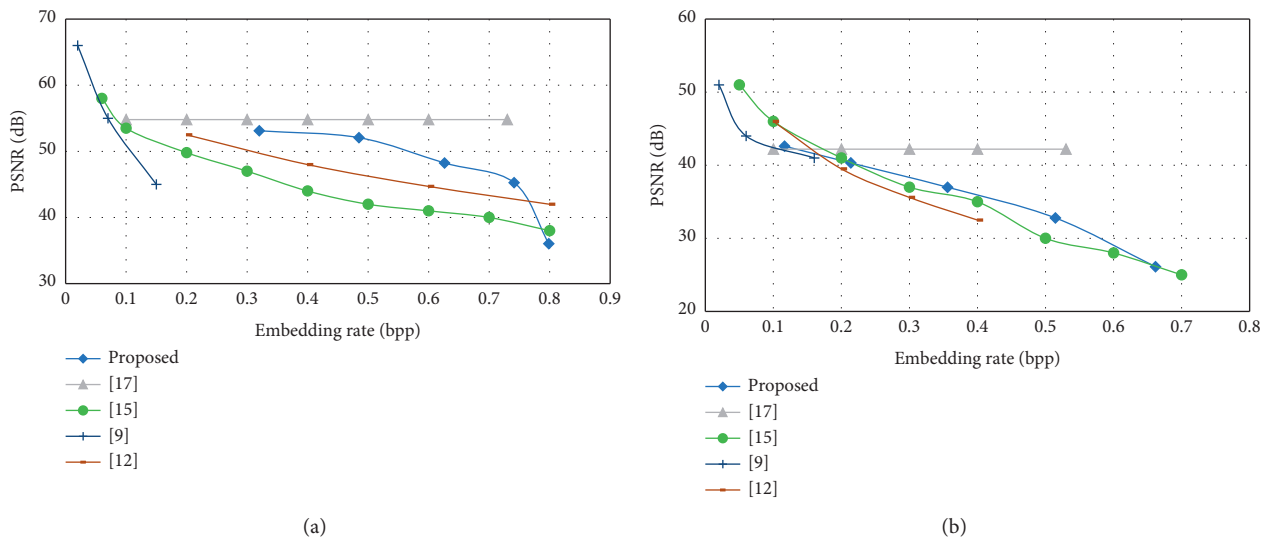


Figure 32: Comparisons of the proposed scheme with related works for typical test images. (a) Lena. (b) Baboon.

of standard test images. At the highest threshold of $g_{th} = 64$, the medical images can reach an embedding rate of 0.8 bpp.

Figure 32 compares the proposed scheme with other related works, including VRAE-based schemes (Zhang et al.'s scheme [9, 12]) and VRBE-based schemes (Cao et al.'s scheme [21] and Malik et al.'s scheme [22]), for two typical test images "Lena" and "Baboon." Note that all these schemes are completely reversible. The PSNR values are compared under the situation of discarding some information in image deciphering. The proposed scheme reaches a good embedding rate with a good visual quality of the approximation image. The scheme proposed in [17] is VRBE-based; therefore, it provides only one fixed embedding rate.

Compared to a recently proposed RDH technique by Anushiadevi et al. [23], the proposed technique offers a

flexible threshold (g_{th}) to allow performance tuning. The user can flexibly select a higher embedding capacity or higher PSNR by adjusting g_{th}; this is not found in [23]. Another recent work that offered homomorphic encryption on RDH was proposed by Anushiadevi et al. [24]. Our proposed technique shows performance on par with them in terms of ER, PSNR, and flexible threshold configuration.

6. Conclusion

Medical images are important assets to the medical industry, and they should be protected to avoid potential infringement of privacy. To achieve this goal, we propose an RDHEI scheme, in which the cover image is encrypted by block permutation using encryption *Key-I* and stream ciphering using encryption *Key-II*. Then, the encrypted patient data is embedded into the encrypted image through *Key-III* to produce a secure CSMI. Experimental results show that the performance of the proposed RDHEI scheme is excellent when applied to medical images. The embedding rate is over 0.8 bpp for typical medical images. An image of high visual quality can be recovered even if the receiver only holds the image encryption keys. Since most processes of the proposed RDHEI scheme are based on the simple Exclusive-OR operation, our scheme can be executed very efficiently. To provide an integrity check, we propose a blockchain system on top of the RDHEI. The hash value of CSMI is stored in a commonly used blockchain system for future verification. The proposed RDHEI blockchain system allows the user to check the integrity of CSMI from time to time, which shows additional benefit compared to the conventional RDHEI schemes.

Data Availability

The data used in the experiments and discussions in the paper are available within this article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the Brain Pool Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (2019H1D3A1A01102607).

References

[1] M. Hadzic, T. Dillon, and E. Chang, "Use of digital ecosystem and ontology technology for standardization of medical records," in *Proceedings of the Inaugural IEEEIES Digital EcoSystems and Technologies Conference*, pp. 595–601, Cairns, Australia, February 2007.

[2] H. Peng, B. Yang, L. Li, and Y. Yang, "Secure and traceable image transmission scheme based on semitensor product compressed sensing in telemedicine system," *IEEE Internet of Things Journal*, vol. 7, no. 3, pp. 2432–2451, 2020.

[3] M. G. R. Alam, M. S. Munir, M. Z. Uddin, M. S. Alam, T. N. Dang, and C. S. Hong, "Edge-of-things computing framework for cost-effective provisioning of healthcare data," *Journal of Parallel and Distributed Computing*, vol. 123, pp. 54–60, 2019.

[4] Y. Yang, X. Xiao, X. Cai, and W. Zhang, "A secure and high visual-quality framework for medical images by contrast-enhancement reversible data hiding and homomorphic encryption," *IEEE Access*, vol. 7, pp. 96900–96911, 2019.

[5] Y. Yang, X. Xiao, X. Cai, and W. Zhang, "A secure and privacy-preserving technique based on contrast-enhancement reversible data hiding and plaintext encryption for medical images," *IEEE Signal Processing Letters*, vol. 27, pp. 256–260, 2020.

[6] J.-J. Li, C.-F. Lee, C.-C. Chang, J.-Y. Lin, and Y.-H. Wu, "Reversible data hiding scheme based on quad-tree and pixel value ordering," *IEEE Access*, vol. 7, pp. 142947–142962, 2019.

[7] F. Aziz, T. Ahmad, A. H. Malik, M. I. Uddin, S. Ahmad, and M. Sharaf, "Reversible data hiding techniques with high message embedding capacity in images," *PLoS One*, vol. 15, Article ID e0231602, 2020.

[8] Z. Wang, N. Luo, and P. Zhou, "Guardhealth: blockchain empowered secure data management and graph convolutional network enabled anomaly detection in smart healthcare," *Journal of Parallel and Distributed Computing*, vol. 142, pp. 1–12, 2020.

[9] R. Zhang, C. Lu, and J. Liu, "A high capacity reversible data hiding scheme for encrypted covers based on histogram shifting," *Journal of Information Security and Applications*, vol. 47, pp. 199–207, 2019.

[10] S. Gao, T. Yu, J. Zhu, and W. Cai, "T-PBFT: an EigenTrust-based practical Byzantine fault tolerance consensus algorithm," *China Communications*, vol. 16, no. 12, pp. 111–123, 2019.

[11] Z. Ni, Y. Q. Shi, N. Ansari, and W. Su, "Reversible data hiding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 16, no. 3, pp. 354–362, 2006.

[12] J. Fridrich, M. Goljan, and R. Du, "Invertible authentication," *Security and Watermarking of Multimedia, Contents*, vol. 3, pp. 197–209, 2001.

[13] J. Fridrich, M. Goljan, and R. Du, "Invertible authentication watermark for jpeg images," in *Proceedings of the International Conference on Information Technology: Coding and Computing*, pp. 223–227, Las Vegas, NV, USA, April 2001.

[14] Y. Du, Z. Yin, and X. Zhang, "Improved lossless data hiding for jpeg images based on histogram modification," *Computer Mater Continua*, vol. 55, no. 3, pp. 495–507, 2018.

[15] W. Wang, J. Ye, T. Wang, and W. Wang, "A high capacity reversible data hiding scheme based on right-left shift," *Signal Processing*, vol. 150, pp. 102–115, 2018.

[16] S. Weng, J.-S. Pan, and L. Li, "Reversible data hiding based on an adaptive pixel-embedding strategy and two-layer embedding," *Information Sciences*, vol. 369, pp. 144–159, 2016.

[17] S. Weng, Y. Liu, J.-S. Pan, and N. Cai, "Reversible data hiding based on flexible block-partition and adaptive block-modification strategy," *Journal of Visual Communication and Image Representation*, vol. 41, pp. 185–199, 2016.

[18] W. Puech, M. Chaumont, and O. Strauss, "A reversible data hiding method for encrypted images," in *Proceedings of the SPIE 6819, Security, Forensics, Steganography, and Watermarking of Multimedia Contents X*, vol. 68191E, San Jose, CA, USA, January 2008.

[19] X. Zhang, "Separable reversible data hiding in encrypted image," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 2, pp. 826–832, 2012.

[20] K. Ma, W. Zhang, X. Zhao, N. Yu, and F. Li, "Reversible data hiding in encrypted images by reserving room before encryption," *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 3, pp. 553–562, 2013.

[21] X. Cao, L. Du, X. Wei, D. Meng, and X. Guo, "High capacity reversible data hiding in encrypted images by patch-level sparse representation," *IEEE Transactions on Cybernetics*, vol. 78, no. 46, pp. 1132–1143, 2016.

[22] A. Malik, H.-X. Wang, Y. Chen, and A. N. Khan, "A reversible data hiding in encrypted image based on prediction-error estimation and location map," *Multimedia Tools and Applications*, vol. 79, no. 17-18, pp. 11591–11614, 2020.

[23] R. Anushiadevi, P. Praveenkumar, J. B. B. Rayappan, and R. Amirtharajan, "Uncover the cover to recover the hidden secret—a separable reversible data hiding framework," *Multimedia Tools and Applications*, 2021.

[24] R. Anushiadevi, P. Praveenkumar, J. B. B. Rayappan, and R. Amirtharajan, "Reversible data hiding method based on pixel expansion and homomorphic encryption," *Journal of Intelligent & Fuzzy Systems*, vol. 39, no. 3, pp. 2977–2990, 2020.

Hindawi
Journal of Healthcare Engineering
Volume 2021, Article ID 5549779, 15 pages
https://doi.org/10.1155/2021/5549779

Research Article

Pathological Myopia Image Recognition Strategy Based on Data Augmentation and Model Fusion

Jianfeng Cui,[1] Xiaoyun Zhang,[2] Feibing Xiong,[2] and Chin-Ling Chen[3,4,5]

[1]*School of Software Engineering, Xiamen University of Technology, Xiamen 361024, China*
[2]*School of Opto-electronic and Communications Engineering, Xiamen University of Technology, Xiamen 361024, China*
[3]*School of Computer and Information Engineering, Xiamen University of Technology, Xiamen 361024, China*
[4]*School of Information Engineering, Changchun Sci-Tech University, Changchun 130600, China*
[5]*Department of Computer Science and Information Engineering, Chaoyang University of Technology, Taichung 41349, Taiwan*

Correspondence should be addressed to Xiaoyun Zhang; 18030043887@163.com and Chin-Ling Chen; clc@mail.cyut.edu.tw

Received 24 February 2021; Revised 2 April 2021; Accepted 27 April 2021; Published 6 May 2021

Academic Editor: Hao-Chun Lu

The automatic diagnosis of various retinal diseases based on fundus images is important in supporting clinical decision-making. Convolutional neural networks (CNNs) have achieved remarkable results in such tasks. However, their high expression ability possibly leads to overfitting. Therefore, data augmentation (DA) techniques have been proposed to prevent overfitting while enriching datasets. Recent CNN architectures with more parameters render traditional DA techniques insufficient. In this study, we proposed a new DA strategy based on multimodal fusion (DAMF) which could integrate the standard DA method, data disrupting method, data mixing method, and autoadjustment method to enhance the image data in the training dataset to create new training images. In addition, we fused the results of the classifier by voting on the basis of DAMF, which further improved the generalization ability of the model. The experimental results showed that the optimal DA mode could be matched to the image dataset through our DA strategy. We evaluated DAMF on the iChallenge-PM dataset. At last, we compared training results between 12 DAMF processed datasets and the original training dataset. Compared with the original dataset, the optimal DAMF achieved an accuracy increase of 2.85% on iChallenge-PM.

1. Introduction

Pathologic myopia (PM) is one of the major causes of visual impairment worldwide [1–3]. As myopia deepens, it is often accompanied by deforming changes in the posterior pole of the eye [3]. These changes are usually manifested as sclerotic atrophy, choroidal atrophy, and growth of the ocular axis, which may be associated with various complications of the eye, such as amblyopia, glaucoma, cataracts, vitreous clouding, and retinal detachment [1]. Complications of pathological myopia are considered to be the main reasons for visual impairment and blindness today, particularly in East Asia [4, 5]. Pathological myopia causes visual impairment due to various pathologies of the macula, peripheral retina, and optic nerve. Structural deformities of the eye, including posterior sclerite uveitis, may accelerate the progression of these diseases [1, 3, 6–9].

According to a summary of 145 studies regarding the global prevalence of myopia and PM, there are approximately 1950 million people with myopia (accounting for 28.3% of the global population) and 277 million people with PM (accounting for 4.0% of the global population), and these numbers are predicted to increase to 4758 million (accounting for 49.8% of the global population) for myopia and 938 million (accounting for 9.8% of the global population) for PM by 2050 [5].

The prevalence of high myopia and pathological myopia shows an increasing yearly trend due to the changes in environmental factors and lifestyle [10–14]. In China, the number of ophthalmologists differs significantly from that in developed countries, but the demand for ophthalmologists

in China is already huge; with the growth of social aging, the number of ophthalmologists has been rising in recent years [10, 12, 14]. In consequence, the visual impairment caused by pathological myopia complications will become increasingly serious in the coming decades. The people's growing medical needs contrast sharply with the increasingly deficient medical resources in the current medical field.

On the one hand, the problem of "difficult and expensive access to medical care" still exists, which is mainly caused by the severe scarcity of talented physicians and the long training period for doctors. On the other hand, with the health problems gradually worsening, people are paying more and more attention to their health, which aggravates the demand for medical services. This is a social problem to which AI technology can offer the medical industry a solution [15–18].

Over the past two decades, with the development of imaging techniques, such as optical coherence tomography, frequency-domain OCT, and 3D magnetic resonance imaging, the complications associated with high myopia have been well known [16, 19–23]. For example, the optic nerve, macula, and neoplastic lesions can be magnified by OCT to extraordinary resolution for evaluation. In addition, myopic retractile macular lesions and domed macular lesions can be utilized in the same way. The advent of new therapies, including antineovascular drug therapies and vasectomy, has led to an improved prognosis for some of the complications associated with high myopia.

Medical artificial intelligence (AI) technologies have been well developed in recent years [20]. A case in point is the application of machine learning-based AI technology to ophthalmology [20, 21]. The diagnosis of many eye diseases relies heavily on the results of ophthalmic examinations, most of which are built on imaging studies. Eye images are delicate, complex, and informative, and diagnostic results are heavily dependent on the doctor's knowledge and clinical experience, which makes diagnosis subjective and time-consuming [24, 25]. The development of medical AI has significantly improved the efficiency of ophthalmic disease diagnosis in clinical work and reduced the burden on ophthalmologists [20–22].

CNNs, which are data-driven and can automatically extract relevant features, have secured better results in image recognition than traditional methods [26]. Therefore, it is considered to be a new choice to introduce CNN techniques into medical image processing. Lately, many studies have been conducted on this matter and applied CNNs to fundus image recognition, whose results generally surpass those based on traditional recognition ways [19, 26–31].

Although the CNN-based fundus image recognition method outperforms traditional methods to a certain extent, there are still some problems. For example, the amount of medical image data is large, while the number of positive samples is small. Nevertheless, a common fact is that the training process of the model mainly depends on the dataset, namely, the training effect of the model can be well improved by the dataset after DA processing [32].

Furthermore, with the continuous development of deep learning in the field of images, CNNs have been increasingly complicated. Each model has its unique advantages, but it is not guaranteed that every aspect of the model will perform well. For instance, the model's lack of expressive ability will lead to the weakness in the recognition of some rare lesion images [32]. To address this problem, researchers have proposed optimized neural network models from different perspectives and achieved effective results. However, there are a lot of difficulties in enhancing the existing models. For example, when the researchers optimize the models (such as widening and deepening models), they cannot predict the effectiveness of the models but just observe whether the optimization operation improves the performance of the original models through the training results. Besides, even if such optimization is effective, it may be computational and time-intensive or needs a long development cycle, so it cannot address the problem effectively [33–35].

In order to deal with these issues, the main contributions of this paper are described as follows:

(1) 12 DAMFs will be designed based on the iChallenge-PM dataset. To our knowledge, these DA methods cover all the operations used in the current DA. The purpose is to increase data characteristics, suppress sample imbalance, and effectively improve the quality of datasets.

(2) Based on AlexNet, VGG-16, GoogLeNet, and ResNet-50 models, different optimizers, loss functions, and learning rates are constructed. Then, the model with the highest accuracy will be adopted as the primary learner and trained based on 12 datasets. The generalization ability of the model will be enhanced through this method.

(3) The abovementioned primary learner prediction will be used as a new input and put into the secondary learner, namely, the hard voting model, and then the fusion model is trained to form the final model.

(4) The model optimized by the above operations achieves high accuracy without transfer learning. More importantly, by using the augmented dataset and the model fusion method, we effectively avoid overfitting and improve the generalization ability of the model when processing various data, which further improves the expressive ability of the model. As a result, the accuracy of the model in recognizing complex and rare case images will be effectively improved.

2. Literature Review

In this section, we mainly reviewed the related literature on automatic disease diagnosis from fundus photography and DA.

2.1. Automatic Disease Diagnosis from Fundus Photography. Most of the conventional ophthalmic diseases can be examined from fundus photography, including PM, age-related macular degeneration (AMD), diabetic retinopathy (DR), and glaucoma. Conventional diagnosis methods tremendously depend on doctors' professional experience and

knowledge, which results in a high misdiagnosis rate and a huge waste of medical data [36, 37]. The deep integration of ophthalmology and AI offers potential for revolutionizing current disease diagnosis patterns and generating a significant clinical impact. As for PM classification, Freire et al. employed Xception as the baseline architecture with ImageNet pretrain weights to diagnose PM from fundus images [38]. Zhang et al. used the feature selection of PM valuable information in the images to improve the training effect of the model [31]. They demonstrated that the new method was much efficient by using less than 25% of the initial candidate feature set. Not only in the field of image classification but DA also occupies a highly important position in the field of target detection. Sun et al. put forward two new DA modules consisting of channel-wise random Gamma correction and channel-wise random vessel augmentation [27]. They argued that their method could ameliorate the performance and robustness of a classic CNN architecture. However, most medical imaging samples are unbalanced. Although there may be a large number of samples, the types of samples are limited. Different from previous works, in this paper, we focused on the first stage of deep learning, which was DA. And by means of this strategy, the defect of data imbalance could be effectively solved and the overfitting in the training process could be suppressed.

2.2. Data Augmentation. DA highlights the characteristics of image data and prevents overfitting in the training effect [33]. In existing studies, researchers divide DA methods into the standard method, data disrupting method, data mixing method, DA method based on reinforcement learning, and fusion-based image augmentation method.

(1) *Standard Data Augmentation Method.* AlexNet [39] preprocesses the data using random cropping and horizontal flipping, and it has been verified on CIFAR-10. Random cropping prevents CNN from overfitting specific features by changing the obvious features in the image. Facebook artificial intelligence research uses another color translation method, called color dithering, to improve the process of ResNet [40]. Color dithering randomly changes the brightness, contrast, and saturation of the image. These DA techniques play an important role in the training of a model. Fu et al. performed extensive experimentation on a large set of images with varying illuminations [41]. The performance is analyzed both quantitatively and qualitatively. However, as the number of parameters increases, the risk of overfitting also increases. Many studies have proposed more complex CNNs. Therefore, more powerful DA strategies are particularly important.

(2) *Data Disrupting Method.* [42] Contrary to (1), a data disrupting method produces unnatural images by destroying images' features. Dropout on the input layer is a DA technique that disturbs and masks the original information of given data by dropping pixels. Pixel dropping functions as injection of noise into an image. It makes the CNN robust to noisy images and contributes to generalization rather than enriching the dataset, randomly erasing a region in an image at every training step. It is an extension of dropout, where the masking of regions behaves like injected noise and makes CNNs robust to noisy images. Under this condition, CNNs need to learn other parts that are usually ignored.

(3) *Data Mixing Method.* This is a special case of (2), where a mixup alpha blends two images to construct a new training image [43]. Mixup can train CNNs on convex combinations of pairs of training samples and their labels and enables CNNs to favor a simple linear behavior in-between training samples. This behavior makes the prediction confidence transit linearly from one class to another class, thus providing smoother estimation and margin maximization. Therefore, the mixup makes CNNs robust to adversarial examples and stabilizes the training of generative adversarial networks.

(4) *Data Augmentation Method Based on Reinforcement Learning* [44]. Autoenhancement is a framework based on reinforcement learning to explore the optimal augmentation combination [45]. Hence, it is not a DA method but is an external framework. It achieves significant results during the CIFAR-10 classification and proves the contribution of reinforcement learning to DA research.

(5) *Fusion-Based Image Augmentation Method.* Fusion is preferred to the direct application of traditional techniques since it involves an amalgamation of traditional techniques, rather than the application of a single technique. Fusion can be done in various ways, out of which multiscale fusion has proved to be one of the best. Parihar et al. presented a detailed analysis of image enhancement techniques based on multifusion, thereby giving an insight into the algorithm used in each method, along with its implementation framework [36, 42].

In this paper, we adopted the concept of (5) and integrated the image enhancement techniques of (1), (2), and (3), aiming to enlarge the features of images from the perspective of color, direction, shape, and so on.

3. Materials and Methods

In this session, we first introduced the dataset of iChallenge-PM [46] composed of 1200 annotated retinal images. In another part of this section, we introduced the primary learner, which included some classical convolutional neural networks and corresponding components. These frameworks have become general arrangements for image classification and have been extensively utilized in different computer vision tasks.

3.1. iChallenge-PM. The iChallenge-PM dataset contained 1200 annotated color fundus photos with non-PM (50%)

and PM (50%) cases. Specifically speaking, the dataset contained 400 images in the training set, validation set, and test set, respectively, while the test set was not public. We reclassified the 800 public image data and divided the validation set into a new validation set and a test set according to the ratio of 50% : 50%. All images data were reshaped to a size of 224 * 224 before using DA. Under the partitioning of iChallenge-PM, the reference standard of PM presence was obtained from the health records, which was not based solely on fundus image but also took OCT, visual test, and other factors into consideration. For the training data, PM, HM, and normal labels were reflected in the image file names, with 0 denoting normal cases while 1 denoting abnormal cases. However, we did not utilize any human-annotated labels during network training. To evaluate the effectiveness of our method, we employed accuracy and loss rate as the evaluation indices. Accuracy was often the most important index for doctors and patients.

3.2. Classifiers and Components

(1) *AlexNet*. The AlexNet network proposed by Krizhevsky et al. was the first to use five convolutional layers and three fully connected expenses to achieve the classification of 1000 classes of images, thus becoming the seminal breakthrough in image classification based on deep learning. Compared with other traditional convolutional neural networks, AlexNet applied various methods to improve deep convolutional networks. For example, the rectified linear unit (ReLU) nonlinear activation function was used to speed up the training of the network, multi-GPU convolutional operations were implemented to address the limitations of insufficient graphic card resources at the time, and the DropOut random inactivation strategy was introduced to reduce overfitting at the full connection layer. Furthermore, strategies such as local response normalization, overlap pooling, as well as augmentation, were proposed by AlexNet to improve the classification and generalization capabilities of the model.

(2) *VGG*. The VGG network was proposed by Simonyan et al., in which filters of 5 * 5 and 7 * 7 were replaced by filters of 3 * 3. It was a basic idea embodied by VGG that the receptive fields of multiple small convolutional layers in series could be in the same size as that of a large convolutional layer. For example, two 3 * 3 convolutional fields in sequences had the same field size as one 5 * 5 convolutional field, and their convolutional effects were the very same. However, multiple small convolutional layers concatenated together had fewer parameters and more nonlinear transformations. This was most effective in better learning the features. At the same time, VGG also increased the network structure to 16 or 19 layers. As the number of layers increased, the network enjoyed better feature representation and better model classification. Being simple and effective, VGG was still commonly used in the field of computer vision for image classification, detection, segmentation, super-resolution, and image styling.

(3) *GoogLeNet*. Google made a significant contribution to the development of deep convolutional neural networks through the proposed inception family. The most significant contribution of Inception-V1 (GoogLeNet) was to propose the inception structure while deepening the depth of the convolutional neural network. The structure increased the width of the network by concatenating multiple convolutional blocks of different sizes, which allowed the convolutional blocks to acquire information from discrete receptive fields. In addition, the structure took full advantage of the 1 * 1 convolutional code to reduce a large number of network parameters, thereby improving the efficiency of computing resources. Inception-V2 proposed an excellent regularization method, namely, batch normalization, which made the data undergo a back normalization process before each convolution and was now the standard for deep convolutional networks. This approach was an excellent solution to the training problem of multilayered networks. The evolution of the inception structures is shown in Figure 1.

(4) *ResNet*. Kaiming He et al. established a deep residual network, namely, ResNet [14], which increased the network depth to 152 layers while ensuring the network accuracy and then further increased the depth to 1000 layers. Theoretically, the deeper the network was, the higher the accuracy should be. However, the authors experimentally found that blindly increasing the depth would lead to the degradation of the network when the depth reached a certain level. The gradient explosion and gradient disappearance of deep networks failed to train the model correctly and led to poor network performance. Inspired by highway network, the authors proposed the residual structure by adding a jump connection between the input and output of the convolutional block, which enabled the input to be passed directly to the output. The residual structure was essentially designed to learn a constant mapping with the nonlinear layer portion of the stack learning another mapping. As shown in Figure 2, if the residual structure was zero, we could easily train a constant mapping. In short, if a network could achieve the desired result by simply setting parameters manually, it was not difficult to train the network to converge to that result so that the added residual structure would at least not degrade the overall performance of the network. ResNet's residual module reduced the difficulty in training deep networks, solved the degradation problem well, and maximized the depth potential of convolutional networks. Eventually, ResNet outperformed human

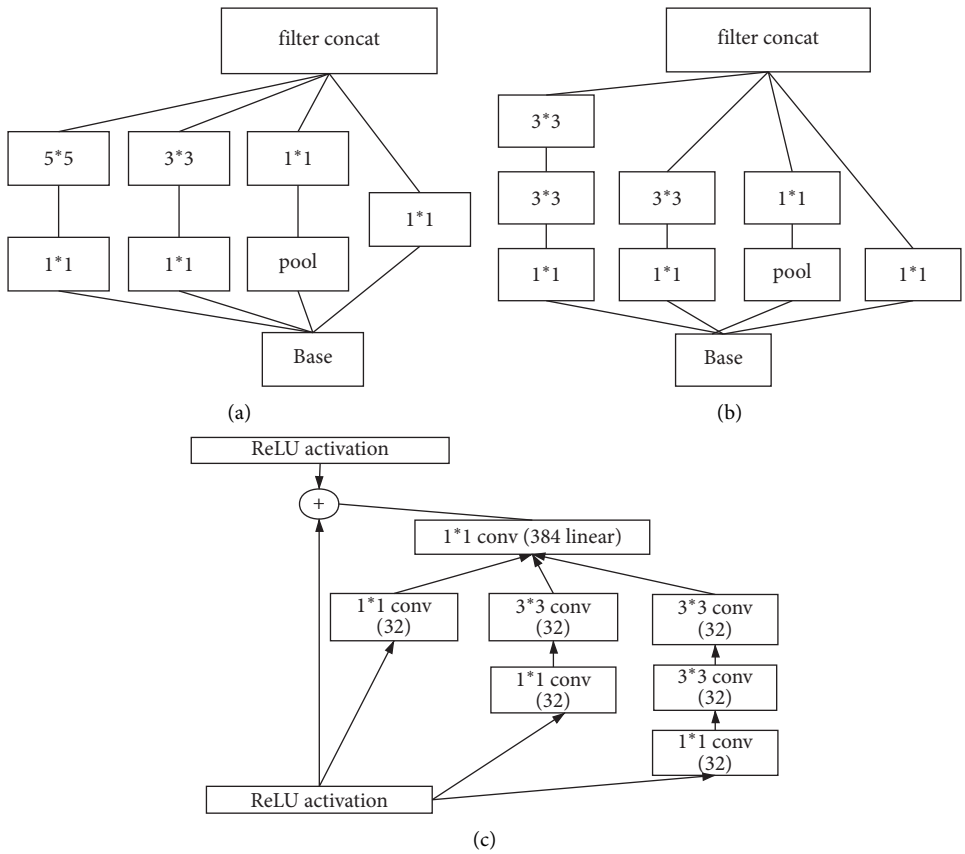


Figure 1: Evolution of inception structures: (a) Inception-V1 chart; (b) Inception-V2 chart; (c) Inception-V3 chart.

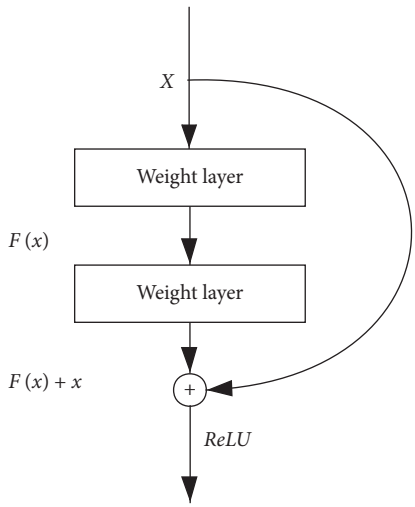


Figure 2: Residual structures in the ResNet network.

performance in terms of the ImageNet classification task for the first time.

(5) *Optimizer*. SGD [47] experienced difficulty exploring gorges. For example, territories, where the surface bent considerably more steeply in one measurement than in another, were basically around neighborhood optima. In these situations, SGD swayed over the gorge's slants and slowly advanced along the valley floor towards the local optimal direction, as indicated in Figure 3(a). The energy was a technique that quickened SGD and hose motions in a pertinent way, as shown in Figure 3(b). As suggested by formula (1), it achieved this by including a part γ of the update vector of the past time venture to the current update vector.

$$v_t = \gamma v_{t-1} + \eta \nabla_\theta J(\theta), \quad \text{where } \theta = \theta - v_t. \tag{1}$$

The energy term γ was usually set to 0.9 or a comparative value. In general, when utilizing energy, we pushed a ball down a slope. The ball collected energy as it moved downhill, getting quicker and quicker in transit (until it arrived at its max speed, if there was air obstruction, for example, $\gamma < 1$). Something very similar happened to our boundary refreshing: the momentum term increases for the latitude where the gradient points in the same direction, while for the dimension where the gradient changes direction, the momentum term will decrease and update. Accordingly, we increased quicker union and diminished swaying.

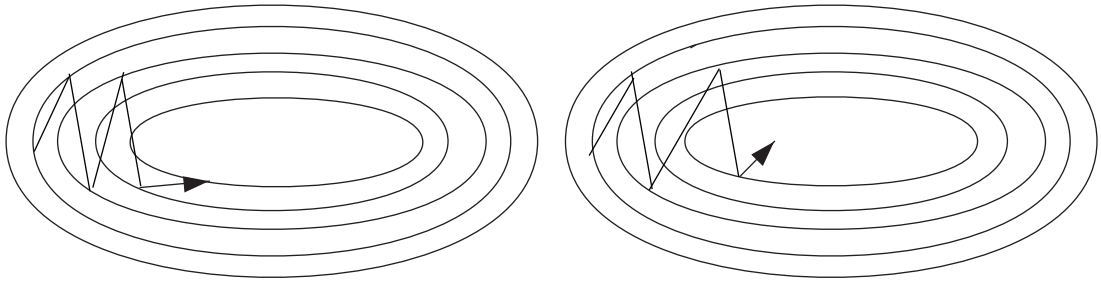


Figure 3: Source: Genevieve B.Orr: (a) SGD without momentum; (b) SGD with momentum.

(6) *Optimizer*. The logistic loss could be calculated by the following formula:

$$\text{loss} = -\text{Labels}^* \log(\text{sigma}(X)) - (1 - \text{Labels})^* \log(1 - \text{sigma}(X)), \quad \text{where sigma}(x) = \frac{1}{1 + \exp(-x)}. \tag{2}$$

After applying it to the above calculation, we got logistic loss formula as follows:

$$\text{loss} = X - X^* \text{Labels} + \log(1 + \exp(-X)). \tag{3}$$

In order to calculate stability and prevent overflowing, the loss function would be calculated using the following formula:

$$\text{loss} = \max(X, 0) - X^* \text{Labels} + \log(1 + \exp(-|X|)). \tag{4}$$

3.3. Voter Model. Voting was a combination strategy aimed at classification problems in ensemble learning. The basic idea was to select the class with the highest output among all machine learning algorithms. There were two types of output judging from a machine learning classification algorithm: one was the direct output of class labels and another was the output of class probabilities. Using the former for voting was called majority/hard voting while using the latter for classification was called soft voting. Hard voting selected the label with the most output of the algorithm. If the number of labels was equal, the selection was made in ascending order. Soft voting used the class probabilities output by each algorithm to select a class. If the weight was input, a weighted average of the class probabilities of each class would be obtained, and the class with a large value would be selected. In this paper, our experiment used the hard voting mechanism.

3.4. Data Augmentation Strategy. Binary coding was used to represent positive and negative samples in this paper. Existing common DA methods consisted of randomly flipping the image (horizontally or vertically), randomly adding noise, rotating the image, changing the brightness, contrast, and saturation of the image, randomly cropping the image, randomly scaling/stretching the image, and randomly changing the clarity of the image. All these methods belonged to (1), (2), and (3) of Section 2.2. Based on these DA methods, we performed 12 different DAMF combinations and made 12 new datasets, as listed in Table 1. Specifically, third-party DA libraries were used in the 11th DAMF and 12th DAMF, respectively. As shown in Figure 4, all the images in the original dataset represented images without corresponding augmentation. The enhanced images were displayed following each original image. Figure 4(b) shows the randomly rotated operation, with the rotation angle at 90/180/270/360 degrees. Figure 4(d) shows the consequence of randomly adding Gaussian white noise to the original operation. Figure 4(f) describes the operation after random adjustment of brightness, saturation, and contrast built on the original image. Figure 4(h) describes the random cropping and stretching based on the original image. Figure 4(j) displays the image after randomly adjusting the sharpness. Figure 4(l) displays the image after randomly adjusting the contrast, saturation, and brightness on the original image and adding random Gaussian white noise. Figure 4(n) shows a randomly rotated, cropped, and stretched image built on the original image. Figure 4(p) shows the effect of randomly superimposing the images after all operations on the original images. Figure 4(r) shows the effect of utilizing the third-party library imaging to mutate the image. Figure 4(t) shows the effect of randomly superimposing all the above special effects.

3.5. Primary Learner Model. AlexNet, GooLeNet, VGG-16, and ResNet-50 were used as the primary learners in this paper. The experiment set the learning rate to 0.001 and utilized the optimizer and loss function described above. Each model was formed for 30 epochs, with each epoch covering all images in the training set.

The highest accuracy of each model was selected after comprehensive training, and the corresponding model parameters were saved. Then, the convolution layer of all primary learners would be frozen, which meant that the data could only be transmitted forward instead of backward after entering the primary learners.

Table 1: 12 DAMFs.

No.	Training set name	DA method	Quantity
1	PALM-Training800-overturning	Original dataset + random flip (4 directions: up, down, left, and right)	800
2	PALM-Training800-noise	Original dataset + Gaussian white noise	800
3	PALM-Training800-color	Original dataset + randomly changing colors (brightness, contrast, saturation)	800
4	PALM-Training800-cropping	Original dataset + random cropping	800
5	PALM-Training800-deforming	Original dataset + random scaling, stretching (stretched into a square by the length or width of the images)	800
6	PALM-Training800-dimming	Original dataset + change clarity	800
7	PALM-Training1600-overturning-noise-color	Randomly stack method 3 or 4 (serial number) on the basis of PALM-Training800-overturning	1600
8	PALM-Training1600-overturning-cropping-deforming	Randomly stack method 5 or 6 (serial number) on the basis of PALM-Training800-overturning	1600
9	PALM-Training1600-overturning-dimming	Randomly stack method 7 (serial number) on the basis of PALM-Training800-overturning	1600
10	PALM-Training3200-overturning-noise-color-cropping-deforming-dimming	Randomly superimpose method 5 or 6 or 7 (serial number) on the basis of PALM-Training800-overturning-noise-color	3200
11	PALM-Training800-imgaug1	Original dataset + random cropping with 0–50 pixels around, 50% probability horizontal flip, Gaussian blur (sigma = 0 to 3.0)	800
12	PALM-Training1600-overturning-dimming-imgaug2	PALM-Training800-overturning-dimming dataset + multiple mixed random overlay	1600

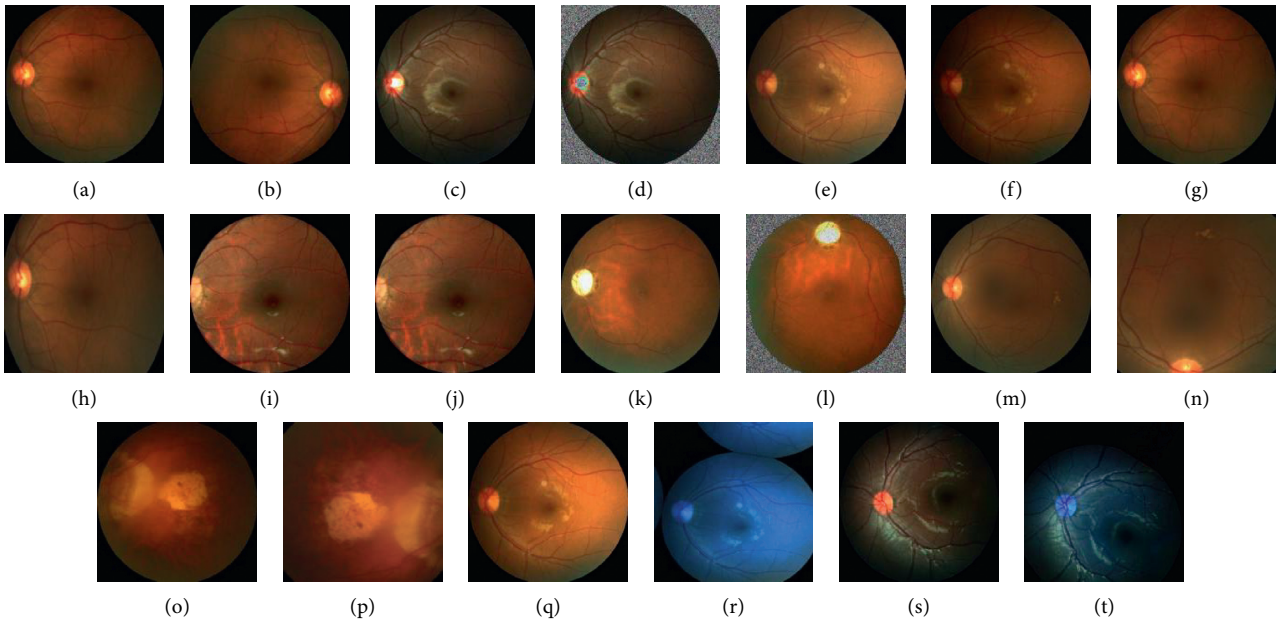


Figure 4: DA effect (original dataset image$_x$: different images in the original dataset: (r), (s), and (t) were mutated from (q)). (a) Original dataset image$_1$. (b) Randomly change direction. (c) Original dataset image$_2$. (d) Randomly add Gaussian noise. (e) Original dataset image$_3$. (f) random color. (g) Original dataset image$_4$. (h) Random stretching. (i) Original dataset image$_5$. (j) Randomly adjust the sharpness. (k) Original dataset image$_6$. (l) Randomly flip, adjust colors, and add Gaussian noise. (m) Original dataset image$_7$. (n) Randomly flip, adjust colors, and add Gaussian noise. (o) Original dataset image$_8$. (p) Random stretching and cropping. (q) Original dataset image$_9$. (r) Random cropping with 0–50 pixels around, 50% probability horizontal flip, and Gaussian blur (sigma = 0 to 3.0). (s) Original dataset image. (t) Randomly stack all operations.

3.6. Staking Model Integration Strategy and Hard Voting Model. Staking, as a hierarchical model integration framework, was one of the main strategies widely used in model integration. Taking two layers as an example, the first layer consisted of multiple base learners, and the original training set was the input of the primary learners. The output of the primary learners was treated as the secondary learners' input, which was the training set of secondary learners. The secondary learners continued training on the above training set to obtain the complete staking model.

As shown in Algorithm 1, processes 1–3 constructed the trained primary learners. Processes 5–9 were the prediction results of the training set using the trained primary learners, and this prediction was used as the training set for the

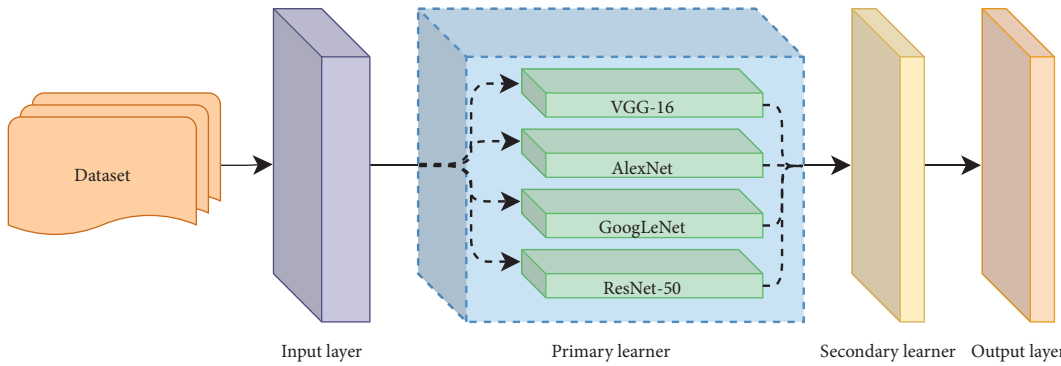


Figure 5: Logic diagram of the fusion model.

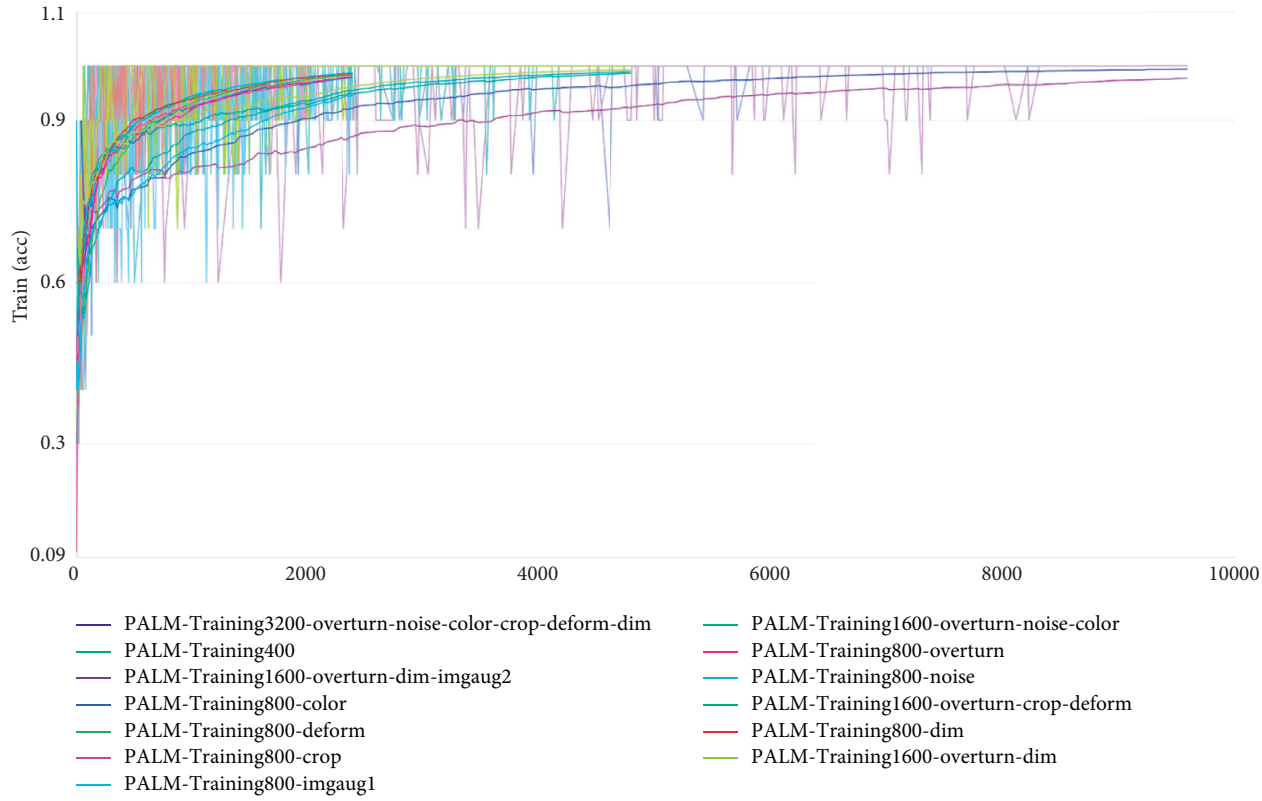


Figure 6: Accuracy of VGG-16 for 30 epochs of training on 13 datasets (original dataset: 2_{num} and DAMF datasets: 1_{num} and $3–13_{num}$).

secondary learners. Process 11 used the prediction results of the primary learners to train the secondary learners to get the fusion model.

The design facilitated the extension of the model. In other words, the hard voting model could be replaced with other secondary learners based on different datasets. Figure 5 displays the framework of the fusion model in this paper. First, the training set normalized the data through the input layer, namely, processed it into a format of the same size ($224 * 224$). Next, the images were input into each classifier, respectively, in the primary learner. The classifier performed 30 times epoch supervised training on the image according to the label of all inputs and then used the training result as the input of the secondary learning, and the final classification result was obtained after voting.

4. Results

4.1. Lab Environment. Hardware environment is as follows: CPU 4 cores, RAM 32 GB, GPU v100, video memory 16 GB, and disk 100 GB.

Environment configuration is as follows: Python version python3.7 and framework version PaddlePaddle 1.8.0.

4.2. Evaluation Indices. The primary reference record was the accuracy of the model forecast. In this paper, recall rate, specificity, and sensitivity were not used as evaluation

Model input:
Training set $D = \{(x_1, y_1), (x_2, y_2), \ldots, (x_m, y_m)\}$
Primary learner $£_1, £_2, \ldots, £_T$
Secondary learner £
Process:
1: **for** $t = 1, 2, \ldots, T$ **do**
2: $h_t = £_T(D)$
3: **end for**
4: $D = \emptyset$;
5: **for** $t = 1, 2, \ldots, m$ **do**
6: **for** $t = 1, 2, \ldots, T$ **do**
7: $Z_{it} = h_t(x_i)$;
8: **end for**
9: $D_j = D_j U ((z_i 1, z_i 2, \ldots, z_{iT}), y_i)$
10: $h_j = (D_j)$
11: **end for**
Output: $H(x) = h_j(h_1(x), h_2(x), h_3(x), \ldots, h_T(x))$

Algorithm 1: Logic diagram of the model fusion algorithm.

Table 2: VGG-16 training results on 13 datasets.

No.	Dataset	Accuracy	Loss
1	PALM-Training1600-overturning-dimming-imgaug2	0.95858336	0.18674079
2	PALM-Training3200-overturning-noise-color-cropping-deforming-dimming	0.95550001	0.27185006
3	PALM-Training1600-overturning-cropping-deforming	0.95266668	0.16523545
4	PALM-Training800-color	0.95033336	0.17019135
5	PALM-Training800-dimming	0.94875002	0.17919912
6	PALM-Training800-cropping	0.94625	0.18303553
7	PALM-Training1600-overturning-dimming	0.94525003	0.23124305
8	PALM-Training1600-overturning-noise-color	0.94008333	0.21350351
9	PALM-Training800-overturning	0.93858335	0.20894363
10	PALM-Training800-deforming	0.93708334	0.20814224
11	PALM-Training800-noise	0.93608335	0.26124661
12	PALM-Training800-imgaug1	0.93391667	0.19876853
13	**PALM-Training400**	**0.93016667**	0.19310093

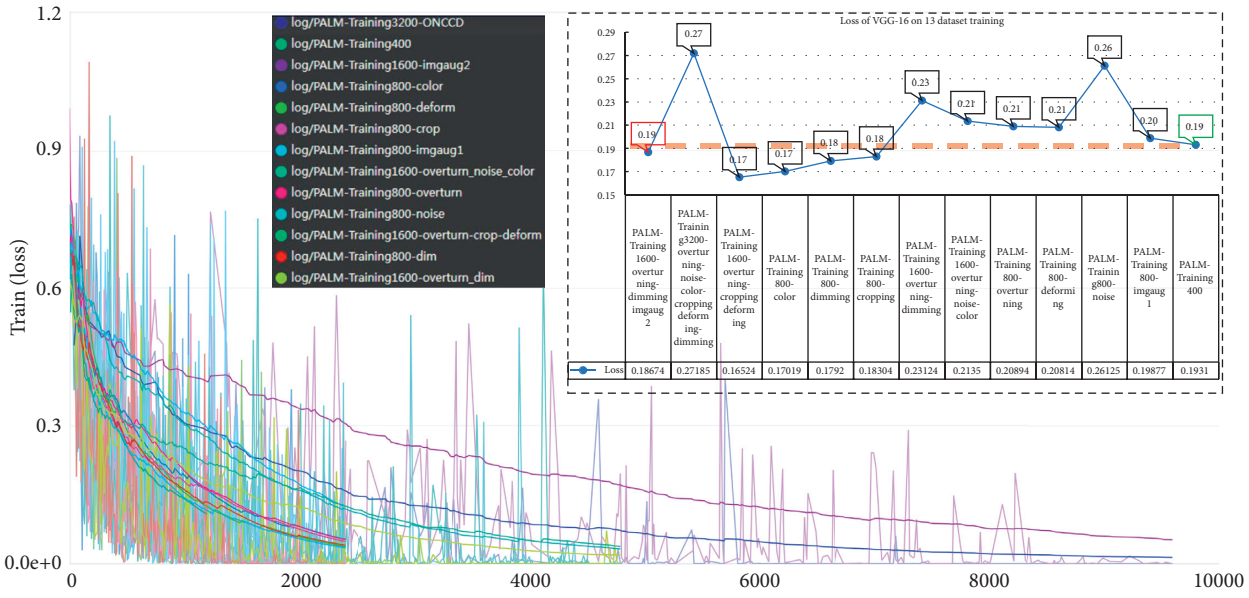


Figure 7: Loss of VGG-16 for 30 epochs of training on 13 datasets.

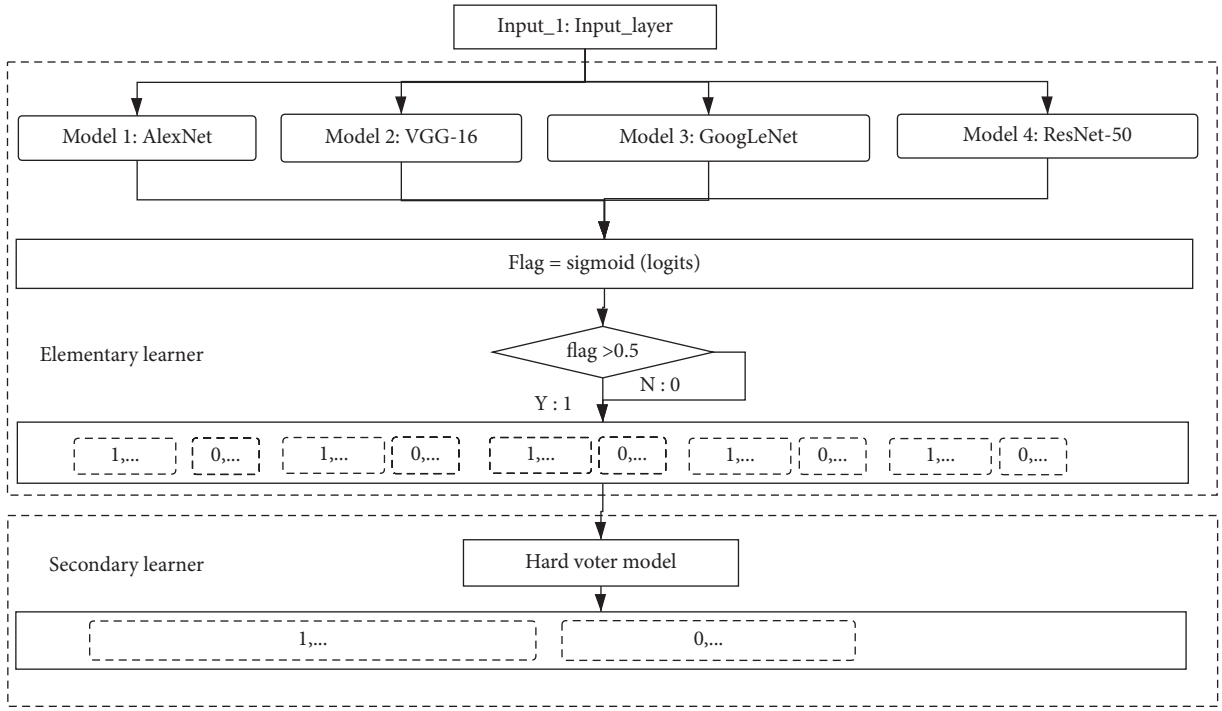


Figure 8: Logic diagram of the fusion model.

indices. The classification performance was mainly evaluated by the classification accuracy, which was defined as follows:

$$\text{accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}, \tag{5}$$

where TP, TN, FP, and FN denoted the true positive, the true negative, the false positive, and the false negative, respectively.

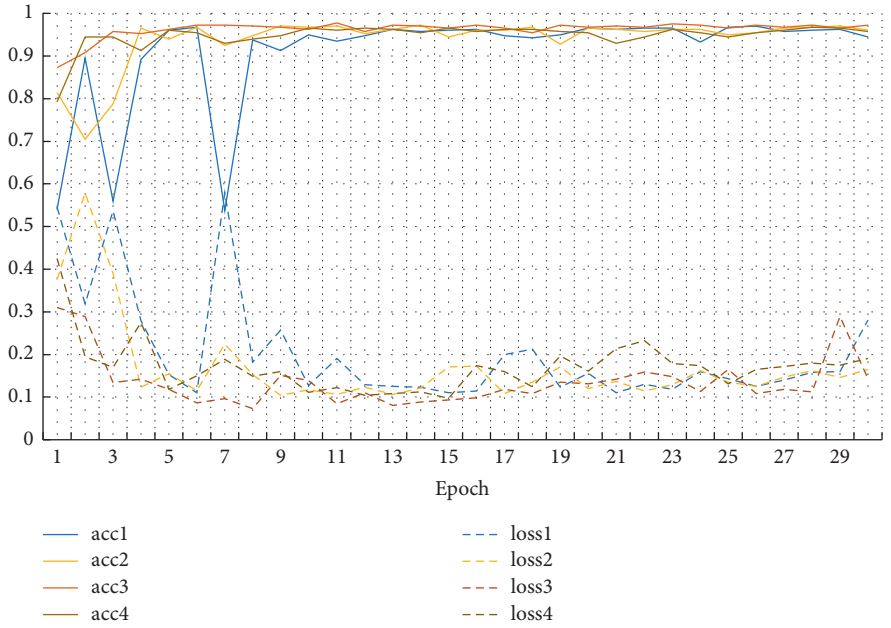
The loss function in the model was measured by the root mean square error, namely, a risk metric corresponding to the expected value of the squared (quadratic) error or loss. If $\hat{y}_i$ was the predicted value of the i-th sample and y_i was the corresponding true value, then the mean squared error (MSE) estimated over n_{samples} was defined as follows:

$$\text{MSE}(y, \hat{y}) = \frac{1}{n_{\text{samples}}} \sum_{i=0}^{n_{\text{samples}}-1} (y_i - \hat{y}_i)^2. \tag{6}$$

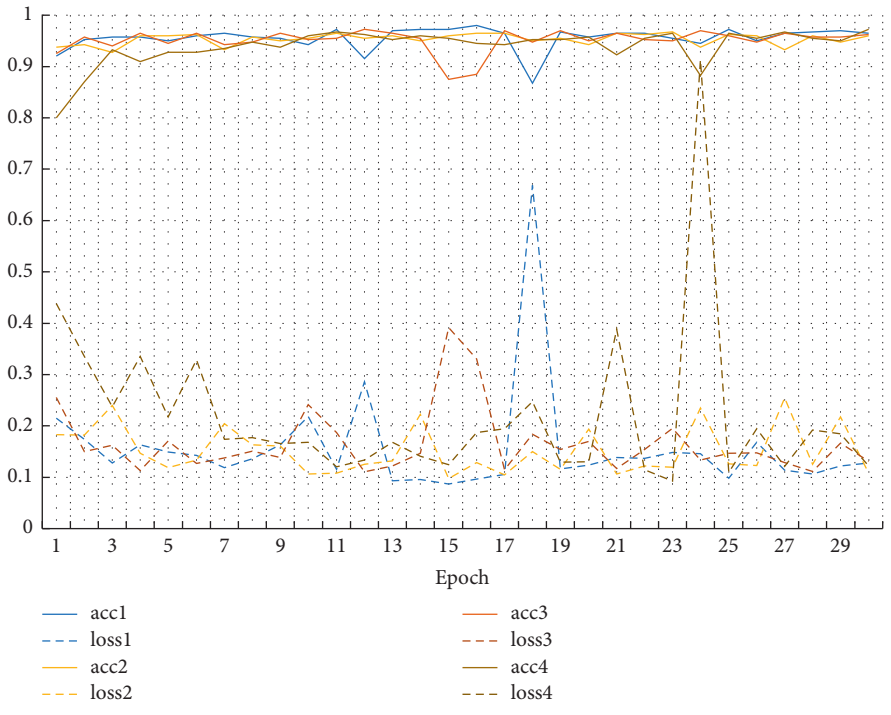
4.3. Primary Learner Training Process and Results. As shown in Figure 6, firstly, VGG-16 was used as a dataset filter, and training was conducted on all datasets. Each dataset was trained for 30 epochs. Each epoch would traverse all the datasets once to form the corresponding trained models on different datasets. These 13 datasets were adopted to make predictions on the test set, and the final results are listed in Table 2. According to Table 2, the overall accuracy of the enhanced dataset was higher than that of the original dataset. To go into detail, the average accuracy of PALM-Training1600-overturning-dimming-imgaug2, PALM-Training3200-overturning-noise-color-cropping-deforming-dimming, PALM-Training1600-overturning-cropping-deforming, and PALM-Training800-color exceeded 95%. Therefore, these 4 datasets were selected to be the candidate datasets. GoogLeNet, AlexNet, and ResNet-50 were also trained on these 4 datasets. Each model was trained on each dataset for 30 epochs, and then the training model was tested on the test set.

Follow-up training was made on the candidate datasets, and the results are shown in Table 3. We used the same parameters during the process of training. It could be seen from Table 3 that the optimal DAMF datasets also varied because of the differences in the expression ability of different models. To be specific, GoogLeNet and ResNet-50 were both trained on the PALM-Training3200-overturning-noise-color-cropping-deforming-dimming dataset. AlexNet and VGG-16 had the highest scores on the PALM-Training1600-overturning-dimming-imgaug2 dataset. The above four models with the highest accuracy were used as primary learners.

In this paper, the accuracy rate was the average of the accuracy rate results of 30 epochs of training. The highest accuracy of the model training set reached 100% (Figure 6). Table 2 still shows the comparison between applying DAMF and not applying DAMF, in which the 13th dataset is the original dataset. We can see that the results of all the datasets processed by DAMF were better than the results of the 13th dataset, and the best DAMF corresponding to the first group of results is 2.84% higher than the 13th group, reached 95.85%. This improvement is obvious. Similarly, we have observed AlexNet, GoogLeNet, and ResNet-50. From the results for training, the accuracy rates of these were, respectively, 95.76%, 96.24%, and 95.60%, which are highly similar to the best training results of VGG-16, i.e., all exceed 95.00%. The above results could show that DAMF is universal.

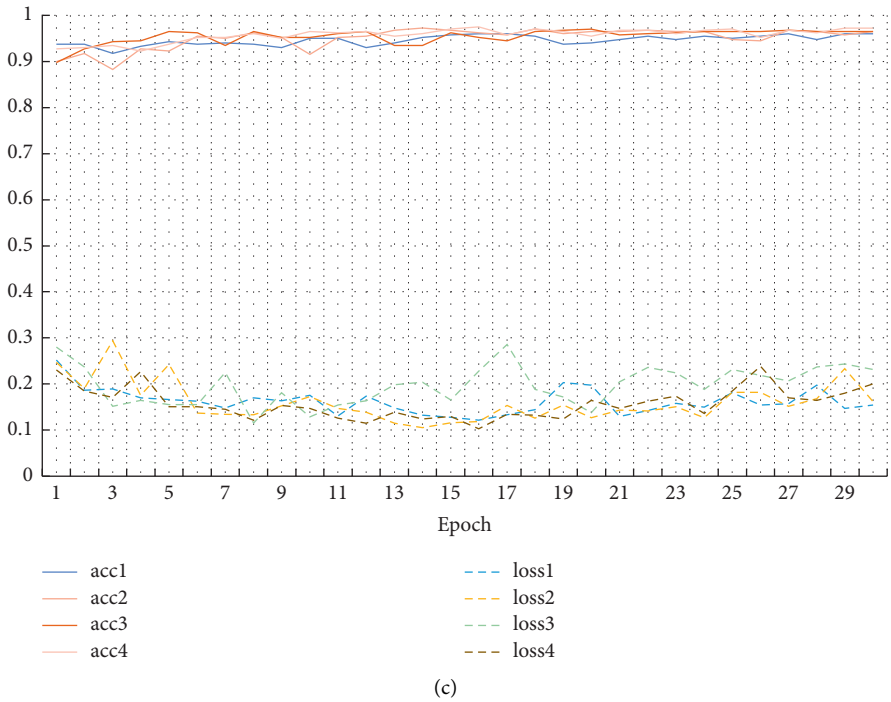


(a)



(b)

Figure 9: Continued.



(c)

Figure 9: The accuracy and loss rate of three primary learners in training process. Four best training results of (a) AlexNet; (b) GoogLeNet; (c) ResNet-50.

Figure 7 displays the loss trend during the training process of VGG-16. Taking the loss rate of the training original dataset as a baseline for comparison, we could find that although some DAMF loss rates were higher than 0.19, the highest loss rate was 0.27, which means the loss rate was acceptable. Among them, the loss rate corresponding to the best DAMF is 0.19, which was the same as the original dataset, which means that the enhanced data did not cause additional loss.

4.4. Fusion Model Training Process and Results. Figure 8 displays the logic diagram of the fusion model, with the most accurate models (AlexNet, GoogLeNet, ResNet-50, and VGG-16) used as the primary learners. The predictions of all primary learners were used as the training dataset for secondary learners. Meanwhile, the original dataset label was treated as the label of the new dataset to build the training set of the secondary model. After the primary model, the hard voting model was inserted as a classifier to form the framework of the secondary learners. In the secondary model, the four prediction results (AlexNet result, GoogLeNet result, ResNet-50 result, and VGG-16 result) in each sample were counted as the final prediction results. After 30 epochs of training, the model was saved and validated on a test set. The final accuracy of the fusion model reached 97.25% (average accuracy), with a maximum accuracy of 98.00%.

5. Discussion

First, 12 DAMF strategies were implemented on the iChallenge-PM dataset, resulting in the formation of 13 datasets, including the original one. Then, the experiment used VGG-16 as a dataset picker to train each of these 13 datasets for 30 epochs, each epoch covering all the data once. The accuracy of the model after training on each dataset was obtained on the validation set. As universally agreed, the four datasets with the best accuracy were selected as the preselected datasets to be used in the training of all remaining models. At the end of the training, the model with the highest prediction accuracy on the validation set was chosen as the primary learner.

Through this experiment, the advantages of DA were evident, and the data augmented datasets generally obtained higher accuracy than that of the original dataset. The average accuracy of the VGG-16 models trained by the four preselected datasets mentioned above was 95.85%, which was 2.84% higher than that of the original dataset.

This study adopted the strategy of model integration. The experiment retrained the output of the first-level model. The average accuracy of the first-level model was 95.86%, and the prediction accuracy of the fusion model was once again improved by 1.39%. Particularly, the greater significance of the fusion model was that the shortcomings of each primary learner were balanced. Therefore, the generalization performance and expression of the model were effectively improved.

We could observe the performance of DAMF in the primary learners again. Figure 9 displays the training process of AlexNet, GoogLeNet, and ResNet-50 on the four optimal DAMF datasets. The details were expressed as follows: the solid line part in each subfigure referred to the accuracy, the dotted line part referred to the loss rate, the abscissa represented the training epoch, and the ordinate represented

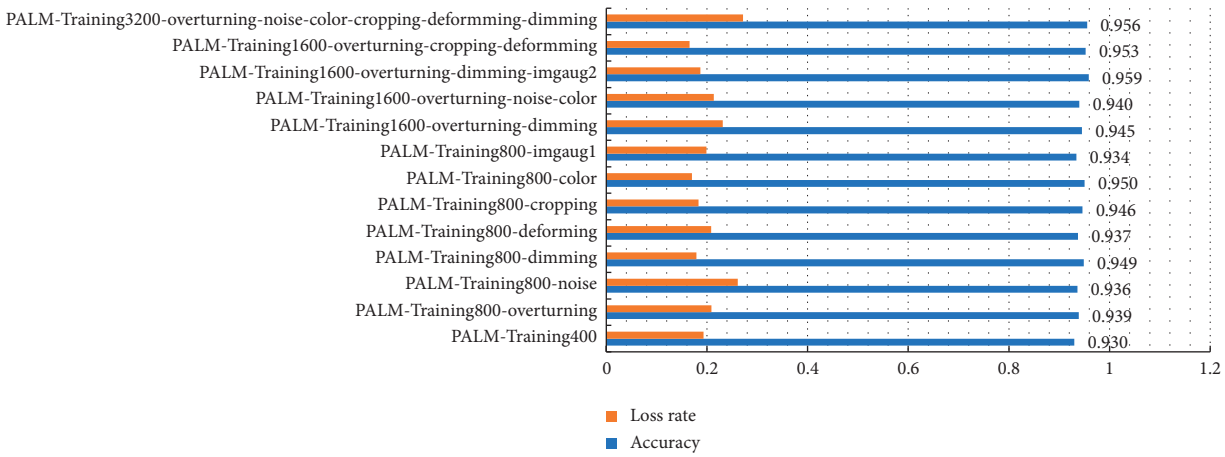


Figure 10: Distribution of accuracy and loss rate of VGG-16 under different MF.

Table 3: Training results of VGG-16, AlexNet, GoogLeNet, and ResNet-50 on the filtered datasets.

Primary learner	PALM-Training800-color		PALM-Training1600-overturning-cropping-deforming		PALM-Training3200-overturning-noise-color-cropping-deforming-dimming		PALM-Training1600-overturning-dimming-imgaug2	
	Accuracy	Loss rate	Accuracy	Loss rate	Accuracy	Loss rate	Accuracy	Loss rate
AlexNet	0.946083	0.162323	0.950833	0.160698	0.954667	0.197096	0.957583	0.157773
GoogLeNet	0.909	0.203461	0.9395	0.169759	0.962417	0.13617727	0.947917	0.170418
ResNet-50	0.9395	0.219571	0.953583	0.151836	0.955917	0.156737	0.95125	0.166358
VGG-16	0.9503	0.17019	0.95267	0.16523	0.9555	0.27185	0.95858	0.18674

Table 4: Results on the iChallenge-PM dataset.

	Accuracy (%)	Methods
Siying Dai [28]	81.82	Optimize network structure + DA
InstDis [48]	95.32	Optimize network structure + DA
Contrastive [49]	96.94	Optimize network structure + DA
Invariant [50]	97.30	Optimize network structure + DA
Xiaomeng Li [51]	98.65	Optimize network structure + DA
Ours	**97.25**	Optimize network structure + DA

the percentage (%). It was worth noting that the accuracy of each model could reach about 95% at the end of training without model fusion, which meant that DAMF played a highly favorable role.

We compared all DAMFs during the training process of VGG-16 by making them learn from all 13 datasets. Figure 10 exhibits the training effect of VGG-16 on the passing accuracy and loss rate of each dataset. It was believed that DAMF should not be as complicated as possible. The best result appeared in PALM-Training1600-overturning-dimming-imgaug1. When DAMF got complicated, the effect would decrease instead. Excessively complex processing of images might destroy valuable features in the image. This meant that DA was a process rather than formula and that we needed to locate the DAMF dataset that best fit each dataset in the dynamic process.

We compared the classification results under different strategies on a fixed dataset (iChallenge-PM). Fully considering the differences of research studies, part of the research focused on the optimization of the network structure, while others focused on the DA. The overall idea of the research is to start from the optimization of the network model and DA direction, with the ultimate goal of model accuracy. Table 4 displays the accuracy of different studies on the iChallenge-PM dataset in recent years. The results showed that although our accuracy rate is not the highest, it is also encouraging. In particular, our calculation cost was low. All training took 19 hours and 56 minutes, and no expensive calculation methods such as transfer learning were used.

6. Conclusion

The in-depth analysis of the discussed image DA techniques based on fusion has taken PM images as the research object and convincingly displayed their wide variety of applications. DAMF has the advantage of effectively improving the accuracy of model training, and the optimal enhanced set of

different datasets can be matched through this strategy. By analyzing the image DA method proposed in this paper, DAMF proves to be better and more effective than other methods. The experiment results have shown that DA still has an optimal complexity in the combination of DAMF. Otherwise, too much complexity may destroy the original features. DAMF can effectively find the best combination of DA which well retains the characteristics of the input images and provides better contrast through 11 contrast combinations. It has also been observed that DAMF can train the model to satisfactory results without using transfer learning or other methods. Arguably, DAMF can be used as an effective DA method during the training of CNNs in the field of fundus image processing.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the Xiamen Science and Technology Program (3502Z20183047) and China Education Research Network (NGII20170620).

References

[1] C. W. Wong, B. J. Fenner, and G. C. M. Cheung, "Epidemiology of myopic and vitreomaculopathies," in *A. Chang, W. F. Mieler, M. Ohji (eds)*, pp. 285–295, Macular Surgery, Springer, Singapore, 2020.

[2] K. Ohno-Matsui, "Pathologic myopia," *Asia-Pacific Journal of Ophthalmology*, vol. 5, no. 6, pp. 415–423, 2016.

[3] T. Y. Wong, A. Ferreira, R. Hughes, G. Carter, and P. Mitchell, "Epidemiology and disease burden of pathologic myopia and myopic choroidal neovascularization: an evidence-based systematic review," *American Journal of Ophthalmology*, vol. 157, no. 1, pp. 9–25, 2014.

[4] S.-K. Jung, J. H. Lee, H. Kakizaki, and D. Jee, "Prevalence of myopia and its association with body stature and educational level in 19-year-old male conscripts in Seoul, South Korea," *Investigative Opthalmology & Visual Science*, vol. 53, no. 9, p. 5579, 2012.

[5] B. A. Holden, T. R. Fricke, D. A. Wilson et al., "Global prevalence of myopia and high myopia and temporal trends from 2000 through 2050," *Ophthalmology*, vol. 123, no. 5, pp. 1036–1042, 2016.

[6] Y.-L. Wong, C. Sabanayagam, Y. Ding et al., "Prevalence, risk factors, and impact of myopic macular degeneration on visual impairment and functioning among adults in Singapore," *Investigative Opthalmology & Visual Science*, vol. 59, no. 11, p. 4603, 2018.

[7] C. W. Wong, N. Brennan, and M. Ang, "Introduction and overview on myopia: a clinical perspective," in *M. Ang, T. Wong (eds) Updates on Myopia*, pp. 1–26, Springer, Singapore, 2020.

[8] S. Vitale, M. F. Cotch, and R. D. Sperduto, "Prevalence of visual impairment in the United States," *JAMA*, vol. 295, no. 18, p. 2158, 2006.

[9] I. G. Morgan, K. Ohno-Matsui, and S.-M. Saw, "Myopia," *The Lancet*, vol. 379, no. 9827, pp. 1739–1748, 2012.

[10] W.-M. Hsu, C.-Y. Cheng, J.-H. Liu, S.-Y. Tsai, and P. Chou, "Prevalence and causes of visual impairment in an elderly Chinese population in Taiwan the authors have no proprietary interest in any aspect of the study," *Ophthalmology*, vol. 111, no. 1, pp. 62–69, 2004.

[11] A. Iwase, M. Araie, A. Tomidokoro, T. Yamamoto, H. Shimizu, and Y. Kitazawa, "Prevalence and causes of low vision and blindness in a Japanese adult population," *Ophthalmology*, vol. 113, no. 8, pp. 1354–1362, 2006.

[12] L. Xu et al., "Causes of blindness and visual impairment in urban and rural areas in beijing: the beijing eye study," *Ophthalmology*, vol. 113, no. 7, pp. 1134.e1–1134.e11, 2006.

[13] M. Yamada, Y. Hiratsuka, C. B. Roberts et al., "Prevalence of visual impairment in the adult Japanese population by cause and severity and future projections," *Ophthalmic Epidemiology*, vol. 17, no. 1, pp. 50–57, 2010.

[14] Q. S. You, L. Xu, H. Yang, Y. X. Wang, and J. B. Jonas, "Five-year incidence of visual impairment and blindness in adult Chinese," *Ophthalmology*, vol. 118, no. 6, pp. 1069–1075, 2011.

[15] J. He, S. L. Baxter, J. Xu, J. Xu, X. Zhou, and K. Zhang, "The practical implementation of artificial intelligence technologies in medicine," *Nature Medicine*, vol. 25, no. 1, pp. 30–36, 2019.

[16] W. Lu, Y. Tong, Y. Yu, Y. Xing, C. Chen, and Y. Shen, "Applications of artificial intelligence in ophthalmology: general overview," *Journal of Ophthalmology*, vol. 2018, Article ID 5278196, 15 pages, 2018.

[17] M. I. Razzak, S. Naz, and A. Zaib, "Deep learning for medical image processing: overview, challenges and the future," in *Lecture Notes in Computational Vision and Biomechanics*, pp. 323–350, Springer International Publishing, Berlin, Germany, 2018.

[18] K.-H. Yu, A. L. Beam, and I. S. Kohane, "Artificial intelligence in healthcare," *Nature Biomedical Engineering*, vol. 2, no. 10, pp. 719–731, 2018.

[19] A. V. Varadarajan, R. Poplin, K. Blumer et al., "Deep learning for predicting refractive error from retinal fundus images," *Investigative Opthalmology & Visual Science*, vol. 59, no. 7, p. 2861, 2018.

[20] D. S. W. Ting, L. R. Pasquale, L. Peng et al., "Artificial intelligence and deep learning in ophthalmology," *British Journal of Ophthalmology*, vol. 103, no. 2, pp. 167–175, 2019.

[21] S. Kuwayama, "Automated detection of macular diseases by optical coherence tomography and artificial intelligence machine learning of optical coherence tomography images," *Journal of Ophthalmology*, vol. 2019, Article ID 6319581, 7 pages, 2019.

[22] D. T. Hogarty, D. A. Mackey, and A. W. Hewitt, "Current state and future prospects of artificial intelligence in ophthalmology: a review," *Clinical & Experimental Ophthalmology*, vol. 47, no. 1, pp. 128–139, 2019.

[23] K. Suzuki, "Overview of deep learning in medical imaging," *Radiological Physics and Technology*, vol. 10, no. 3, pp. 257–273, 2017.

[24] Z. Li, S. Keel, C. Liu et al., "An automated grading system for detection of vision-threatening referable diabetic retinopathy on the basis of color fundus photographs," *Diabetes Care*, vol. 41, no. 12, pp. 2509–2516, 2018.

[25] J. De Fauw, J. R. Ledsam, B. Romera-Paredes et al., "Clinically applicable deep learning for diagnosis and referral in retinal disease," *Nature Medicine*, vol. 24, no. 9, pp. 1342–1350, 2018.

[26] K. Shankar, Y. Zhang, Y. Liu, L. Wu, and C.-H. Chen, "Hyperparameter tuning deep learning for diabetic retinopathy fundus image classification," *IEEE Access*, vol. 8, pp. 118164–118173, 2020.

[27] X. Sun, X. Cao, Y. Yang, L. Wang, and Y. Xu, "Robust retinal vessel segmentation from a data augmentation perspective," 2020, https://arxiv.org/abs/2007.15883.

[28] S. a. C. Dai, Leiting, T. Lei, C. Zhou, and Y. Wen, "Automatic detection of pathological myopia and high myopia on fundus images," in *Proceedings of the IEEE International Conference on Multimedia and Expo (ICME)*, pp. 1–6, London, UK, 2020.

[29] J. Cheng, Z. Li, Z. Gu, H. Fu, D. W. K. Wong, and J. Liu, "Structure-preserving guided retinal image filtering and its application for optic disk analysis," *IEEE Transactions on Medical Imaging*, vol. 37, no. 11, pp. 2536–2546, 2018.

[30] L. Wang, "Comparative analysis of image classification methods for automatic diagnosis of ophthalmic images," *Scientific Reports*, vol. 7, no. 1, p. 41545, 2017.

[31] Z. Zhang, J. Cheng, J. Liu, Y. C. M. Sheri, C. C. Kong, and S. S. Mei, "Pathological myopia detection from selective fundus image features," in *Proceedings of the 7th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, pp. 1742–1745, Singapore, 2012.

[32] J. Jiang, "Automatic diagnosis of imbalanced ophthalmic images using a cost-sensitive deep convolutional neural network," *BioMedical Engineering OnLine*, vol. 16, no. 1, 2017.

[33] H. Ačkar, A. A. Allmisreb, and M. A. Saleh, "A review on image enhancement techniques," *International Journal of Engineering and Applied Computer Science (IJEACS)*, vol. 2, no. 7, 2017.

[34] R. Bernardes, P. Serranho, and C. Lobo, "Digital ocular fundus imaging: a review," *Ophthalmologica*, vol. 226, no. 4, pp. 161–181, 2011.

[35] N. V. Chawla, A. Lazarevic, L. O. Hall, and K. W. Bowyer, "SMOTEBoost: improving prediction of the minority class in boosting," *Knowledge Discovery in Databases: PKDD 2003*, Springer, Berlin, Germany, pp. 107–119, 2003.

[36] A. S. Parihar, K. Singh, H. Rohilla, G. Asnani, and H. Kour, "A comprehensive analysis of fusion-based image enhancement techniques," in *Proceedings of the 4th International Conference on Intelligent Computing and Control Systems (ICICCS)*, IEEE, Melur, India, 2020.

[37] A. Maheshwari and Y. Gupta, "Empirical aspect of big data to enhance medical images using HIPI," in *Proceedings of the 2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS)*, IEEE, Melur, India, 2018.

[38] F. c. c. o. Chollet, "Xception: deep learning with depthwise separable convolutions," 2017, https://arxiv.org/abs/1610.02357.

[39] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.

[40] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," 2015, https://arxiv.org/abs/1512.03385.

[41] X. Fu, D. Zeng, Y. Huang, Y. Liao, X. Ding, and J. Paisley, "A fusion-based enhancing method for weakly illuminated images," *Signal Processing*, vol. 129, pp. 82–96, 2016.

[42] R. Takahashi, T. Matsubara, and K. Uehara, "Data augmentation using random image cropping and patching for deep CNNs," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 30, no. 9, pp. 2917–2931, 2020.

[43] H. Zhang, M. Cisse, Y. N. Dauphin, and D. Lopez-Paz, "Mixup: beyond empirical risk minimization," 2018, https://arxiv.org/abs/1710.09412.

[44] B. Zoph and Q. V. Le, "Neural architecture search with reinforcement learning," 2017, https://arxiv.org/abs/1611.01578.

[45] E. D. Cubuk, B. Zoph, D. Mane, V. Vasudevan, and Q. V. Le, "AutoAugment: learning augmentation policies from data," 2019, https://arxiv.org/abs/1805.09501.

[46] H. Fu, J. Cheng, Y. Xu, D. Wong, J. Liu, and X. Cao, "Joint optic disc and cup segmentation based on multi-label deep network and polar transformation," *IEEE Transactions on Medical Imaging*, vol. 37, no. 7, pp. 1597–1605, 2018.

[47] B. Normalization, "Accelerating deep network training by reducing internal covariate shift," 2015, https://arxiv.org/abs/1502.03167.

[48] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, "A simple framework for contrastive learning of visual representations," in *Proceedings of the 37th International Conference on Machine Learning*, Vienna, Austria, 2020, http://proceedings.mlr.press/v119/chen20j.html.

[49] Q. V. Le, "Building high-level features using large scale unsupervised learning," in *Proceedings of the 2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, IEEE, Vancouver, Canada, 2013.

[50] H. Lee, R. Grosse, R. Ranganath, and A. Y. Ng, "Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations," in *Proceedings of the 26th Annual International Conference on Machine Learning*, Montreal, Canada, 2009.

[51] Z. Wu, Y. Xiong, S. X. Yu, and D. Lin, "Unsupervised feature learning via non-parametric instance discrimination," in *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, IEEE, Salt Lake City, UT, USA, 2018.

Hindawi
Journal of Healthcare Engineering
Volume 2021, Article ID 9917545, 13 pages
https://doi.org/10.1155/2021/9917545

Research Article

Progressive Transmission of Medical Images via a Bank of Generative Adversarial Networks

Ching-Chun Chang,[1] **Xu Wang**,[2] **Ji-Hwei Horng**,[3] **and Isao Echizen**[1]

[1]*National Institute of Informatics, Tokyo, Japan*
[2]*Department of Information Engineering and Computer Science, Feng Chia University, Taichung 40724, Taiwan*
[3]*Department of Electronic Engineering, National Quemoy University, Kinmen 89250, Taiwan*

Correspondence should be addressed to Xu Wang; xu.wang.phd@gmail.com

Received 13 March 2021; Revised 29 March 2021; Accepted 9 April 2021; Published 28 April 2021

Academic Editor: Chi-Hua Chen

The healthcare sector is currently undergoing a major transformation due to the recent advances in deep learning and artificial intelligence. Despite a significant breakthrough in medical imaging and diagnosis, there are still many open issues and undeveloped applications in the healthcare domain. In particular, transmission of a large volume of medical images proves to be a challenging and time-consuming problem, and yet no prior studies have investigated the use of deep neural networks towards this task. The purpose of this paper is to introduce and develop a deep-learning approach for the efficient transmission of medical images, with a particular interest in the progressive coding of bit-planes. We establish a connection between bit-plane synthesis and image-to-image translation and propose a two-step pipeline for progressive image transmission. First, a bank of generative adversarial networks is trained for predicting bit-planes in a top-down manner, and then prediction residuals are encoded with a tailored adaptive lossless compression algorithm. Experimental results validate the effectiveness of the network bank for generating an accurate low-order bit-plane from high-order bit-planes and demonstrate an advantage of the tailored compression algorithm over conventional arithmetic coding for this special type of prediction residuals in terms of compression ratio.

1. Introduction

With the development of digital imaging equipment, more and more medical images have been produced and a medical image dataset usually contains a large number of images. Transmission of abundant medical images over a low narrow bandwidth public network will no doubt cause critical network pressure and further delay the timing of diagnosis [1].

To increase the speed of medical image transmission, some schemes have been proposed to build up more effective special networks. For instance, in [2], a unique pair of interfaces was proposed to provide a fast thoroughfare for medical data transmission between two DICOM applications and could speed up 1.22 and 13 times, respectively, in the local area networks (LAN) and wide area networks (WAN). Later, they further proposed an improved network [3] to combine the local area network (LAN) and wide area network (WAN) and used parallel TCP connections to optimize the DICOM protocol. This network can speed up about 2.2 to 3.5 times when transmitting magnetic resonance images.

On the other aspect, some new image transmission techniques have been proposed. An intuitive way is to reduce the resolution of medical images before transmission in a poor bandwidth network. However, doctors suffer from misdiagnosis with the lower resolution medical images. Therefore, the progressive image transmission (PIT) [4] technique emerges as the times require. The conventional raster scan ordered pixel by pixel image transmission

technique is inefficient, and the key region of the image which indicates the disease may be received at the end of the transmission makes doctors out of patience and affects the treatment. The progressive image transmission technique usually transmits a lower resolution image at first and gradually transmits the details of the image to further increase the resolution until the original image is totally transmitted. Through the transmission, receivers can decide whether they want to receive the details of the image or not in a short time. As a consequence, doctors can control the final image resolution and the required storage space. Over the past decades, some mature PIT techniques have been proposed. The bit-plane progressive image transmission method (BPM) [5, 6] is one of the low computational complexity and intuitive technique to transmit the image bit-plane by bit-plane which is detailed in Section 2. Another sort of progressive image transmission techniques is that, in [7–12], the image is divided into nonoverlapping blocks, and a quantized pixel value is used to represent the whole block, then, progressively dividing the image into finer blocks to improve the image resolution. Some schemes [13–17] transform the image into a wavelet domain and transmit the coefficients instead of pixels. Lossy compression [18] and lossless compression [19] methods are also adopted in some schemes to reduce the file size in each round of transmission.

The deep learning technique has matured in recent years and drawn massive research which is widely used in industry [20]. Some advanced deep learning architectures have also been introduced into medical image processing. Medical image classification is the most important area where deep learning makes a great contribution. The CNN network has been widely used in medical image classification and achieved a higher accuracy [21–23]. The deep learning technique is also suitable for medical image detection. The methods for organ, region, and landmark localization [24–26] and object and lesion detection [27–29] all adopt the deep learning architectures. Then, the medical image segmentation is naturally developed in the detected images [30–32]. The other tasks on medical images processing such as content-based image retrieval [33]; image generation and enhancement [34]; and description of medical images with shape, margin, and density [35] have also been proposed. Recently, the deep learning architecture with adversarial learning: generative adversarial network (GAN) has been proposed by Goodfellow et al. [36] that has the ability to generate realistic-looking images. Based on the GAN network, a powerful architecture pix2pix [37] was set up as an image bit-plane predictor to translate the image styles and has been adopted in different scenarios such as steganography [38] and even in medical images [39, 40]. However, so far, these deep learning techniques have not been applied to the medical image transmission. By leveraging the prediction ability of the GAN network, the medical images can be compressed before transmission and thus the efficiency of transmission can be improved.

In this paper, we adopt the framework of the bit-plane PIT method to transmit medical images. To reduce the amount of data before transmission, a GAN-based predictor bank is designed to predict the bit-planes. By progressively predicting the bit-plane to be sent based on previously transmitted bit-planes, we need only to send the residual provided that the receiver can execute the same prediction. The proposed GAN-based predictor bank produces a prediction of high accuracy so that the residual can be efficiently compressed before transmission. Thus, the efficiency of bit-plane PIT is greatly improved, and the transmission cost can be reduced. The main contributions of our scheme are summarized as follows:

(1) Introduce a deep neural network of pix2pix GAN network into the research of progressive medical image prediction and transmission

(2) Design a GAN-based predictor bank to help bit-plane compression of medical images

(3) Propose an adaptive residual bit-plane compression technique of high compression ratio

Experimental results validate the effectiveness and efficiency of the proposed prediction network and compression technique. The rest of this paper is organized as follows. Section 2 briefly introduces the bit-plane method of the progressive image transmission. Section 3 details the proposed adversarial learning for progressive bit-plane prediction scheme and the adaptive residual bit-plane compression technique. Section 4 provides the experimental results to demonstrate the accuracy of our prediction scheme and the efficiency of the proposed compression technique. In addition, the performance of the proposed residual bit-plane compression is compared with the conventional arithmetic coding. The conclusions are offered in Section 5.

2. Bit-Plane Method of PIT

In this section, the bit-plane method of progressive image transmission (BPM) technique which has been widely used to transmit a large volume of high-resolution medical images, is briefly introduced. The progressive image transmission (PIT) technique is an effective way to transmit images over narrow bandwidth channels. Among the PIT techniques, BPM is a simple and intuitive way of transmission, which divides an image into bit-planes and sends it plane by plane. For the grayscale images, each pixel is valued from 0 to 255, that is, 256 gray levels, and can be represented by 8 bits. The original image can be decomposed into eight bit-planes from the most significant bit-plane b_8 to the least significant bit-plane b_1, so that b_8 records the main content and b_1 records the subtle details, respectively, of the original image. Therefore, in the bit-plane PIT method, bit-planes from b_8 to b_1 are sequentially transmitted and the receiver can progressively recover the image until all bit-planes are sent or the current image quality is satisfied. The detailed BPM method is described below.

In the most significant bit-plane b_8, the bit "0" and "1" indicate the corresponding pixels are ranging from "0 to 127" and "128 to 255," respectively. Therefore, after receiving b_8, the mean values of these two intervals, that is, 64 and 192, are used to coarsely represent them. When receiving the next bit-plane b_7, the receiver updates the image by using finer

quantization levels of (32, 96, 160, 224) to represent corresponding codes "00," "01," "10," and "11." This process proceeds until all of the eight bit-planes are sent. An example of BPM transmission is illustrated in Figure 1, where the progressive combination of bit-planes and their corresponding recovered images are presented. As the number of received bit-planes increases, the available quantization levels also increase and a finer recovered image can be obtained.

3. Progressive Bit-Plane Prediction and Compression

Based on BPM progressive image transmission, we propose a bit-plane prediction method and a compression technique to improve the transmission efficiency of medical images. In this section, the proposed image transmission scheme and a pix2pix model-based bit-plane prediction method are introduced first. Then, a new technique for compression of the prediction error bit-plane is presented.

3.1. Adversarial Learning for Progressive Bit-Plane Prediction. For the purpose of progressively predicting image bit-plane by bit-plane, the generative adversarial network is introduced in this paper to predict and synthesize bit-planes. GAN architecture is a classical neural network to generate synthetic images of a specific style through training. The pix2pix model is a famous model of a GAN-based network which can be exploited to make an image-to-image translation. In the pix2pix model, the generator G is a U-Net, which is designed for biomedical image segmentation, instead of the conventional encoder-decoder net. The U-Net consists of a contracting path to downsample the input image, and an expansive path to generate an output image of full resolution. The contracting path consists of four steps connected by 2x2 max pooling layer with stride 2; each step contains three repeated 3×3 convolutions with ReLU. The expansive path also consists of four steps connected by 2×2 upconvolution and halves of the feature maps are cropped from the contracting path. For the discriminator, the PatchGAN is adopted to discriminate that the overlapping patches in the generated image are real or fake. Different from the conventional GAN discriminator which only output one result, that is, real or fake, PatchGAN outputs a matrix and each element reflects the pixel differences of the current patch between the synthetic image and the original image in the current patch region. Additionally, to obtain a generated image with higher similarity to the input original image, the Manhattan distance (ℓ_1 norm) which is more suitable for high dimensional images is chosen to evaluate the difference between input and output images. Therefore, in the pix2pix model, assume x and y are the input and output images; the difference between input and output is evaluated by

$$\mathscr{L}_{L1}(G) = \mathbb{E}\left[\|y - G(x)\|_1\right]. \tag{1}$$

Then, the final objective is obtained by mixing the ℓ_1 loss with the traditional GAN loss by

$$\begin{aligned} G^* &= \arg\min_G \max_D \mathscr{L}_{\mathrm{GAN}}(G, D) + \lambda \mathscr{L}_{L1}(G), \\ \mathscr{L}_{\mathrm{GAN}}(G, D) &= \mathbb{E}[\log D(y)] + \mathbb{E}[\log(1 - D(G(x)))], \end{aligned} \tag{2}$$

where λ is a hyperparameter to balance two losses to achieve better performances.

Through the above architecture, the pix2pix model can translate the input image into different styles such as color images, day-to-night, and image inpainting. Inspired by pix2pix image-to-image translation model, in this paper, we train the pix2pix model to predict bit-planes of medical images. The proposed scheme is based on the BPM of progressive image transmission. A medical image is divided into eight bit-planes and transmitted plane by plane. On the receiver side, bit-planes are also received plane by plane and the medical image can be progressively reconstructed. During transmission, we leverage the pix2pix model to generate a predicted version of the bit-plane to be sent based on the previous bit-planes. Benefiting from the high accuracy of prediction, the prediction error of bit-plane (residual bit-plane) can be compressed efficiently. Thus, a bit-plane can be transmitted by sending its residual instead. The predictors of different layers of bit-plane are trained separately, organized into a predictor bank, and shared with the receiver. When receiving the residual, the receiver can do the same prediction to obtain the predicted bit-plane, and then, the original bit-plane can be recovered according to the received residual bit-plane and the predicted bit-plane. The framework of our proposed progressive medical image transmission scheme is shown in Figure 2.

In detail, assuming the bit-planes from the MSB to LSB are b_8 to b_1, we feed all b_8 bit-planes of training images into the model as the input images to generate output images and set their corresponding b_7 bit-planes as the ground truth images. The standard format of input and output for the pix2pix model is an image of 8-bit depth. To fit this format, zeros are appended behind the provided bit-planes. When training the model for predicting b_6, b_8 and b_7 are concatenated as the input. An illustration for this case is shown in Figure 3. The detailed prediction models for the remaining bit-planes are trained in a similar manner. Then, a series of pretrained predictors are organized into a predictor bank as shown in Figure 4, which contains seven predictors for predicting bit-planes of different layers. For instance, the b_6 predictor can predict bit-plane b_6 by using the bit-planes b_8 and b_7. Therefore, each bit-plane from b_7 to b_1 can be progressively predicted by its previous bit-planes. On the application phase, the residual bit-plane can be calculated by comparing the bit-plane to be sent and the synthetic one by

$$\mathscr{R}_{i,j} = \begin{cases} 0, & \text{if } \mathscr{B}_{i,j} = \mathscr{S}_{i,j}, \\ 1, & \text{if } \mathscr{B}_{i,j} \neq \mathscr{S}_{i,j}, \end{cases} \tag{3}$$

where $\mathscr{R}$, $\mathscr{B}$, and $\mathscr{S}$ represent the residual, original, and synthetic bit-planes, respectively. Assuming the image is sized $M \times M$, i, j denote the coordinates and range from 1 to M. In this paper, the seven residual bit-planes are denoted by symbols r_7 to r_1. On the receiver side, the predicted bit-plane

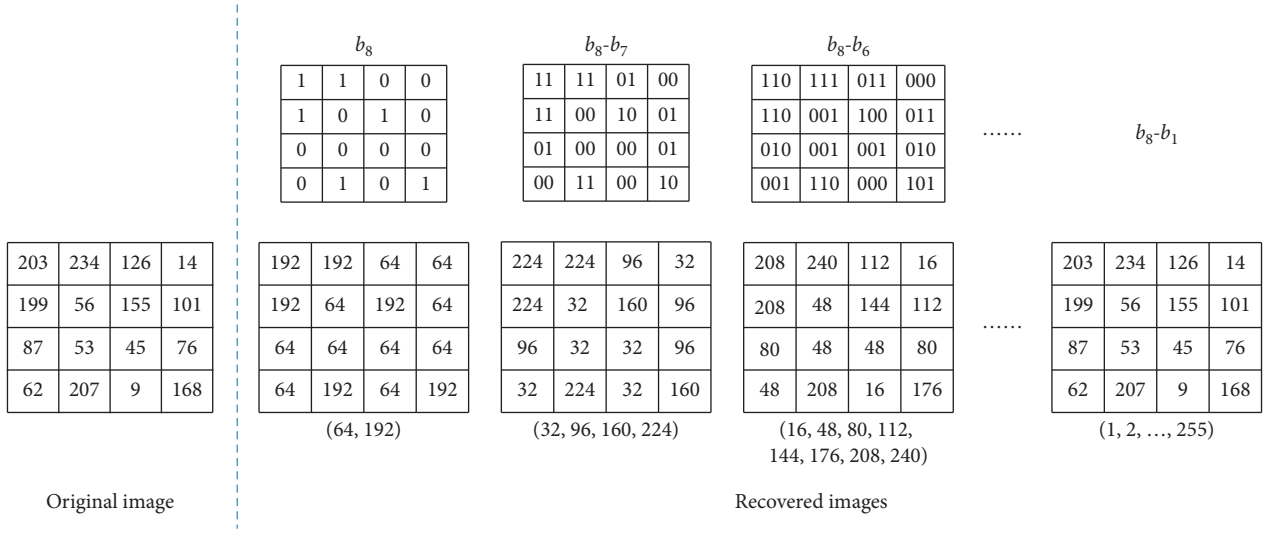


Figure 1: An illustration of the bit-plane PIT method.

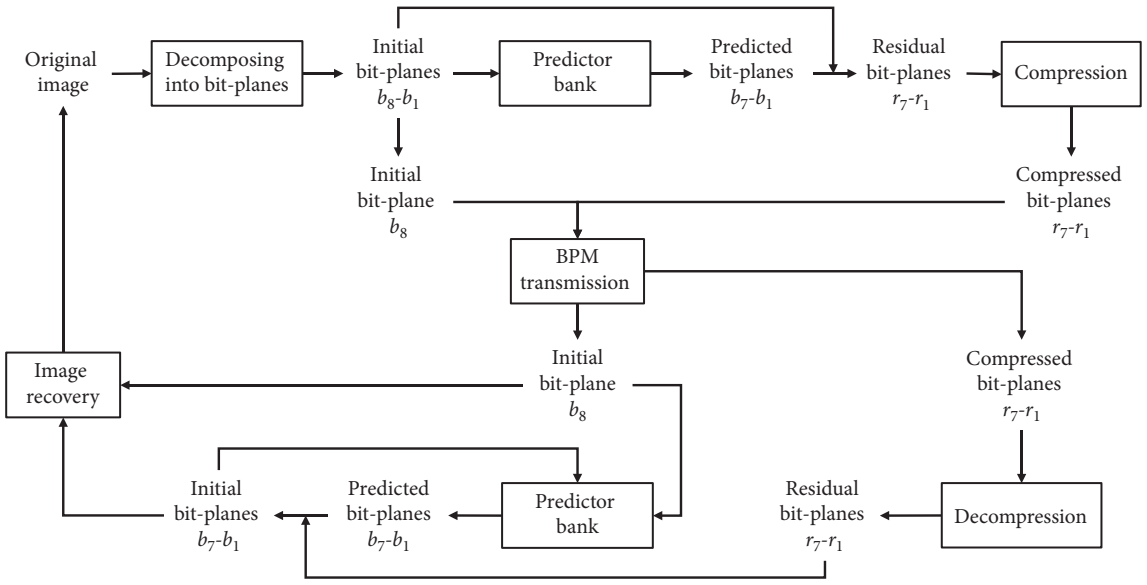


Figure 2: The framework of our proposed scheme.

can be obtained by the preshared predictor bank. After receiving the residual bit-plane, the real bit-plane can be recovered by

$$\mathscr{B}_{i,j} = \begin{cases} \mathscr{S}_{i,j}, & \text{if } \mathscr{R}_{i,j} = 0, \\ \overline{\mathscr{S}_{i,j}}, & \text{if } \mathscr{R}_{i,j} = 1, \end{cases} \tag{4}$$

where $\overline{\mathscr{S}_{i,j}}$ denotes the binary bit-flipped value of $\mathscr{S}_{i,j}$, that is, change "1" to "0" or change "0" to "1."

3.2. Adaptive Residual Bit-Plane Compression. The size of the prediction residual and its original bit-plane is the same. The major difference is that the values in the residual bit-plane are mostly zero. So, it can be compressed more efficiently. In addition, the error bits of our GAN-based bit-plane prediction are highly concentrated in the vicinity of image edges or complex details. According to the features of the residual bit-planes, we further propose an adaptive compression technique to reduce the data size. The detailed procedures are as follows:

Step 1: decompose a residual bit-plane $\mathscr{R}$ sized $M \times M$ into M^2/m^2 nonoverlapping blocks $\mathscr{R}^n$ sized $m \times m$, where n ranges from 1 to M^2/m^2.

Step 2: initialize a location map L with the length of M^2/m^2.

Step 3: scan the residual blocks in the raster scan order. If values in the block $\mathscr{R}^n$ are all "0" or all "1," go to Step 4; otherwise, go to Step 5.

Step 4: set the n-th bit L^n of the location map L to "0," and record the block value by one bit, that is, $z^n = \{b\}$,

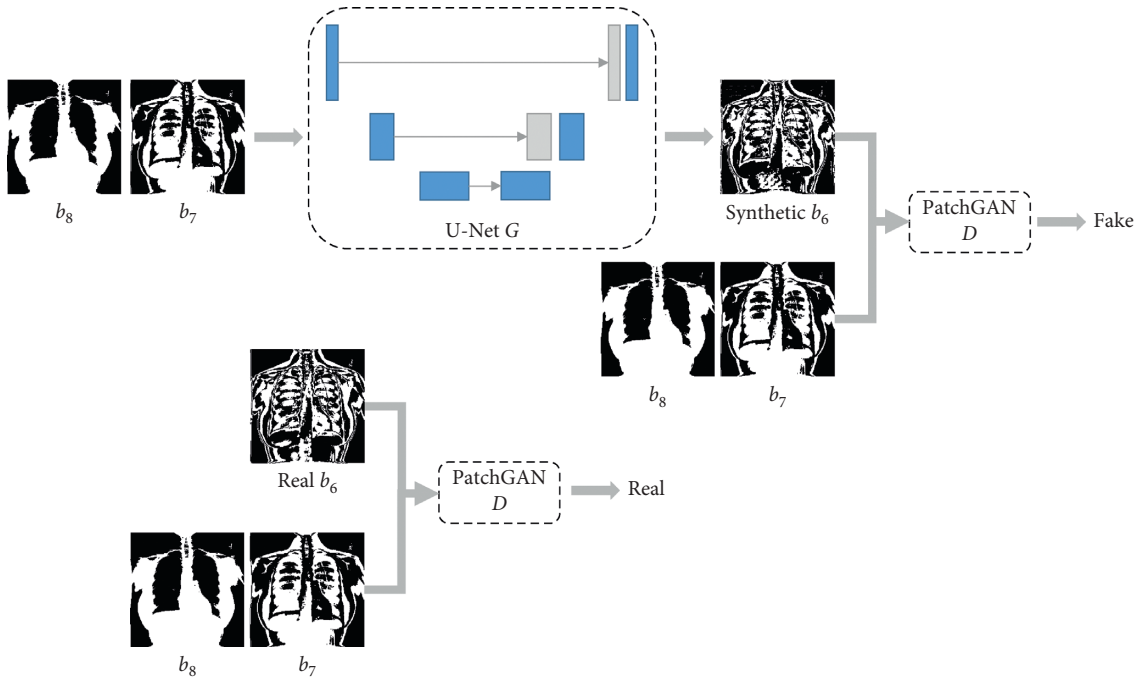


Figure 3: The training workflow of the network architecture.

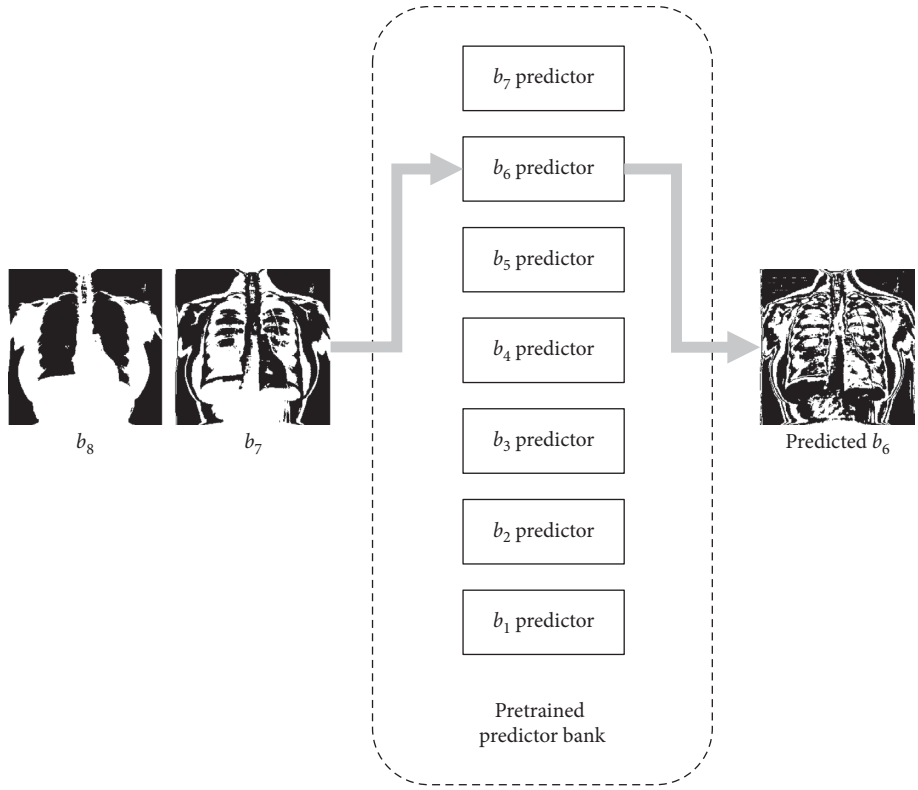


Figure 4: The predictor bank for bit-plane prediction.

where z^n denotes the n-th segment of the recovery sequence and b is the common value of the block's elements.

Step 5: set L^n to "1," and record all m^2 bits of the current residual block into the recovery sequence r^n in raster scan order.

Step 6: return to Step 3 until all blocks have been scanned to obtain the location map L and the recovery sequence r.

Step 7: compress the location map L by arithmetic coding to acquire the final location map L'. Then, the final compressed residual bit-plane is obtained by concatenating L' and r.

When receiving the compressed residual bit-plane, the initial residual bit-plane can be obtained by the following decompression procedures:

Step 1: decompress the location map L' by arithmetic decoding to obtain the M^2/m^2 sized location map L.

Step 2: if L^n is "0," all bits in the current block $\mathscr{R}^n$ are the same and go to Step 3; otherwise, go to Step 4.

Step 3: extract one-bit z^n from the recovery sequence r and set all m^2 bits in $\mathscr{R}^n$ to r^n.

Step 4: extract m^2 bits z^n from r and rearrange these bits in the raster scan order to recover $\mathscr{R}^n$.

Step 5: return to Step 2 until all blocks have been scanned. Then, the initial residual bit-plane $\mathscr{R}$ can be obtained by tiling the recovered residual blocks in the raster scan order.

A simple example is given in Figure 5. In this example, the block size m is set to 2. Since the first, third, and fourth blocks are uniform valued blocks, the corresponding bits in the location map are set to 0. Meanwhile, just one sample bit is recorded into the recovery sequence for each of them. For the rest blocks, their corresponding bits in the location map are set to 1 and all bit values are recorded. When recovering these blocks, referring to the location map and the recovery sequence, all blocks can be totally recovered.

4. Experimental Results

In this section, we evaluate the performance of our proposed bit-plane prediction scheme with generative adversarial networks and compare our proposed residual bit-plane compression scheme with the conventional arithmetic coding compression technique. The dataset used for training and testing, the training details, and the evaluation metrics are first introduced, and then, the performance of the proposed scheme is evaluated.

4.1. Dataset and Training Details. The image samples for training and testing are the greyscale chest X-ray images collected from the National Institutes of Health under the U.S. Department of Health and Human Services. In our experiments, 10,000 images from the first compressed file are used for training, and the other 10 images from the fifth compressed file as shown in Figure 6 are used for performance evaluating and analyzing. Before training and testing, all the images are first downsampled into 256×256 sized images to fit our proposed network architecture.

When training the proposed model, each bit-plane prediction is trained over 100 epochs and the initial learning rate and the batch size are set to 2×10^{-4} and 32, respectively. The learning rate is enforced halfway during training and the hyperparameter λ is set to 10^3. The model parameters are updated and optimized by the Adam function.

4.2. Evaluation Metrics. To evaluate the performance of the predicted bit-planes, two important metrics, error rate (ER) and the compression ratio (CR), are introduced in this paper to conduct some experiments. For a predicted bit-plane, referring to its original bit-plane, if the two bits in the same place are the same, the corresponding bit in the residual bit-plane is set to "0" and set to "1" when the predicted bit is incorrect. The error rate (ER) is defined as the ratio of "1's" in the residual bit-plane and can be calculated by

$$\mathrm{ER} = \frac{\varepsilon}{M \times M}\,(\%), \tag{5}$$

where ε denotes the number of error bits. After obtaining the residual bit-plane, different compression techniques are adopted to compress it. The compression ratio (CR) denotes the ratio between the data sizes of the original and the compressed versions that can be calculated by

$$\mathrm{CR} = \frac{M \times M}{\delta}, \tag{6}$$

where δ represents the data size, in bits, of the compressed residual. Generally, the compression ratio is greatly dependent on the error rate, and a residual of a low error rate can be compressed more efficiently. Thus, a prediction of high accuracy can improve the transmission efficiency.

4.3. Performances. First, the test image "Image 05" is used to demonstrate the experimental results of the original bit-planes, prediction bit-planes, and corresponding residual bit-planes.

From Figure 7, when receiving the bit-plane 8, the rest bit-planes can be progressively synthesized to obtain the predicted bit-planes. In the residual bit-plane, the white pixels are the error predictions, and most of them are concentrated in the edges of the image. The error rate of the bit-plane 7 is the lowest because this bit-plane is relatively smooth and just a few edges are presented. The error rate gradually increases as the details of a bit-plane become more complex. Nevertheless, the error rate of bit-plane 1 is still lower than 35% and most bits can be accurately predicted. Therefore, the proposed scheme can progressively predict bit-planes with high accuracy. To further verify the prediction accuracy, Table 1 lists all error rates of different layers in ten test images.

The error rate monotonously increases from r_7 to r_1. Two exceptions occur in "Image 04" and "Image 06": the error rate of r_6 is lower than r_7. The lowest error rate and highest error rate are 5.9723% and 36.1099%, respectively, in r_7 of "Image 07" and r_1 of "Image 09." Generally, the average error rate monotonously increases from high bit-plane to low bit-plane. The average error rate for an entire image is 19.2673%.

The compression ratio is greatly dependent on the error rate and the distribution of error bits. A residual bit-plane

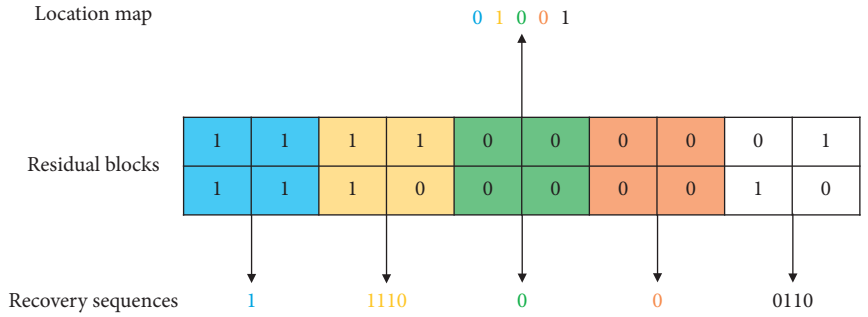


Figure 5: An example of the proposed adaptive residual bit-plane compression.

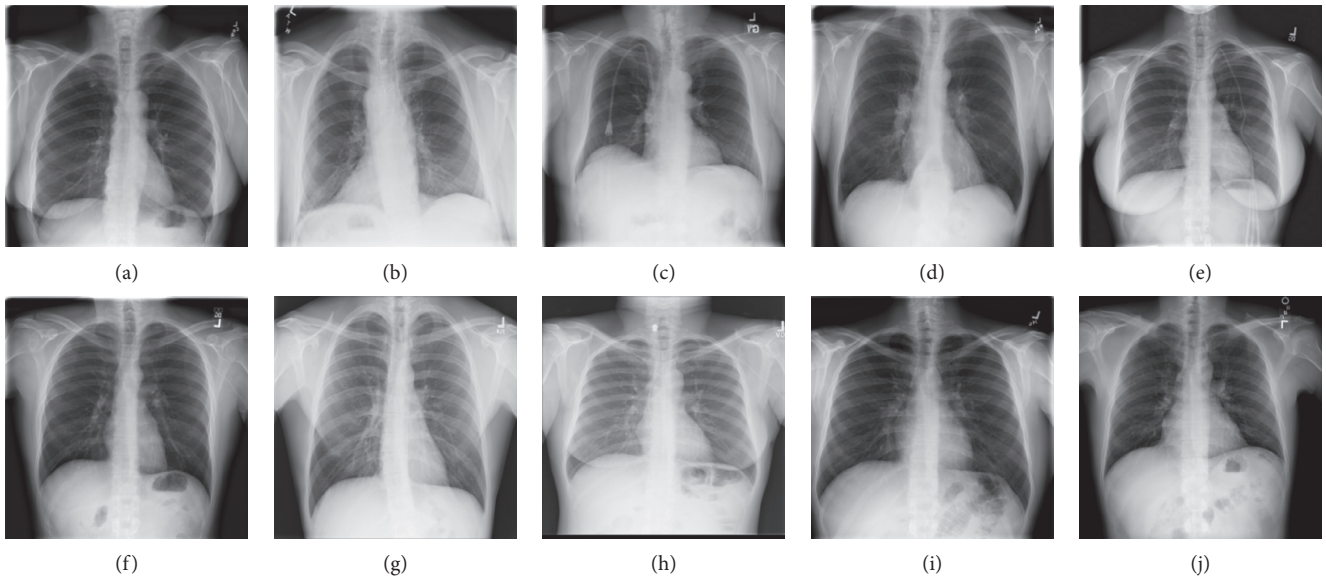


Figure 6: Ten test images: (a) Image 01, (b) Image 02, (c) Image 03, (d) Image 04, (e) Image 05, (f) Image 06, (g) Image 07, (h) Image 08, (i) Image 09, and (j) Image 10.

with a few concentrated error bits can be compressed efficiently. The compression ratio of the residual bit-planes is plotted in Figure 8. Since the number of error bits increases from r_7 to r_1, the compression ratio decreases as the bit layer decreases. Even in the exception cases of "Image 04" and "Image 06," where the error bits of r_7 are more than r_6, the compression ratio of r_7 is also greater due to the concentrated distribution of error bits.

To investigate the effect of block size setting, four different sizes are applied to compress the residual. The compression ratio for different block sizes and bit layers is listed in Table 2. The compression ratio of size 4 is significantly greater than the other sizes for all layers of bit-planes.

Additionally, our proposed residual bit-plane compression technique is also compared with the arithmetic coding technique. As shown in Figure 9, the average compression ratio of our proposed scheme is significantly better in the bit-planes of r_7 and r_6, while the performance of arithmetic coding is slightly better in the bit-planes of r_4 and r_3. That is because, in the rear residual bit-planes such as r_4 to r_1, the errors have no continuous distribution to make our compression technique not efficient. For compression of an entire image, the proposed scheme is significantly better than the arithmetic coding. To further verify the applicability in the medical image domain, the compression ratio for all test images is listed in Table 3. In Table 3, $\mathscr{R}$ and $\mathscr{B}$ denote that the sources of compression are residuals and original bit-planes, respectively. Additionally, since the original first bit-plane should be transmitted at the beginning, in two types of bit-plane $\mathscr{R}$ and $\mathscr{B}$, the original first bit-plane is transmitted and compressed by different compression methods. As shown in Table 3, the compression ratio of residual is greater than that of the original version. It implies the devised GAN predictor bank is helpful. In addition, the proposed bit-plane compression technique outperforms the arithmetic coding in the scheme of bit-planewise compression. As

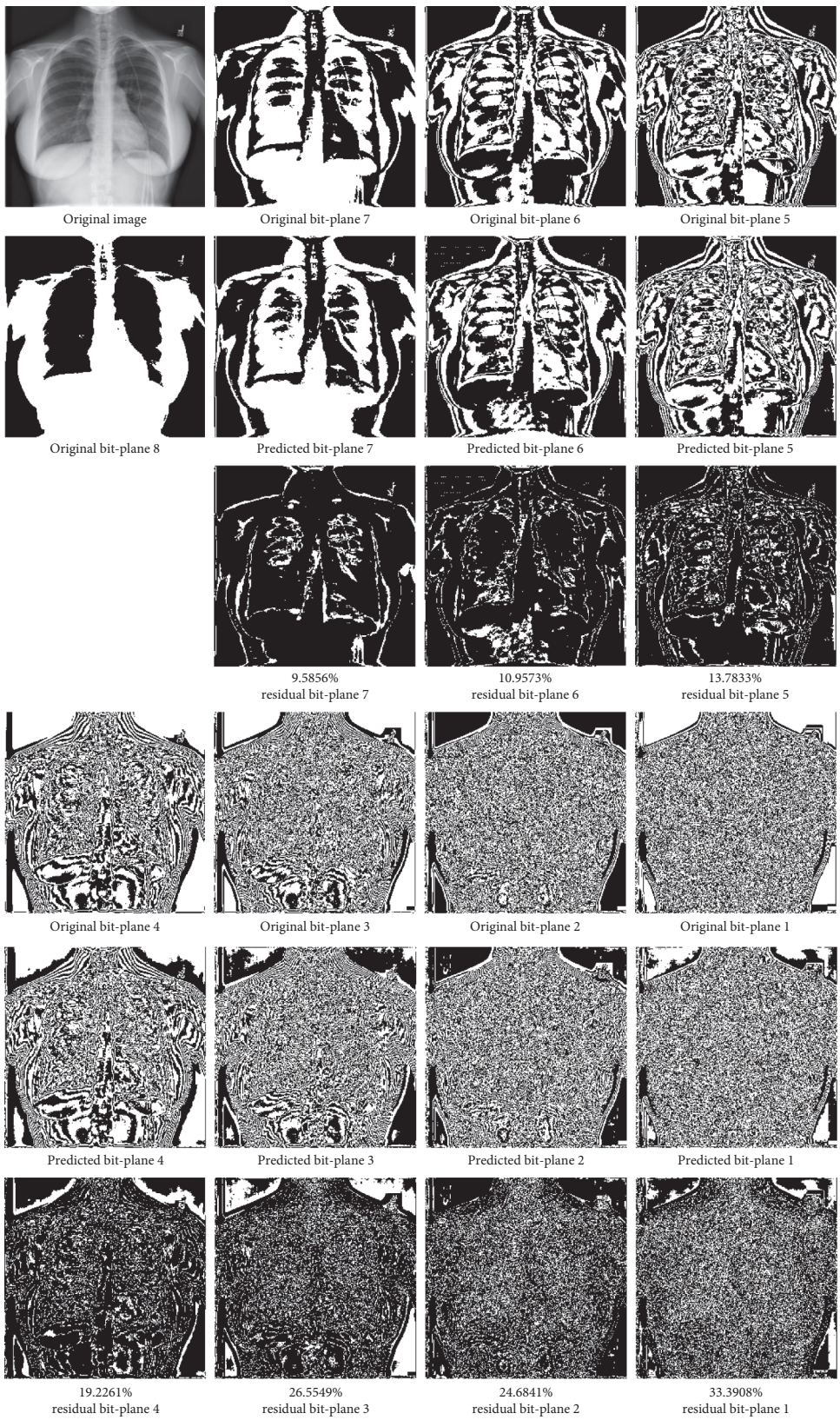
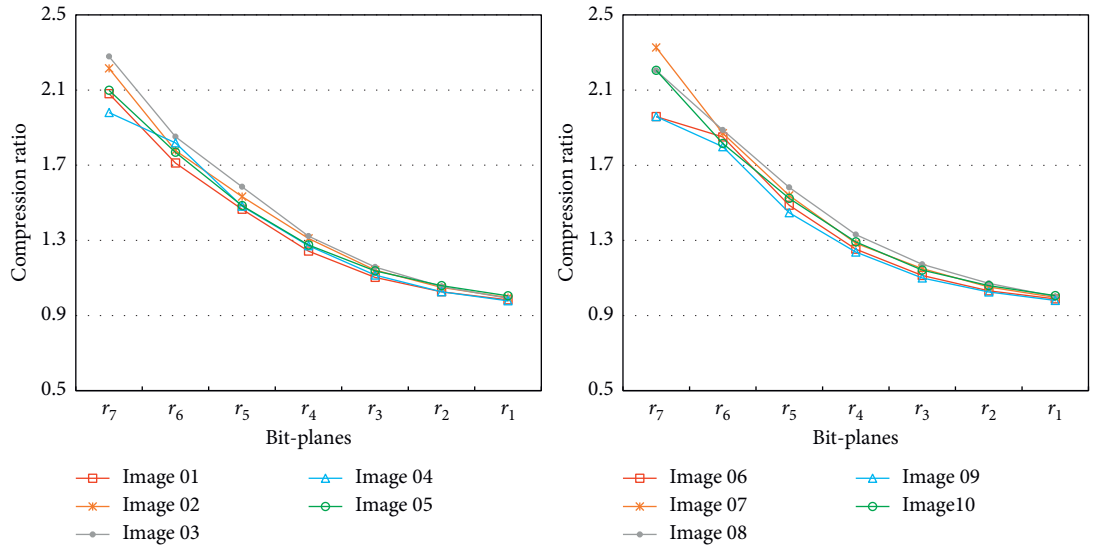


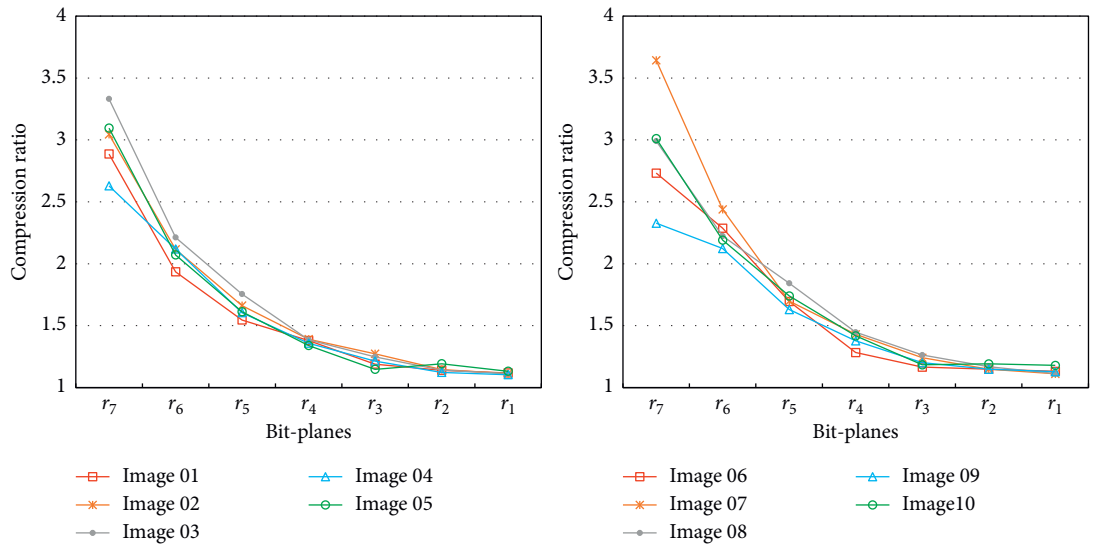
Figure 7: The original bit-planes (first row), the prediction bit-planes (second row), and the corresponding residual bit-planes with error rates (third row) of the test image "Image 05" (top-left).

Table 1: Error rates ER of residual bit-planes 7-1 in test images.

Images	r_7	r_6	r_5	r_4	r_3	r_2	r_1	Average
Image 01	10.4553	12.5137	14.1861	17.9123	22.0245	26.9897	34.8251	19.8438
Image 02	11.8851	14.6881	15.9241	18.6523	21.9116	25.9674	31.4346	20.0662
Image 03	8.4518	10.7819	12.5763	18.4158	21.4813	24.8047	31.7764	18.3269
Image 04	13.2202	11.4502	14.5264	18.1915	21.5698	26.7654	34.0698	19.9705
Image 05	9.5856	10.9573	13.7833	19.2261	26.5549	24.6841	33.3908	19.7403
Image 06	10.2814	9.4879	13.6139	22.3846	22.3145	26.4420	33.3206	19.6921
Image 07	5.9723	9.6817	14.3890	16.8106	20.3964	25.3586	31.3263	17.7050
Image 08	11.3373	11.6348	12.8296	16.2766	19.9966	24.0417	31.6544	18.2530
Image 09	16.1194	12.0071	14.6622	17.5919	21.8216	27.3041	36.1099	20.8023
Image 10	10.5713	12.0087	13.3682	16.1804	20.0516	24.5636	31.1661	18.2728
Average	10.7880	11.5211	13.9859	18.1642	21.8123	25.6921	32.9074	19.2673



(a)



(b)

Figure 8: Continued.

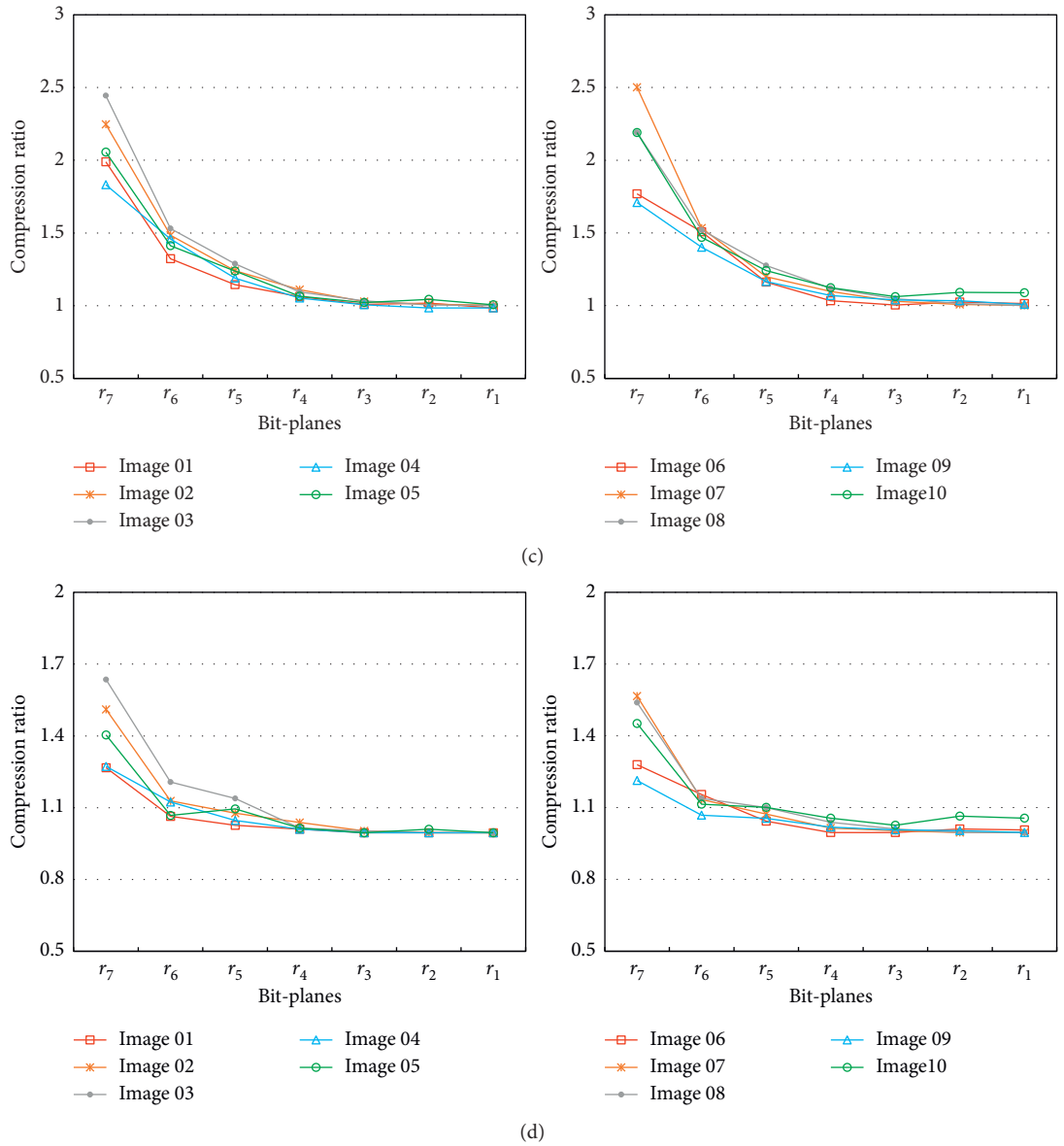


Figure 8: Compression ratio of each bit-plane in ten test images when the block sizes are (a) $m = 2$, (b) $m = 4$, (c) $m = 8$, and (d) $m = 16$.

Table 2: Average bit-plane compression ratio (CR) for different block sizes.

Block size	r_7	r_6	r_5	r_4	r_3	r_2	r_1
2	2.1295	1.8124	1.5112	1.2814	1.1327	1.0449	0.9922
4	**2.9095**	**2.2059**	**1.7053**	**1.3731**	**1.2614**	**1.2137**	**1.1426**
8	2.1585	1.5553	1.2788	1.1166	1.0435	1.0313	1.0147
16	1.4372	1.1383	1.0775	1.0266	1.0036	1.0072	1.0027

shown in the experimental results, the compression ratio is severely affected by the smoothness of residual bit-planes. In medical images, most of the textures are concentrated in the region of interest, which is particularly suitable for the proposed predictor bank to work efficiently. Additionally, the parameters of the pretrained predictor bank are crucial for the sender and receiver, so that the predictor bank needs to be protected to ensure applicability and security.

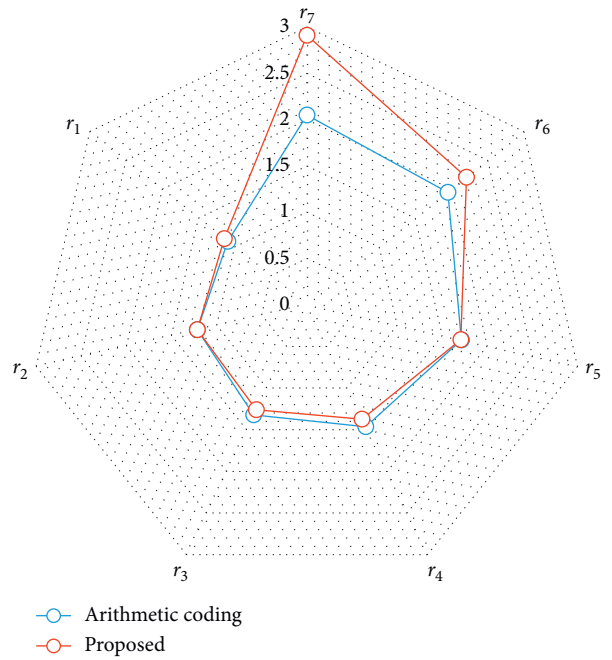


Figure 9: Average compression ratio comparison in each bit-plane.

Table 3: Comparison of compression ratio for ten test images.

	Image 01		Image 02		Image 03		Image 04		Image 05	
	$\mathscr{R}$	$\mathscr{B}$	$\mathscr{R}$	$\mathscr{B}$	$\mathscr{R}$	$\mathscr{B}$	$\mathscr{R}$	$\mathscr{B}$	$\mathscr{R}$	$\mathscr{B}$
Proposed	1.3923	1.2352	1.4347	1.3180	1.4566	1.3087	1.3970	1.2453	1.4272	1.2650
Arithmetic coding	1.3739	1.0062	1.3475	1.0098	1.4255	1.0020	1.3629	1.0022	1.3783	1.0101
	Image 06		Image 07		Image 08		Image 09		Image 10	
	$\mathscr{R}$	$\mathscr{B}$	$\mathscr{R}$	$\mathscr{B}$	$\mathscr{R}$	$\mathscr{B}$	$\mathscr{R}$	$\mathscr{B}$	$\mathscr{R}$	$\mathscr{B}$
Proposed	1.4003	1.2409	1.4609	1.2722	1.4615	1.2953	1.3917	1.2376	1.4601	1.2835
Arithmetic coding	1.3827	1.0052	1.4585	1.0051	1.4166	1.0006	1.3349	1.0055	1.4158	1.0116

5. Conclusions

In this paper, we introduce a GAN predictor bank and a bit-plane compression technique into the progressive transmission of medical images. Experimental results validate the proposed predictor bank can effectively predict bit-planes under the progressive scheme. Furthermore, the proposed adaptive bit-plane compression is more efficient than the arithmetic coding for the current application. This paper intends to introduce deep learning techniques into the classical progressive image transmission and we hope this paper can give some illuminations for future research. In our future works, we will further introduce and improve more convolutional neural networks into the compression and transmission of medical images.

Data Availability

The data used to support the findings of the study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was partially supported by the JSPS KAKENHI (Grant nos. JP16H06302 and JP18H04120) and JST CREST (Grant nos JPMJCR18A6 and JPMJCR20D3), Japan.

References

[1] H. U. Lemke, "Communication networks for medical image transmission," *Strahlentherapie und Onkologie: Organ der Deutschen Röntgengesellschaft*, vol. 169, no. 9, pp. 512–520, 1993.

[2] R. Maani, S. Camorlinga, N. Arnason, and R. Eskicioglu, "A practical fast method for medical imaging transmission based on the DICOM protocol," *Medical Imaging 2010: Advanced PACS-Based Imaging Informatics and Therapeutic Applications*, vol. 7628, Article ID 76280M, 2010.

[3] R. Maani, S. Camorlinga, and N. Arnason, "A parallel method to improve medical image transmission," *Journal of Digital Imaging*, vol. 25, no. 1, pp. 101–109, 2012.

[4] K.-H. Tzou, "Progressive image transmission: a review and comparison of techniques," *Optical Engineering*, vol. 26, no. 7, Article ID 267581, 1987.

[5] C.-C. Chang, F.-C. Shiue, and T.-S. Chen, "A new scheme of progressive image transmission based on bit-plane method,"

in *Proceedings of the Fifth Asia-Pacific Conference on and Fourth Optoelectronics and Communications Conference on Communications*, pp. 892–895, Singapore, October 1999.

[6] T.-S. Chen, H.-C. Wu, H.-F. Tsai, M. Hsieh, and S.-F. Chiou, "Progressive transmission of two-dimensional gel electrophoresis image based on context features and bit-plane method," in *Proceedings of the IEEE International Conference on Networking, Sensing and Control*, pp. 1241–1246, Taipei, Taiwan, March 2004.

[7] L. Wang and M. Goldberg, "Progressive image transmission using vector quantization on images in pyramid form," *IEEE Transactions on Communications*, vol. 37, no. 12, pp. 1339–1349, 1989.

[8] C.-C. Chang, H.-C. Hsia, and T.-S. Chen, "A progressive image transmission scheme based on block truncation coding," in *The Human Society and the Internet Internet-Related Socio-Economic Issues*pp. 383–397, Berlin, Heidelberg, 2001.

[9] K.-L. Chung and S.-Y. Tseng, "New progressive image transmission based on quadtree and shading approach with resolution control," *Pattern Recognition Letters*, vol. 22, no. 14, pp. 1545–1555, 2001.

[10] C.-C. Chang, G.-X. Xiao, and T.-S. Chen, "A simple prediction method for progressive image transmission," in *Distributed Multimedia Databases: Techniques and Applications, 2002*, The Idea Group Publishing, Hershey, PA, USA, http://www.igi-global.com/chapter/distributed-multimedia-databases/8626.

[11] H. Kim, R. Annavajjala, P. Cosman, and L. Milstein, "Source-channel rate optimization for progressive image transmission over block fading relay channels," *IEEE Transactions on Communications*, vol. 58, no. 6, pp. 1631–1642, 2010.

[12] B. C. Dhara and B. Chanda, "A fast progressive image transmission scheme using block truncation coding by pattern fitting," *Journal of Visual Communication and Image Representation*, vol. 23, no. 2, pp. 313–322, 2012.

[13] L. Wang and M. Goldberg, "Progressive image transmission by transform coefficient residual error quantization," *IEEE Transactions on Communications*, vol. 36, no. 1, pp. 75–87, 1988.

[14] W. J. Hwang and H. Derin, "Multiresolution multiresource progressive image transmission," *IEEE Transactions on Image Processing*, vol. 4, no. 8, pp. 1128–1140, 1995.

[15] Y. Lu, J. Zheng, Y. Jiang, M. Yang, B. Fu, and W. Hou, "Progressive image transmission for medical applications based on wavelet transform with a non-uniform scalar quantization scheme," in *Proceedings of the 2006 International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 4795–4798, New York, NY, USA, August 2006.

[16] C.-C. Chang and T.-C. Lu, "A wavelet-based progressive digital image transmission scheme," in *Proceedings of the First International Conference on Innovative Computing, Information and Control-Volume I (ICICIC'06)*, pp. 681–684, Beijing, China, September 2006.

[17] C.-P. Huang and C.-C. Li, "Secure and progressive image transmission through shadows generated by multiwavelet transform," *International Journal of Wavelets, Multiresolution and Information Processing*, vol. 6, no. 6, pp. 907–931, 2008.

[18] R. S. Dilmaghani, A. Ahmadian, M. Ghavami, and A. H. Aghvami, "Progressive medical image transmission and compression," *IEEE Signal Processing Letters*, vol. 11, no. 10, pp. 806–809, 2004.

[19] K.-L. Hung, C.-C. Chang, and I.-C. Lin, "Lossless compression-based progressive image transmission scheme," *The Imaging Science Journal*, vol. 52, no. 4, pp. 212–224, 2004.

[20] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[21] R. Neelapu, G. Lavanya Devi, and K. Srinivasa Rao, "Deep learning based conventional neural network architecture for medical image classification," *Traitement du signal*, vol. 35, no. 2, pp. 169–182, 2018.

[22] L. Faes, S. K. Wagner, D. J. Fu et al., "Automated deep learning design for medical image classification by health-care professionals with no coding experience: a feasibility study," *The Lancet Digital Health*, vol. 1, no. 5, pp. e232–e242, 2019.

[23] J. Wagner, Y. Xie, Q. Wu, and Y. Xia, "Medical image classification using synergic deep learning," *Medical Image Analysis*, vol. 54, pp. 10–19, 2019.

[24] H. Chen, D. Ni, J. Qin et al., "Standard plane localization in fetal ultrasound via domain transferred deep neural networks," *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 5, pp. 1627–1636, 2015.

[25] Y. Ni, M. Landis, D. T. Laidley, A. Kornecki, A. Lum, and S. Li, "Multi-modal vertebrae recognition using transformed deep convolution network," *Computerized Medical Imaging and Graphics*, vol. 51, pp. 11–19, 2016.

[26] A. Kumar, "Plane identification in fetal ultrasound images using saliency maps and convolutional neural networks," in *Proceedings of the 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, pp. 791–794, Prague, Czech Republic, April 2016.

[27] A. Teramoto, H. Fujita, O. Yamamuro, and T. Tamaki, "Automated detection of pulmonary nodules in PET/CT images: ensemble false-positive reduction using a convolutional neural network technique," *Medical Physics*, vol. 43, no. 6, pp. 2821–2827, 2016.

[28] M. J. J. P. Van Grinsven, B. van Ginneken, C. B. Hoyng, T. Theelen, and C. I. Sánchez, "Fast convolutional neural network training using selective data sampling: application to hemorrhage detection in color fundus images," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1273–1284, 2016.

[29] J. M. Wolterink, T. Leiner, B. D. de Vos, R. W. van Hamersvelt, M. A. Viergever, and I. Išgum, "Automatic coronary artery calcium scoring in cardiac CT angiography using paired convolutional neural networks," *Medical Image Analysis*, vol. 34, pp. 123–136, 2016.

[30] P. F. Christ, "Automatic liver and lesion segmentation in CT using cascaded fully convolutional neural networks and 3D conditional random fields," in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 415–423, Athens, Greece, October 2016.

[31] M. Gao, "Segmentation label propagation using deep convolutional neural networks and dense conditional random field," in *Proceedings of the 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, pp. 1265–1268, Prague, Czech Republic, April 2016.

[32] K. Kamnitsas, C. Ledig, V. F. J. Newcombe et al., "Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation," *Medical Image Analysis*, vol. 36, pp. 61–78, 2017.

[33] Y. Ledig, I. Kogan, E. Gelbart, O. Geva, and H. Greenspan, "Visualizing and enhancing a deep learning framework using patients age and gender for chest x-ray image retrieval," in *Proceedings of the SPIE on Medical Imaging*, p. 978510, San Diego, CA, USA, 2016.

[34] A. Janowczyk, A. Basavanhally, and A. Madabhushi, "Stain normalization using sparse autoencoders (StaNoSA):

application to digital pathology," *Computerized Medical Imaging and Graphics*, vol. 57, pp. 50–61, 2017.

[35] P. Kisilev, E. Sason, E. Barkan, and S. Hashoul, "Medical image description using multi-task-loss CNN," in *Deep Learning and Data Labeling for Medical Applications*, Springer, Cham, Switzerland, 2016.

[36] I. J. Goodfellow, "Generative adversarial networks," ArXiv14062661 Cs Stat, 2014, http://arxiv.org/abs/1406.2661.

[37] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1125–1134, Honolulu, HI, USA, July 2017.

[38] C.-C. Chang, "Adversarial learning for invertible steganography," *IEEE Access*, vol. 8, pp. 198425–198435, 2020.

[39] M. Sato, K. Hotta, A. Imanishi, M. Matsuda, and K. Terai, "Segmentation of cell membrane and nucleus by improving Pix2pix," *Biosignals*, pp. 216–220, 2018.

[40] M. Mori, T. Fujioka, L. Katsuta et al., "Feasibility of new fat suppression for breast MRI using pix2pix," *Japanese Journal of Radiology*, vol. 38, no. 11, pp. 1075–1081, 2020.

Hindawi
Journal of Healthcare Engineering
Volume 2021, Article ID 6633832, 14 pages
https://doi.org/10.1155/2021/6633832

Research Article

Utilizing Machine Learning Techniques to Predict the Efficacy of Aerobic Exercise Intervention on Young Hypertensive Patients Based on Cardiopulmonary Exercise Testing

Fangwan Huang[1], **Xiuyu Leng**[2], **Mohan Vamsi Kasukurthi**[3], **Yulong Huang**[4], **Dongqi Li**[3], **Shaobo Tan**[3], **Guiying Lu**[2], **Juhong Lu**[1], **Ryan G. Benton**[3], **Glen M. Borchert**[5] and **Jingshan Huang**[3,5]

[1]*College of Mathematics and Computer Science, Fuzhou University, Fuzhou 350108, China*
[2]*Department of Cardiology, First Affiliated Hospital of Sun Yat-sen University, Guangzhou 510080, China*
[3]*School of Computing, University of South Alabama, Mobile, AL 36688, USA*
[4]*College of Allied Health Professions, University of South Alabama, Mobile, AL 36688, USA*
[5]*Department of Pharmacology, College of Medicine, University of South Alabama, Mobile, AL 36688, USA*

Correspondence should be addressed to Xiuyu Leng; lengxy@mail.sysu.edu.cn

Received 27 December 2020; Revised 8 March 2021; Accepted 5 April 2021; Published 22 April 2021

Academic Editor: Hao Chun Lu

Recently, the incidence of hypertension has significantly increased among young adults. While aerobic exercise intervention (AEI) has long been recognized as an effective treatment, individual differences in response to AEI can seriously influence clinicians' decisions. In particular, only a few studies have been conducted to predict the efficacy of AEI on lowering blood pressure (BP) in young hypertensive patients. As such, this paper aims to explore the implications of various cardiopulmonary metabolic indicators in the field by mining patients' cardiopulmonary exercise testing (CPET) data before making treatment plans. CPET data are collected "breath by breath" by using an oxygenation analyzer attached to a mask and then divided into four phases: resting, warm-up, exercise, and recovery. To mitigate the effects of redundant information and noise in the CPET data, a sparse representation classifier based on analytic dictionary learning was designed to accurately predict the individual responsiveness to AEI. Importantly, the experimental results showed that the model presented herein performed better than the baseline method based on BP change and traditional machine learning models. Furthermore, the data from the exercise phase were found to produce the best predictions compared with the data from other phases. This study paves the way towards the customization of personalized aerobic exercise programs for young hypertensive patients.

1. Introduction

As a prevalent chronic disease, hypertension has been widely considered as a major risk factor for cardio-cerebrovascular events [1]. Strikingly, hypertension incidence is increasing most dramatically in young adults [2, 3]. As an alternative to antihypertensive drugs, lifestyle adjustments, including body weight control, diet, and exercise, can also be used to lower blood pressure (BP) [4, 5]. In particular, aerobic exercise not only directly reduces BP but also indirectly achieves similar effects by controlling body weight, reducing stress, and improving vascular endothelial function, along with other mechanisms [6–8]. Therefore, aerobic exercise intervention (AEI) has been widely recommended for the treatment of hypertension [9, 10]. Unfortunately, specific guidelines for effectively administering aerobic exercise aimed at anti-hypertension have not been widely accepted as there is significant individual variation in BP lowering achieved by the same exercise program, with the same exercise type, time, frequency, and duration [11–13]. Understanding the individual responsiveness to AEI before formulating comprehensive hypertension management plans will help to

improve both effectiveness and efficiency of BP management. To our knowledge, research in this field is still very limited, thus motivating us to perform the work conducted in this paper.

For the clinical feasibility and practicality, this work provided an investigation on the feasibility of utilizing machine learning techniques to predict the efficacy of AEI on young hypertensive patients. Taking into account the prognostic ability of key cardiopulmonary variables, data mining was performed based on the data generated by cardiopulmonary exercise testing (CPET) before treatment. CPET provides a comprehensive physiological assessment of multiorgan system function, including not only cardiovascular and pulmonary but also musculoskeletal and hematopoietic systems [14]. It can help clinicians identify the severity of the disease and evaluate the response to treatments, thus playing an important role in formulating aerobic exercise training prescription and cardiac rehabilitation [15, 16]. In this paper, CPET being used is an electric bicycle with many sensors (see Figure 1) as the main ergometer to measure the changes of various cardiopulmonary metabolic indicators over time. To provide the best measure of the response to exercise, these data were collected "breath by breath" by an oxygenation analyzer attached to a mask. The specific test scheme guided by clinicians included four phases: (1) resting for 1 minute to relieve the patient's tension; (2) load-free cycling (no resistance on the pedals) for 3 minutes to warm up; (3) exercise for 5–12 minutes with increasing resistance on the pedals (20–35 watt/min increment) until maximal exertion; and (4) recovery for 6 minutes with the first 3 minutes of load-free cycling and the second 3 minutes of sitting still.

Based on the professional advice of clinicians, this paper first utilized a simple method as the baseline to predict the BP-lowering effect of AEI for young hypertensive patients. Just to be clear, BP in this paper was equal to the sum of systolic blood pressure (SBP) and diastolic blood pressure (DBP). This method compared BP at the 6th minute of recovery (R6BP) with BP at the pre-exercise resting (PEBP) in a single CPET before AEI. Patients with R6BP $\leq$ PEBP were predicted to be strong responders to AEI. If the converse was true, they were predicted to be weak responders. Subsequent experiments showed that the accuracy of this method was typically 50%–60%, closely approximating a random guess, and far beneath the requirement for making effective and accurate clinical exercise prescriptions. To meet this challenge, machine learning techniques were utilized to fully capitalize on the information present within several cardiopulmonary metabolic indicators provided by CPET. As such, this work provides useful insights into the formulation of personalized AEI prescriptions for young hypertensive patients. The main contributions of this paper are as follows:

(i) A sparse representation classifier based on analytic dictionary learning was designed to accurately predict the efficacy of AEI on BP lowering. This model can not only alleviate the interference of redundant information and noise brought by breath-by-breath collection but also overcome the deficiency of the existing sparse representation-based classifier which needs a large number of training samples.

(ii) The significance of various cardiopulmonary metabolic indicators at different phases of CPET for this task was discussed through comparative experiments. The results showed that the data from the exercise phase can produce the best predictions compared with the data from other phases. Among various metabolic indicators, oxygen pulse (i.e., oxygen intake per heartbeat) was recommended as a powerful indicator for predicting the individual responsiveness to AEI.

Figure 1: Illustration of cardiopulmonary exercise testing. To protect privacy, the patient's eyes were partially blurred.

The remainder of the paper is structured as follows. Section 2 introduces various metabolic indicators of CPET used in this paper. Section 3 briefly introduces the related works, including the development of application scenarios and research methods. Section 4 describes the designed model in detail based on the shortcomings of the existing model. Section 5 reports the experimental results along with analyses. Finally, conclusions and future works are summarized in Section 6.

2. Main Metabolic Indicators of CPET

CPET provides time-varying information regarding multiple indicators related to circulation, respiration, and gas metabolism at different levels of exercise intensity [17]. The nine indicators recommended by professional clinicians for this work are briefly described in the following:

(1) Heart rate (HR): the number of heartbeats per minute. Normally, HR is 60–100 beats per minute at rest. HR varies individually according to age, sex, and other physiological factors.

(2) Stroke volume (SV): the volume of blood ejected from either ventricle of the heart in a single beat. The main affecting factors of SV are myocardial contractility, venous return blood volume (preload), arterial BP (afterload), and so forth.

(3) Cardiac output (CO): the volume of blood that flows out of the heart in a given period, usually denoted as liters per minute. It can be obtained by multiplying the average SV per beat by HR, varying with metabolism and activity. For example, it increases with muscle movement, emotional agitation, pregnancy, and so forth.

(4) Oxygen pulse (VO_2/HR): the volume of the oxygen intake per heartbeat. Hence, it is the amount of oxygen that the tissues of the body extract from oxygen carried by each SV. A higher oxygen pulse suggests better cardiopulmonary function. This can be used as a comprehensive index to determine the cardiopulmonary function.

(5) Oxygen consumption/kilogram (VO_2/kg): the volume of oxygen consumed by the metabolic processes of the body over a period of time, usually denoted as milliliters of oxygen per kilogram of body weight per minute. It reflects the body's ability to use oxygen and is usually determined by the maximum cardiac output, arterial oxygen content, cardiac output to the distribution index of the exercise muscle, and muscle oxygen capacity.

(6) Tidal volume (VT): the volume of air inhaled or exhaled during a normal breath. It is related to age, sex, volume and surface, breathing habits, body metabolism, and so forth.

(7) Ventilation volume/minute (VE): the volume of air inhaled or exhaled from the lungs in a minute, which can be obtained by multiplying VT by the respiratory rate.

(8) Respiratory exchange ratio (R): the ratio of the carbon dioxide (CO_2) output to the oxygen (O_2) uptake (i.e., VCO_2/VO_2) during the same period. It reflects not only the exchange of tissue metabolism of gas but also the influence of transient change in gas storage.

(9) Carbon dioxide ventilation equivalent (VE/VCO_2): the ability of the body to discharge carbon dioxide, calculated as the ratio between the required ventilation volume and carbon dioxide output.

To illustrate the characteristics of these indicators more vividly, Figure 2 shows a visualization of the above nine indicators for a patient during the exercise phase of a CPET before AEI. Since each breath represents a sampling point, the information of each metabolic indicator collected by the breath-by-breath technique can be stored as a time series [18].

3. Related Works

CPET is a dynamic, noninvasive diagnostic method to evaluate cardiopulmonary function during increasing load exercise. Recently, the application of CPET in clinical decision-making for various diseases has been significantly developed. For example, CPET is playing a growing role in cardiology, including heart failure, valve diseases, and ischemic heart disease [19]. Buys et al. evaluated the predictive value of CPET for the incidence of hypertension in patients undergoing aortic coarctation surgery and determined the high-risk boundary as VE/VCO_2 slope ≥ 27 and peak SBP ≥ 220 mmHg through Cox regression analysis [20]. Keller et al. suggested that BP overresponse in CPET might be a diagnostic tool for identifying high-risk groups of hypertension [21]. Besides, CPET can be used as a tool for preoperative risk stratification of patients (not limited to cardiopulmonary surgery) to predict postoperative adverse outcomes [22, 23]. Currently, one of the most impressive advances is that the integration of CPET and other tests has been introduced to diagnose several diseases [24]. Exercise stress echocardiography and CPET have been successfully combined in the dynamic assessment of heart failure for hypertensive patients [25]. Similarly, CPET combined with echocardiography of the right ventricle was applied to predict the prognosis of patients with pulmonary arterial hypertension [26].

From the perspective of research methods, in addition to traditional statistical analysis, data mining of CPET using machine learning techniques is gradually becoming a research hotspot. Leopold et al. developed a greedy heuristic algorithm based on feature clustering to study the ability of CPET to predict the anaerobic mechanical power outputs [27]. Braccioni et al. used a random forest algorithm to analyze the relationship between symptoms and cardiopulmonary parameters of lung transplant recipients based on incremental CPET [28]. Sakr et al. evaluated the performance of six machine learning techniques in predicting the individuals at risk of hypertension through treadmill stress tests on a massive crowd [29]. Unfortunately, the above work only selected some special values of cardiopulmonary metabolic indicators (such as peaks or slope) as features for analysis, without taking into account their time-varying characteristics. Our previous work has proved that time-varying data of some metabolic indicators obtained through CPET could be used to predict the efficacy of AEI [30], but how to further improve the predictive accuracy is still a challenge, especially in the case of insufficient training samples. This encourages us to perform the research conducted in this paper.

In fact, the prediction of the BP-lowering effect of AEI by using a certain metabolic indicator can be transformed into time series classification (TSC) for data mining. To date, researchers have proposed hundreds of approaches for TSC in different application scenarios. TSC algorithms can be roughly divided into seven categories: (1) the whole-series-based method, (2) the interval-based method, (3) the shaped method, (4) the word-frequency-based method, (5) the model-based method, (6) the integration-based

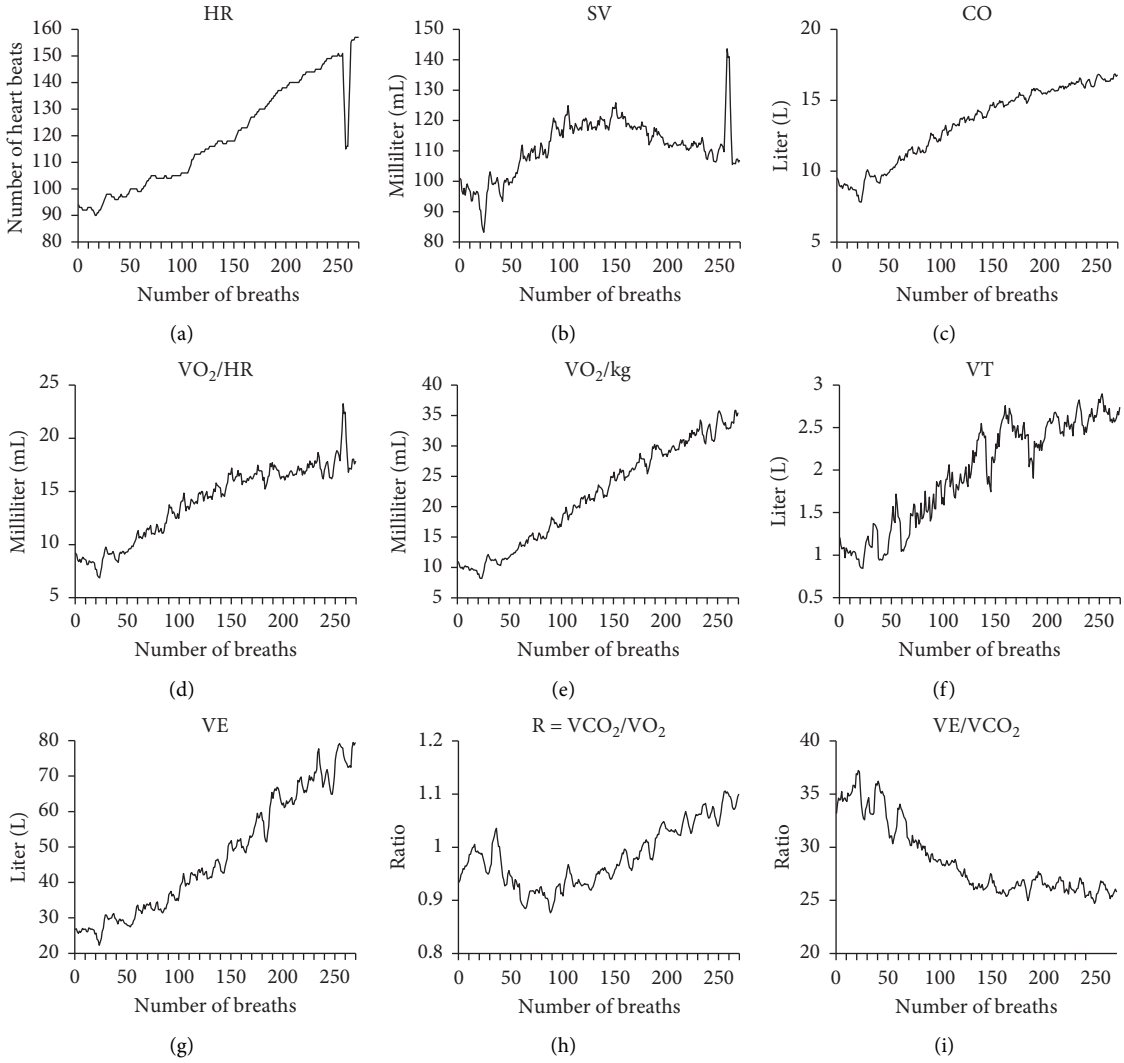


Figure 2: Visualization of several indicators for a patient during the exercise phase of a CPET. (a) Heart rate. (b) Stroke volume. (c) Cardiac output. (d) Oxygen pulse. (e) Oxygen consumption/kilogram. (f) Tidal volume. (g) Ventilation volume/minute. (h) Respiratory exchange ratio. (i) Carbon dioxide ventilation equivalent.

method, and (7) the deep learning-based method. Bagnall et al. evaluated the latest progress of TSC algorithms on 85 datasets in the University of California, Riverside (UCR) archive [31]. They recommended 1-nearest neighbor with dynamic time warping (1NN-DTW) and random forest (RF) as the baseline classifiers for comparison with other classifiers. Besides, they also concluded that the integration-based method can achieve high accuracy by utilizing multiple classifiers on one or more feature spaces. For example, Bagnall et al. integrated 35 classifiers on the time, frequency, change, and shapelet transformation domains [32]. On this basis, Lines et al. added two new classifiers, two additional transformation domains, and a hierarchical structure of probability voting to further improve the performance [33]. Recently, the method based on deep learning has gradually become a research hotspot [34]. Deep learning is characterized by learning hidden and more abstract representations of data from the original time series to achieve better classification performance. This method is widely used for end-to-end learning including methods such as convolutional neural networks (CNNs) [35] and echo state network (ESN) [36]. The common disadvantage of these methods is that they require a large amount of data and computational cost for model training. As this work represents the first stage in a larger experiment, the relatively small number of samples means that the above approach is not appropriate. Moreover, the robustness of the method to signal-to-noise ratio also needs to be considered because the process involved in collecting CPET data is usually very noisy. For the above reasons, a classifier based on sparse representation is recommended for the task in this paper.

4. Sparse Representation-Based Classifier

In this section, a sparse representation classifier based on dictionary learning was designed to accurately predict the efficacy of AEI on BP lowering. This method firstly

eliminated redundant information and reduced noise by feature extraction based on the sparse representation. At the same time, it took advantage of learning of an analytic dictionary without requiring as many training samples as the existing sparse representation-based classifier.

4.1. Brief Introduction for Sparse Representation. Recently, sparse representation has received increasing attention in many fields. While initially developed for use in image analysis and signal processing, sparse representation has been successfully utilized for dealing with more general tasks in the machine learning field [37]. Specifically, given a signal $x \in R^m$ of m observations and an overcomplete dictionary $D \in R^{m \times n}$ $(n \gg m)$ in which the column vector d_i $(1 \le i \le n)$ is known as an atom, the main goal of the sparse representation is the reconstruction of a signal perfectly with the least possible number of atoms. Its objective function is as follows:

$$\min_\alpha \|\alpha\|_0 \ s.t.\ x = D\alpha \ , \tag{1}$$

where $\alpha \in R^n$ is the sparse representation (or sparse solution) of x and $\|\cdot\|_0$ refers to the number of nonzero elements in α. Due to the noise in the real signal, the solution of equation (1) can be approximated by either of the following two equations:

$$\min_\alpha \|\alpha\|_0 \ s.t.\ \|x - D\alpha\|_2^2 \le \delta, \tag{2}$$

$$\min_\alpha \|x - D\alpha\|_2^2 \ s.t.\ \|\alpha\|_0 \le k, \tag{3}$$

where δ can be considered as noise or a reconstruction residual; the sparse factor k is a predefined integer not less than 1. Besides, based on the Lagrange multiplier theorem, solving sparse representation can be equivalently transformed into an unconstrained minimization problem:

$$\min_\alpha \|x - D\alpha\|_2^2 + \lambda \|\alpha\|_0, \tag{4}$$

where λ is a positive constant used to achieve a tradeoff between the reconstruction residual and the sparse solution.

It should be noted that since obtaining the optimal solution with l_0-norm minimization is an NP-hard problem, many algorithms have been proposed to deal with it. The strategies commonly used in these algorithms mainly include greedy pursuit strategy and convex relaxation strategy [38, 39]. The greedy pursuit strategy represented by the orthogonal matching pursuit (OMP) algorithm is to gradually approach the optimal solution through the sequential selection of column vectors (atoms) until the end of iteration [40]. For the convex relaxation strategy, the main idea is to replace the l_0-norm minimization term with the l_1-norm minimization term. Taking equation (3) as an example, it can be approximately equivalent to the lasso problem:

$$\min_\alpha \|x - D\alpha\|_2^2 \ s.t.\ \|\alpha\|_1 \le \varepsilon, \tag{5}$$

where $\|\cdot\|_1$ represents the sum of the absolute values of nonzero elements in α and ε is a positive constant given beforehand. The advantage of this strategy is that the l_1-norm minimization problem has an analytical solution and can be effectively solved by several methods, such as least angle regression (LAR) [41], coordinate descent algorithm (CDA) [42], iterative shrinkage-thresholding algorithm (ISTA) [43], and many variations of them.

4.2. The Existing Sparse Representation-Based Classifier. Proposed by Wright et al., a sparse representation-based classifier (SRC) was first applied in the field of face recognition and then successfully extended to TSC [44, 45]. Specifically, the sparse representation of an unlabeled sample is first solved based on the dictionary composed of all labeled samples. Then, the reconstruction residuals of each class are calculated by using the samples of each class and the corresponding elements in the sparse representation. Finally, the classification is performed by examining which class leads to the minimum residual of the unlabeled sample. The steps to implement SRC are as follows:

(1) The l_2-norm normalization is preprocessed for each sample of the whole dataset with a class number of c.

(2) A dictionary $D = [D_1, \cdots, D_j, \cdots, D_c]$ is generated, where D_j $(1 \le j \le c)$ is a subdictionary composed of jth-class normalized samples in the training set as column vectors (atoms).

(3) The sparse representation α of the unlabeled sample y is obtained by using the algorithm described above.

(4) The unlabeled sample y is reconstructed, respectively, using each D_j and corresponding α_j, where α_j $(1 \le j \le c)$ is a subvector consisting of the elements in α corresponding to all atoms in D_j. The label is determined based on the minimum residual, as shown in the following equation:

$$\text{Label}_y = \text{class}_j \ s.t.\ \forall 1 \le i \le c \text{ and } i \ne j, \quad \left\|y - D_i\alpha_i\right\|_2^2 > \left\|y - D_j\alpha_j\right\|_2^2. \tag{6}$$

Figure 3 shows the SRC schematic for a two-class problem. The success of the SRC depends on the hypothesis that the unlabeled sample can be best reconstructed by a linear representation of samples within the same class. However, once the samples of different classes look similar to each other, the performance of SRC is very unstable [46]. Besides, the dictionary cannot satisfy the overcompleteness if the number of labeled samples is less than the dimension of samples, which will also affect the performance of the SRC [47]. To overcome the shortcoming of the SRC, a sparse representation classifier based on an analytic dictionary was designed, and then its accuracy was improved by using dictionary learning. For the sake of simplicity, the model was called SRC-AL for short. The principle is described in the following.

4.3. The Designed Sparse Representation-Based Classifier. In the application domain of sparse representation, an overcomplete dictionary can be usually generated using data implementation or analytic approach [48]. The approach based on data implementation is to construct an explicit dictionary

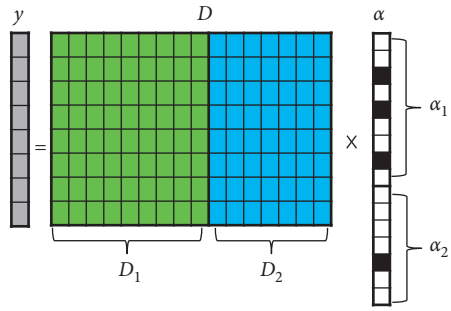


Figure 3: The SRC schematic for a two-class problem, where y is the unlabeled sample. D_1 and D_2 represent subdictionaries composed of all normalized samples belonging to class 1 and class 2, respectively. The sparse representation α can be divided into two subvectors (α_1 and α_2) according to the number of columns in D_1 and D_2. The black-filled blocks in α represent nonzero elements.

directly by using the raw data. This is exactly the way adopted by the SRC, intending to obtain the residuals of the unlabeled sample reconstructed by the samples of different classes. Unlike SRC, SRC-AL generates an implicit dictionary based on the analytic approach as the initial dictionary. This approach generally utilizes some fixed transformations, such as discrete Fourier transform (DFT), discrete cosine transform (DCT), and discrete wavelet transform (DWT) [49]. Compared to the data implementation, the analytic approach has the advantage of allowing an overcomplete dictionary of any size without being limited by the number of labeled samples. However, due to the poor adaptability, the analytic dictionary often requires further optimization through dictionary learning. K-singular value decomposition (K-SVD) is a popular algorithm for dictionary learning, which updates the used atoms one by one in an iterative manner to train the overcomplete dictionary most suitable for the training set [50].

Inspired by the sparse representation predictor for time series proposed by our previous work [51], the workflow of SRC-AL consists of the following six steps:

(1) Generate an initial dictionary $D \in R^{(m+c)\times n}$ by utilizing the analytic approach, where m is the dimension of the sample, c is the number of classes, and n is an arbitrary integer much larger than $(m+c)$. The upper and lower parts of the dictionary are represented by $D_{up} \in R^{m\times n}$ and $D_{lw} \in R^{c\times n}$, respectively.

(2) Normalize each sample of the training dataset with l_2-norm, and convert its label into one-hot encoding. Combine the above two parts into the new training sample $x \in R^{(m+c)}$.

(3) According to the training set composed of new samples, update the initial dictionary through dictionary learning, with the purpose of better reconstructing the samples. The objective function of dictionary learning can be described as

$$\min_{\alpha^i, D} \sum_{i=1}^{r} \left\| x^i - D\alpha^i \right\|_2^2 \quad \text{s.t.} \quad \left\| \alpha^i \right\|_0 \le k, \tag{7}$$

where r is the number of samples in the training set and α^i is a sparse representation of sample x^i.

(4) Normalize the unlabeled sample $y \in R^m$ with l_2-norm, and then obtain its sparse representation $\alpha_y \in R^n$ based on the upper part of the learned dictionary ($D_{up}' \in R^{m\times n}$).

(5) Multiply the lower part of the learned dictionary ($D_{lw}' \in R^{c\times n}$) by the sparse representation $\alpha_y \in R^n$ to obtain the label vector $L_y \in R^c$.

(6) Determine the label of y according to the index of the element with the largest absolute value in L_y, as shown in the following equation:

$$\text{Label}_y = \text{class}_j \ \text{s.t.} \ \forall\, 1 \le i \le c \ \text{and} \ i \ne j, \quad \left| L_y(i) \right| < \left| L_y(j) \right|, \tag{8}$$

where $L_y(i)$ represents the ith element in vector L_y.

Figure 4 shows the SRC-AL schematic for a two-class problem. Assuming that sample x_1 belongs to class 1, the green-filled blocks represent the normalized sample, and the following "10" represents the one-hot encoding of the label. Similarly, the blue-filled blocks represent the normalized sample of class 2, and the following "01" represents the one-hot encoding of its label. The dictionary filled with orange is generated by the analytic approach. To better reconstruct all training samples, a dictionary-learning algorithm (such as K-SVD) should be applied to constantly update the dictionary. Based on the upper part of the learned dictionary (D_{up}'), the sparse representation α_y of the unlabeled sample y (grey-filled blocks) is solved, and then $D_{lw}' \times \alpha_y$ is used to obtain the label vector L_y. Finally, the element with the largest absolute value in L_y is set to 1, and the other elements are set to 0. This one-hot encoding is used to replace the question mark in Figure 4 to achieve the classification of y.

5. Experiments and Results

CPET data from 24 young patients with stage I hypertension before AEI treatment were used for the experiments. The dataset was provided by the Department of Cardiology, First Affiliated Hospital of Sun Yat-sen University, China. The whole exercise process of all the people was completed under the supervision of professional medical staff in the hospital. Blood pressure before and after exercise was assessed using both dynamic and exercise blood pressure results. Although the cost of each sample is very large, the data are highly comparable and reliable due to the guaranteed amount of exercise and more comprehensive monitoring indicators.

The performance of various machine learning models based on the data from the exercise phase was compared with the baseline method given by the clinician. Note that the baseline method only focused on BP change between pre-exercise and postexercise within a single CPET, while the machine learning model took into account the time series of metabolic indicators during CPET. After verifying the effectiveness of the designed model, the significance of the data from different phases in CPET for

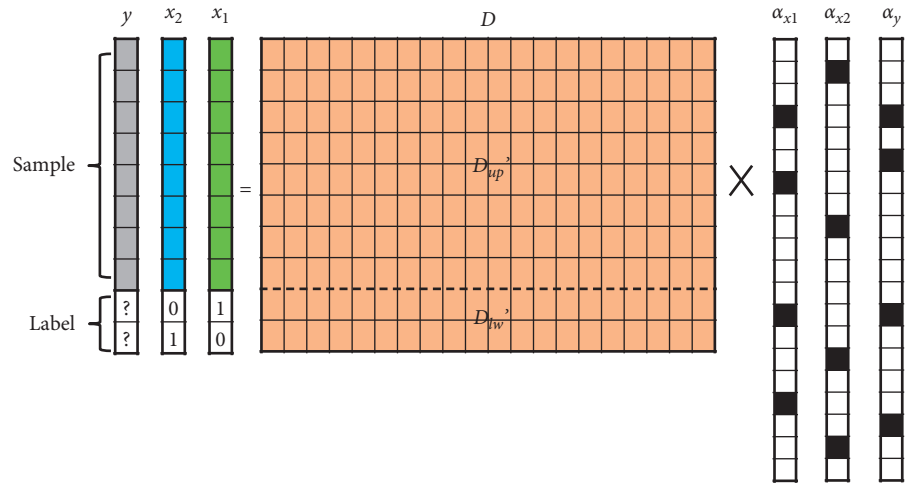


Figure 4: The SRC-AL schematic for a two-class problem, where x_1 and x_2 represent the samples of class 1 and class 2, respectively, and y is the unlabeled sample. Due to the limited space, only two labeled samples are drawn in the figure. In fact, all labeled samples are used for dictionary learning. D_{up}' and D_{lw}' represent the upper and lower parts of the learned dictionary. α_{x1}, α_{x2}, and α_y are the sparse representations of x_1, x_2, and y where the black-filled blocks represent nonzero elements.

predicting the efficacy of AEI on BP lowering was further evaluated.

5.1. Description of the Dataset

(1) Inclusion criteria: between the ages of 18 and 45; stage I hypertension (SBP: 140–160 mmHg; DBP: 90–100 mmHg) either without medication or with discontinuation of antihypertensive drugs for more than two weeks and still presenting stage I hypertension; no regular exercise for four months prior to admission; willingness to participate in follow-ups for more than 6 months.

(2) Treatment prescription: patients underwent aerobic exercise with an Italian COSMED K4 electric bicycle. Training intensity corresponded to the metabolic equivalent of task (MET) of 70% of maximal oxygen consumption (VO_{2max}). Get aerobic exercise 5 times per week, each time 45 minutes (exercise intensity equivalent to 2,000–3,000 kcal per week), lasting 12 weeks.

(3) Classification standard: patients were categorized as strong or weak responders of AEI treatment according to the therapeutic effect. The classification process is as follows:

  (1) All patients received 24-hour dynamic BP monitoring before and after AEI to obtain their daily mean BP.

  (2) The rate of BP change before and after treatment was calculated for each patient: $r_i = \|\text{MBPB} - \text{MBPA}\|/\text{MBPB}$, where MBPB and MBPA indicated the mean BP of 24 hours before and after treatment, respectively.

  (3) Z-score standardization was performed for r_i as follows: $z_i = (r_i - \mu)/\sigma$, where μ and σ were the mean and standard deviation, respectively. The role of z_i was to determine whether the antihypertensive efficacy of the ith patient was above average.

  (4) Classify according to z_i. Patients with $z_i > 0$ (14 individuals in total) were identified as the strong antihypertensive responders of AEI, while patients with $z_i < 0$ (10 individuals in total) were classified as weak responders. The real labels of 24 patients are detailed in Table 1.

As can be seen from Table 1, all patients except the last one exhibited certain antihypertensive effects following 12 weeks of AEI treatment. The average antihypertensive change rate was 7.582%. The individual showing the best antihypertensive effect exhibited a 40 mmHg (or 16.529%) BP decrease after AEI. However, the absence of obvious changes in BP of some individuals also proved that the efficacy of AEI is significantly different in hypertensive patients.

5.2. Experimental Results. In this paper, accuracy and F1-score (the harmonic average of precision and recall) obtained by the confusion matrix (see Figure 5) were used to evaluate the performance of the model. For them, higher values indicate positive benefits.

5.2.1. The Performance of the Baseline Method Based on BP Change. An intuitive way to predict the BP-lowering effect of AEI is to determine whether the BP of patients after exercise is lower than that before exercising in CPET. Specifically, the pre-exercise resting BP (PEBP) was subtracted from BP at the 6th minute of the recovery phase (R6BP) to obtain BP change ($\triangle$BP) for each patient. A patient with $\triangle$BP less than 0 was considered to be unable to benefit from AEI, meaning the predicted label was weak. Conversely, a patient would exhibit a strong, beneficial antihypertensive response to AEI. The predicted labels of the

Table 1: Real labels of 24 young patients with stage I hypertension.

Sample	MBPB (mmHg)	MBPA (mmHg)	r_i (%)	z_i	Label
1	242	202	16.529	2.209	Strong
2	229	195	14.847	1.792	Strong
3	221	192	13.122	1.365	Strong
4	241	212	12.033	1.094	Strong
5	235	213	9.362	0.432	Strong
6	253	230	9.091	0.365	Strong
7	223	203	8.969	0.334	Strong
8	249	227	8.835	0.301	Strong
9	209	191	8.612	0.246	Strong
10	244	223	8.607	0.245	Strong
11	214	196	8.411	0.196	Strong
12	246	226	8.130	0.127	Strong
13	204	188	7.843	0.055	Strong
14	244	225	7.787	0.041	Strong
15	244	226	7.377	−0.060	Weak
16	231	214	7.359	−0.065	Weak
17	240	223	7.083	−0.133	Weak
18	231	215	6.926	−0.172	Weak
19	214	207	3.271	−1.079	Weak
20	221	214	3.167	−1.104	Weak
21	222	216	2.703	−1.220	Weak
22	211	208	1.422	−1.537	Weak
23	207	205	0.966	−1.650	Weak
24	207	208	0.483	−1.770	Weak

Confusion matrix		Real label		$\text{Accuracy} = \frac{A+D}{A+B+C+D}$
		Strong	Weak	$\text{F1-score} = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$
Predicted label	Strong	A	B	$\text{Precision} = \frac{A}{A+B}$
	Weak	C	D	$\text{Recall} = \frac{A}{A+C}$

Figure 5: Illustration of the confusion matrix.

baseline method based on BP change are shown in Table 2. Using the confusion matrix, the accuracy of the baseline method was 0.542, and F1-score was 0.56. This meant that the baseline method was only slightly superior to the random guess (accuracy = 0.5), far less than the requirement for clinical applications.

5.2.2. The Performance of Machine Learning Models Based on the Metabolic Indicators. Time series of the nine metabolic indicators described above during the exercise phase were selected for analysis using machine learning models. Of note, patients had distinct exercise durations based on different physical conditions, resulting in different numbers of sampling points for individuals (ranging from 85 to 270). As most machine learning models required samples to have the same dimension, linear interpolation was first applied to unify the sampling numbers of all patients to 270 points. Afterward, SRC-AL presented herein was compared with SRC and some popular models of TSC, including 1NN-DTW, random forest (RF), and support vector machine (SVM). Due to the limited samples, the leave-one-out cross-validation was adopted to carry out the experiments [52]. All the above models were implemented by MATLAB. For SRC and SRC-AL, OMP and K-SVD algorithms in the SPAMS toolbox were used to solve the sparse representation and dictionary learning, respectively. Besides, the optimal sparse factor was obtained by grid search in a specific interval. Finally, for SRC-AL, the size of the initial dictionary was defined as a matrix where the number of columns was twice the number of rows, which was realized by the discrete cosine transform. The experimental results of each model are shown in Tables 3 and 4, where the last column of each table shows the average performance of each metabolic indicator based on different machine learning models.

5.2.3. The Performance of SRC-AL Based on the Data from Different Phases of CPET. Since SRC-AL performed best in the above model, it was directly used to evaluate the significance of the data generated in the three important phases of CPET for predicting the individual responsiveness to AEI. These three phases included warm-up, exercise, and recovery. Similar to the exercise phase, the data dimensions of different patients in the other two phases were also inconsistent. For the warm-up phase, the shortest time series of metabolic indicators had only 38 sample values, while the longest had 81 sample values. For the recovery phase, the shortest one had only 113 sample values, while the longest one had 195 sample values. Therefore, linear interpolation should be used first to unify the data dimensions of different patients into the same. Besides, the dictionary learned in the exercise phase cannot be applied to the other two phases due to different data dimensions. The experimental results of SRC-AL based on the data of the above three phases of CPET are shown in Table 5.

5.3. Analyses of Experimental Results. This work investigated the ability of metabolic indicators to discriminate between strong and weak responses to AEI in patients. Through the analysis of the above experimental results, the following insights can be obtained to help clinicians predict the efficacy of AEI on young hypertensive patients based on CPET.

(1) From Tables 3 and 4, SRC-AL and SRC were superior to other traditional classifiers in predicting the individual responsiveness to AEI based on the time series of metabolic indicators. This is mainly because the process of collecting these metabolic indicator data is prone to generate many interference signals, while the sparse representation can effectively extract the main features of time series and enhance the robustness to noise to the maximum extent.

(2) The performance of SRC-AL was significantly better than that of SRC regardless of the time series based on any indicator, although both were based on sparse representation. This indicates that SRC needs an adequate set of training samples to form an overcomplete dictionary for better performance. On the contrary, SRC-AL can always guarantee the overcompleteness because it generates dictionaries

Table 2: Predicted labels of the baseline method based on BP change, where △BP = PEBP − R6BP.

Sample	PEBP (mmHg)	R6BP (mmHg)	△BP (mmHg)	Predicted label	Real label
1	242	267	−25	Weak	Strong
2	233	260	−27	Weak	Strong
3	199	170	29	Strong	Strong
4	236	216	20	Strong	Strong
5	190	223	−33	Weak	Strong
6	238	246	−8	Weak	Strong
7	209	205	4	Strong	Strong
8	274	256	18	Strong	Strong
9	201	203	−2	Weak	Strong
10	224	214	10	Strong	Strong
11	204	211	−7	Weak	Strong
12	262	255	7	Strong	Strong
13	216	218	−2	Weak	Strong
14	219	194	25	Strong	Strong
15	222	232	−10	Weak	Weak
16	245	229	16	Strong	Weak
17	247	250	−3	Weak	Weak
18	176	173	3	Strong	Weak
19	233	225	8	Strong	Weak
20	219	227	−8	Weak	Weak
21	205	206	−1	Weak	Weak
22	223	235	−12	Weak	Weak
23	216	242	−26	Weak	Weak
24	224	185	39	Strong	Weak

Table 3: Accuracy of various machine learning models based on the data from the exercise phase of CPET.

Indicator	SRC-AL	SRC	1NN-DTW	RF	SVM	Mean
VO_2/HR	1.000	0.792	0.625	0.583	0.667	0.733
VE	0.875	0.708	0.625	0.458	0.542	0.642
VO_2/kg	0.917	0.667	0.500	0.500	0.583	0.633
CO	0.917	0.667	0.500	0.583	0.792	0.692
HR	0.833	0.625	0.500	0.500	0.417	0.575
SV	0.667	0.542	0.458	0.542	0.542	0.550
VE/VCO_2	0.708	0.542	0.500	0.458	0.458	0.533
VT	1.000	0.458	0.583	0.500	0.333	0.575
R	0.958	0.417	0.375	0.500	0.583	0.567

Table 4: F1-score of various machine learning models based on the data from the exercise phase of CPET.

Indicator	SRC-AL	SRC	1NN-DTW	RF	SVM	Mean
VO_2/HR	1.000	0.815	0.667	0.667	0.714	0.773
VE	0.903	0.741	0.640	0.581	0.593	0.692
VO_2/kg	0.933	0.692	0.539	0.600	0.643	0.681
CO	0.923	0.692	0.571	0.667	0.828	0.736
HR	0.846	0.640	0.539	0.600	0.462	0.617
SV	0.667	0.667	0.552	0.645	0.621	0.630
VE/VCO_2	0.778	0.667	0.600	0.606	0.581	0.646
VT	1.000	0.381	0.615	0.571	0.429	0.599
R	0.966	0.462	0.516	0.571	0.643	0.632

employing the analytic approach. Through dictionary learning, the initial dictionary can be gradually updated to better fit the training samples and their labels.

(3) According to the last column of Table 3, except for the indicator VE/VCO_2, the average accuracy of all the other metabolic indicators based on the five machine learning models was higher than that of the baseline method based on BP change (accuracy = 0.542). However, if evaluated by the average F1-score, all metabolic indicators were superior to BP change alone (F1-score = 0.56), as shown in the last column of Table 4. This interesting finding suggests that the multipoint characteristics of cardiopulmonary metabolic indicators formed by collecting breath data can more accurately reflect the individual responsiveness to AEI. Figure 6 visualizes the comparison between the predictive performance of each indicator obtained by machine learning models and that of BP change obtained by the baseline method, where Figure 6(a) shows the average/optimal accuracy and Figure 6(b) shows the average/optimal F1-score. Note that the optimal performance of all metabolic indicators was obtained by SRC-AL designed herein.

(4) Table 5 illustrates the significance of data from different phases in CPET for predicting the BP-lowering effect of AEI. VO_2/HR, VE, VO_2/kg, VT, and R had the best predictive effect by using the time series of the exercise phase, while HR, SV, and VE/VCO_2 performed better according to the time series of the warm-up phase. The performance of CO was consistent in both the exercise and the warm-up phases. Finally, the data in the recovery phase were less important than in the previous two phases. The reason may be that the patient is only active for the

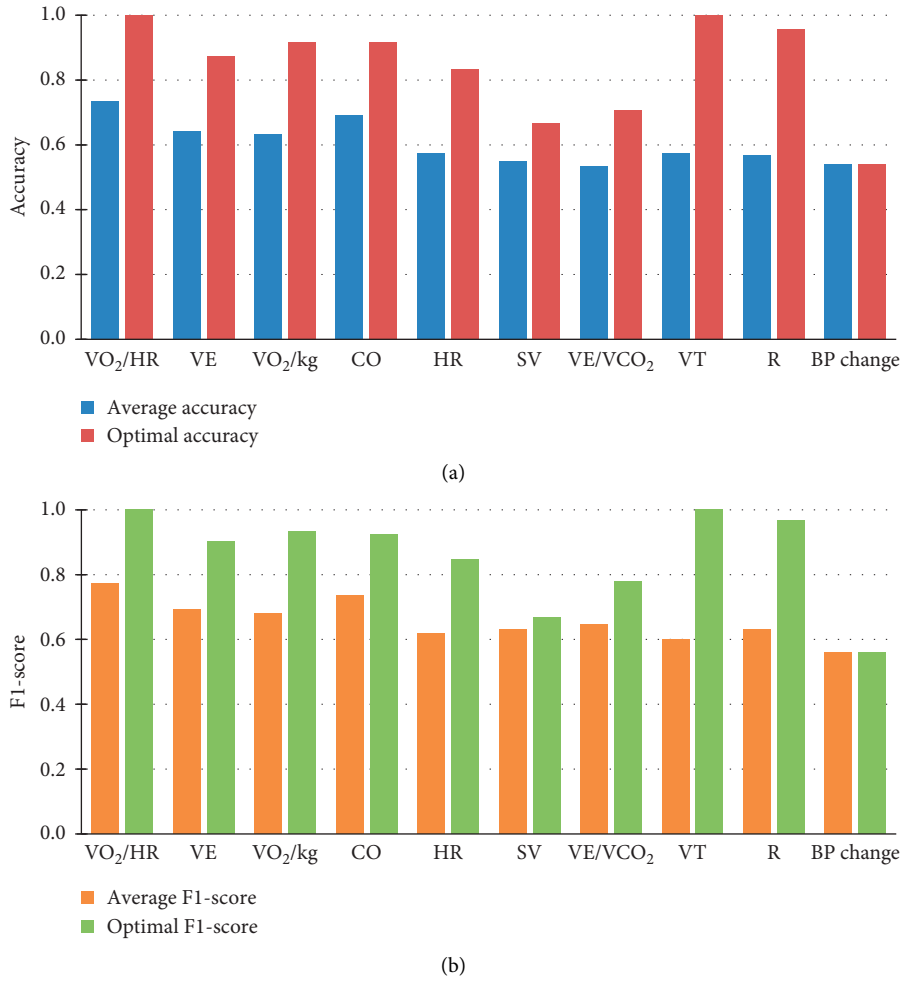


Figure 6: Comparison between the predictive performance of each indicator obtained by machine learning models and that of BP change obtained by the baseline method. (a, b) The average and optimal performance of metabolic indicators, respectively.

Table 5: The performance of SRC-AL based on the data from the three phases of CPET.

Indicator	Warm-up accuracy	Exercise accuracy	Recovery accuracy	Warm-up F1-score	Exercise F1-score	Recovery F1-score
VO_2/HR	0.583	1.000	0.625	0.737	1.000	0.743
VE	0.625	0.875	0.833	0.757	0.903	0.875
VO_2/kg	0.750	0.917	0.750	0.824	0.933	0.824
CO	0.917	0.917	0.708	0.923	0.923	0.720
HR	0.833	0.833	0.792	0.857	0.846	0.828
SV	0.958	0.667	0.625	0.963	0.667	0.757
VE/VCO_2	0.750	0.708	0.708	0.800	0.778	0.788
VT	0.583	1.000	0.792	0.737	1.000	0.828
R	0.917	0.958	0.708	0.923	0.966	0.759

first three minutes during the recovery phase and remains inactive for the next three minutes. In other words, the metabolic data of patients in the active state are more significant for predicting the individual responsiveness to AEI.

5.4. Additional Experiments. Considering that the sample size of the aforementioned experiments is rather limited, six datasets from the UCR Time Series Classification Archive were selected for additional experiments to further verify the effectiveness of our proposed model [53]. The common characteristics of these datasets include the following: (1) the number of classes is two and (2) the number of training samples is less than or close to the length of the sample. This results in the dictionary based on data implementation not being overcomplete, which may reduce the classification accuracy of SRC. Nevertheless, the proposed model based on analytic dictionary learning (SRC-AL) should not be affected. The detailed description

Table 6: Brief description of six datasets from the UCR Time Series Classification Archive.

Type	Dataset	Classes	Length	Training set	Testing set
ECG	ECGFiveDays	2	136	23	861
ECG	ECG2000	2	96	100	100
Sensor	SonyAIBORobotSurface1	2	70	20	601
Spectro	Ham	2	431	109	105
Image	Herring	2	512	64	64
Image	BeetleFly	2	512	20	20

Table 7: Classification accuracy of various machine learning models on UCR datasets.

Dataset	SRC-AL	SRC	1NN-DTW	RF	SVM
ECGFiveDays	0.974	0.971	0.768	0.787	0.974
ECG2000	0.920	0.900	0.770	0.819	0.770
SonyAIBORobotSurface1	0.890	0.757	0.725	0.733	0.677
Ham	0.762	0.619	0.467	0.722	0.619
Herring	0.672	0.609	0.531	0.572	0.625
BeetleFly	0.900	0.650	0.700	0.825	0.900

Table 8: Accuracy of improved versions of various machine learning models based on the data from the exercise phase of CPET.

Indicator	SRC-AL	1NN-sharpDTW (RAWS)	1NN-sharpDTW (DWT)	1NN-sharpDTW (slope)	TSF
VO_2/HR	1.000	0.625	0.583	0.542	0.675
VE	0.875	0.583	0.583	0.667	0.392
VO_2/kg	0.917	0.458	0.500	0.500	0.450
CO	0.917	0.833	0.833	0.667	0.558
HR	0.833	0.500	0.542	0.542	0.517
SV	0.667	0.583	0.625	0.542	0.592
VE/VCO_2	0.708	0.625	0.583	0.542	0.542
VT	1.000	0.542	0.542	0.625	0.392
R	0.958	0.583	0.583	0.625	0.517

of the datasets is shown in Table 6. According to the results demonstrated in Table 7, SRC-AL achieved the best classification in all the datasets, indicating that SRC-AL is particularly suitable for datasets with fewer training samples than the sample length.

In addition, considering that SRC-AL is an extended sparse representation classifier, an interesting question is whether or not other machine learning models can be modified to handle the problem addressed in this paper with better performance. To answer this question, the improved versions of some machine learning models were used to be compared with SRC-AL. For example, in order to reduce the huge feature space of the random forest, time series forest (TSF) was used to divide a time series into $\sqrt{m}$ random intervals (m is the length of the time series), and then the mean, standard deviation, and slope of each interval were all taken as features for classification [54]. Similarly, in order to improve the classification accuracy of 1NN-DTW, 1NN-sharpDTW was first adopted to convert the time series into a sequence of shape descriptors, and then the locally similar structures were paired [55]. Aiming to extract different characteristics of the domain data, three description functions were utilized to encode local shape information in this paper: raw subsequence (RAWS), discrete wavelet transform (DWT), and slope. Specifically, RAWS was applied to directly take a subsequence of the data around a sampling point of a time series as its shape descriptor. On this basis, DWT was used to decompose each subsequence into three levels, and then all the coefficients were serialized into a shape descriptor. Alternatively, the slope function was first adopted to divide each subsequence into several intervals, and then the slopes of the fitting lines of all the intervals were concatenated into a shape descriptor. According to the results shown in Table 8, SRC-AL performed best in all the improved versions. This fully demonstrates the significance of sparse representation in feature extraction and noise reduction of CPET data.

6. Conclusions and Future Works

In recent years, the incidence of hypertension has shown a clear trend towards presenting in younger patients. Note that AEI has been recognized as an effective treatment among young hypertensive patients. Unfortunately, research regarding how to predict the individual responsiveness to AEI for young hypertensive patients is still lacking. As such, a sparse representation classifier based on analytic dictionary learning, a.k.a. SRC-AL, was designed to mine the time series of multiple cardiopulmonary metabolic indicators from CPET data to accurately estimate the effectiveness of AEI on patients' BP management.

In summary, the experimental results first showed that the machine learning model, especially SRC-AL, which is based on the time series of metabolic indicators, can better

predict the individual responsiveness to AEI than the baseline method that is based on scalar values of BP change alone. Secondly, data from the exercise phase in CPET are the first choice for data mining, with the second choice being data from the warm-up phase. Thirdly, VO_2/HR is strongly recommended as a powerful, new prognostic indicator for predicting aerobic exercise efficacy as an antihypertensive, with an average accuracy of about 75% and up to 100%. Besides, CO is also a good choice not only because its average performance is second only to VO_2/HR but also due to the fact that its performance is very stable in both warm-up and exercise phases. As such, this will likely prove to be useful to clinicians for more accurately selecting comprehensive antihypertensive treatment measures without requiring extra clinical testing.

Note that the predictive model in this study is a qualitative prediction that predicts whether or not an individual hypertensive patient's response to aerobic exercise intervention is ideal. In future work, the quantitative prediction model of BP reduction caused by AEI is planned to be studied. Besides, BP defined in the current model is the sum of SBP and DBP. It may make more sense to analyze SBP and DBP separately in the subsequent work. Finally, the work presented here includes data generated from 24 young patients with stage I hypertension. Due to the limited sample size of this dataset, more samples should be collected in the future to prove the robustness of the proposed method. At the same time, further optimization can be attempted through the data augmentation technologies.

Data Availability

The data used to support the findings of this study cannot be made freely available in order to protect patient privacy. Requests for access to these data should be made to the corresponding author.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The authors wish to thank the research team of the Department of Cardiology, First Affiliated Hospital of Sun Yat-sen University, for providing the experimental data. This work was supported by the National Natural Science Foundation of China (Grant nos. 61772136 and 61672159) and the Natural Science Foundation of Fujian Province (Grant no. 2018J07005).

References

[1] R. M. Carey and P. K. Whelton, "Prevention, detection, evaluation, and management of high blood pressure in adults: synopsis of the 2017 American college of cardiology/American heart association hypertension guideline," *Annals of Internal Medicine*, vol. 168, no. 5, pp. 351–358, 2018.

[2] Y. Yano, J. P. Reis, L. A. Colangelo et al., "Association of blood pressure classification in young adults using the 2017 American College of Cardiology/American Heart Association blood pressure guideline with cardiovascular events later in life," *Journal of the American Medical Association*, vol. 320, no. 17, pp. 1774–1782, 2018.

[3] S. Wu, Y. Song, S. Chen et al., "Blood pressure classification of 2017 associated with cardiovascular disease and mortality in young Chinese adults," *Hypertension*, vol. 76, no. 1, pp. 251–258, 2020.

[4] R. D. Brook, L. J. Appel, M. Rubenfire et al., "Beyond medications and diet: alternative approaches to lowering blood pressure," *Hypertension*, vol. 61, no. 6, pp. 1360–1383, 2013.

[5] H. Wen and L. Wang, "Reducing effect of aerobic exercise on blood pressure of essential hypertensive patients: a meta-analysis," *Medicine*, vol. 96, no. 11, p. e6150, 2017.

[6] L. Cao, X. Li, P. Yan et al., "The effectiveness of aerobic exercise for hypertensive population: a systematic review and meta-analysis," *The Journal of Clinical Hypertension*, vol. 21, no. 7, pp. 868–876, 2019.

[7] M. L. Pedralli, B. Eibel, G. Waclawovsky et al., "Effects of exercise training on endothelial function in individuals with hypertension: a systematic review with meta-analysis," *Journal of the American Society of Hypertension*, vol. 12, no. 12, pp. e65–e75, 2018.

[8] I. Gorostegi-Anduaga, P. Corres, A. MartinezAguirre-Betolaza et al., "Effects of different aerobic exercise programmes with nutritional intervention in sedentary adults with overweight/obesity and hypertension: EXERDIET-HTA study," *European Journal of Preventive Cardiology*, vol. 25, no. 4, pp. 343–353, 2018.

[9] B. K. Pedersen and B. Saltin, "Exercise as medicine-evidence for prescribing exercise as therapy in 26 different chronic diseases," *Scandinavian Journal of Medicine & Science in Sports*, vol. 25, pp. 1–72, 2015.

[10] S. Lopes, J. Mesquita-Bastos, A. J. Alves, and F. Ribeiro, "Exercise as a tool for hypertension and resistant hypertension management: current insights," *Integrated Blood Pressure Control*, vol. 11, pp. 65–71, 2018.

[11] C. Hacke, D. Nunan, and B. Weisser, "Do exercise trials for hypertension adequately report interventions? A reporting quality study," *International Journal of Sports Medicine*, vol. 39, no. 12, pp. 902–908, 2018.

[12] C. Ozemek and R. Arena, "Precision in promoting physical activity and exercise with the overarching goal of moving more," *Progress in Cardiovascular Diseases*, vol. 62, no. 1, pp. 3–8, 2019.

[13] R. Ross, B. H. Goodpaster, L. G. Koch et al., "Precision exercise medicine: understanding exercise response variability," *British Journal of Sports Medicine*, vol. 53, no. 18, pp. 1141–1153, 2019.

[14] K. Albouaini, M. Egred, A. Alahmar, and D. J. Wright, "Cardiopulmonary exercise testing and its application," *Postgraduate Medical Journal*, vol. 83, no. 985, pp. 675–682, 2007.

[15] G. J. Balady, R. Arena, K. Sietsema et al., "Clinician's guide to cardiopulmonary exercise testing in adults," *Circulation*, vol. 122, no. 2, pp. 191–225, 2010.

[16] J.-C. Youn and S.-M. Kang, "Cardiopulmonary exercise test in patients with hypertension: focused on hypertensive response to exercise," *Pulse*, vol. 3, no. 2, pp. 114–117, 2015.

[17] A. Mezzani, "Cardiopulmonary exercise testing: basics of methodology and measurements," *Annals of the American Thoracic Society*, vol. 14, no. Supplement_1, pp. S3–S11, 2017.

[18] U. Drescher, J. Koschate, and U. Hoffmann, "Oxygen uptake and heart rate kinetics during dynamic upper and lower body exercise: an investigation by time-series analysis," *European Journal of Applied Physiology*, vol. 115, no. 8, pp. 1665–1672, 2015.

[19] M. Guazzi, F. Bandera, C. Ozemek, D. Systrom, and R. Arena, "Cardiopulmonary exercise testing," *Journal of the American College of Cardiology*, vol. 70, no. 13, pp. 1618–1636, 2017.

[20] R. Buys, A. Van De Bruaene, J. Müller et al., "Usefulness of cardiopulmonary exercise testing to predict the development of arterial hypertension in adult patients with repaired isolated coarctation of the aorta," *International Journal of Cardiology*, vol. 168, no. 3, pp. 2037–2041, 2013.

[21] K. Keller, K. Stelzer, M. A. Ostad, and F. Post, "Impact of exaggerated blood pressure response in normotensive individuals on future hypertension and prognosis: systematic review according to PRISMA guideline," *Advances in Medical Sciences*, vol. 62, no. 2, pp. 317–329, 2017.

[22] P. O. Older and D. Z. H. Levett, "Cardiopulmonary exercise testing and surgery," *Annals of the American Thoracic Society*, vol. 14, no. Supplement_1, pp. S74–S83, 2017.

[23] D. J. Stubbs, L. A. Grimes, and A. Ercole, "Performance of cardiopulmonary exercise testing for the prediction of postoperative complications in non-cardiopulmonary surgery: a systematic review," *PLoS One*, vol. 15, no. 2, p. e0226480, 2020.

[24] C. Santoro, R. Sorrentino, R. Esposito et al., "Cardiopulmonary exercise testing and echocardiographic exam: an useful interaction," *Cardiovascular Ultrasound*, vol. 17, no. 1, p. 29, 2019.

[25] I. Nedeljkovic, M. Banovic, J. Stepanovic et al., "The combined exercise stress echocardiography and cardiopulmonary exercise test for identification of masked heart failure with preserved ejection fraction in patients with hypertension," *European Journal of Preventive Cardiology*, vol. 23, no. 1, pp. 71–77, 2016.

[26] R. Badagliacca, S. Papa, G. Valli et al., "Echocardiography combined with cardiopulmonary exercise testing for the prediction of outcome in idiopathic pulmonary arterial hypertension," *Chest*, vol. 150, no. 6, pp. 1313–1322, 2016.

[27] E. Leopold, D. Navot-Mintzer, E. Shargal et al., "Prediction of the Wingate anaerobic mechanical power outputs from a maximal incremental cardiopulmonary exercise stress test using machine-learning approach," *PLoS One*, vol. 14, no. 3, p. e0212199, 2019.

[28] F. Braccioni, D. Bottigliengo, A. Ermolao et al., "Dyspnea, effort and muscle pain during exercise in lung transplant recipients: an analysis of their association with cardiopulmonary function parameters using machine learning," *Respiratory Research*, vol. 21, no. 1, pp. 1–11, 2020.

[29] S. Sakr, R. Elshawi, A. Ahmed et al., "Using machine learning on cardiorespiratory fitness data for predicting hypertension: the Henry Ford ExercIse Testing (FIT) Project," *PLoS One*, vol. 13, no. 4, p. e0195344, 2018.

[30] G. Yang, X. Leng, F. Huang et al., "Use CPET data to predict the intervention effect of aerobic exercise on young hypertensive patients," in *Proceedings of 2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pp. 1699–1702, IEEE, San Diego, CA, USA, November 2019.

[31] A. Bagnall, J. Lines, A. Bostrom, J. Large, and E. Keogh, "The great time series classification bake off: a review and experimental evaluation of recent algorithmic advances," *Data Mining and Knowledge Discovery*, vol. 31, no. 3, pp. 606–660, 2017.

[32] A. Bagnall, J. Lines, J. Hills, and A. Bostrom, "Time-series classification with COTE: the collective of transformation-based ensembles," *IEEE Transactions on Knowledge and Data Engineering*, vol. 27, no. 9, pp. 2522–2535, 2015.

[33] J. Lines, S. Taylor, and A. Bagnall, "Time series classification with HIVE-COTE: the hierarchical vote collective of transformation-based ensembles," *ACM Transactions on Knowledge Discovery from Data*, vol. 12, no. 5, p. 52, 2018.

[34] H. I. Fawaz, G. Forestier, J. Weber et al., "Deep learning for time series classification: a review," *Data Mining and Knowledge Discovery*, vol. 33, no. 4, pp. 917–963, 2019.

[35] B. Zhao, H. Lu, H. Chen, J. Liu, and D. Wu, "Convolutional neural networks for time series classification," *Journal of Systems Engineering and Electronics*, vol. 28, no. 1, pp. 162–169, 2017.

[36] Q. Ma, L. Shen, W. Chen, J. Wang, J. Wei, and Z. Yu, "Functional echo state network for time series classification," *Information Sciences*, vol. 373, pp. 1–20, 2016.

[37] Z. Zhang, Y. Xu, J. Yang, X. Li, and D. Zhang, "A survey of sparse representation: algorithms and applications," *IEEE Access*, vol. 3, pp. 490–530, 2015.

[38] J. A. Tropp, A. C. Gilbert, and M. J. Strauss, "Algorithms for simultaneous sparse approximation. Part I: greedy pursuit," *Signal Processing*, vol. 86, no. 3, pp. 572–588, 2006.

[39] J. A. Tropp, "Algorithms for simultaneous sparse approximation. Part II: convex relaxation," *Signal Processing*, vol. 86, no. 3, pp. 589–602, 2006.

[40] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4655–4666, 2007.

[41] B. Efron, T. Hastie, I. Johnstone et al., "Least angle regression," *The Annals of Statistics*, vol. 32, no. 2, pp. 407–499, 2004.

[42] T. T. Wu and K. Lange, "Coordinate descent algorithms for lasso penalized regression," *The Annals of Applied Statistics*, vol. 2, no. 1, pp. 224–244, 2008.

[43] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM Journal on Imaging Sciences*, vol. 2, no. 1, pp. 183–202, 2009.

[44] J. Wright, A. Y. Yang, A. Ganesh et al., "Robust face recognition via sparse representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 2, pp. 210–227, 2008.

[45] Z. Chen, W. Zuo, Q. Hu, and L. Lin, "Kernel sparse representation for time series classification," *Information Sciences*, vol. 292, pp. 15–26, 2015.

[46] L. Zhang, M. Yang, and X. Feng, "Sparse representation or collaborative representation: which helps face recognition," in *Proceedings of International Conference on Computer Vision*, pp. 471–478, IEEE, Barcelona, Spain, November 2011.

[47] W. Deng, J. Hu, and J. Guo, "Extended SRC: undersampled face recognition via intraclass variant dictionary," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 9, pp. 1864–1870, 2012.

[48] R. Rubinstein, A. M. Bruckstein, and M. Elad, "Dictionaries for sparse representation modeling," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 1045–1057, 2010.

[49] P. Wang, L. Kong, T. Du, and L. Wang, "Orthogonal sparse dictionary based on Chirp echo for ultrasound imaging," *Applied Acoustics*, vol. 156, pp. 359–366, 2019.

[50] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4311–4322, 2006.

[51] Z. Yu, X. Zheng, F. Huang et al., "A framework based on sparse representation model for time series prediction in smart city," *Frontiers of Computer Science*, vol. 15, no. 1, pp. 1–13, 2020.

[52] M. Alkhodari, D. K. Islayem, F. A. Alskafi, and A. H. Khandoker, "Predicting hypertensive patients with higher risk of developing vascular events using heart rate variability and machine learning," *IEEE Access*, vol. 8, pp. 192727–192739, 2020.

[53] H. A. Dau, A. Bagnall, K. Kamgar et al., "The UCR time series archive," *IEEE/CAA Journal of Automatica Sinica*, vol. 6, no. 6, pp. 1293–1305, 2019.

[54] H. Deng, G. Runger, E. Tuv, and M. Vladimir, "A time series forest for classification and feature extraction," *Information Sciences*, vol. 239, pp. 142–153, 2013.

[55] J. Zhao and L. Itti, "Shapedtw: shape dynamic time warping," *Pattern Recognition*, vol. 74, pp. 171–184, 2018.

Hindawi
Journal of Healthcare Engineering
Volume 2021, Article ID 5591660, 14 pages
https://doi.org/10.1155/2021/5591660

Research Article

PSR: Unified Framework of Parameter-Learning-Based MR Image Superresolution

Huanyu Liu[1,2], **Jiaqi Liu**[1,2], **Junbao Li**[1,2], **Jeng-Shyang Pan**[3] and **Xiaqiong Yu**[4]

[1]*School of Electronics and Information Engineering, Harbin Institute of Technology, Harbin 150001, China*
[2]*Center of AI Perception, AI Research Institute, Harbin Institute of Technology, Harbin 150001, China*
[3]*College of Computer Science and Engineering, Shandong University of Science and Technology, Qingdao 266590, China*
[4]*32021 Troops of the PLA, Beijing 100094, China*

Correspondence should be addressed to Junbao Li; lijunbao@hit.edu.cn

Received 25 February 2021; Revised 27 March 2021; Accepted 9 April 2021; Published 22 April 2021

Academic Editor: Hao Chun Lu

Magnetic resonance imaging has significant applications for disease diagnosis. Due to the particularity of its imaging mechanism, hardware imaging suffers from resolution and reaches its limit, and higher radiation intensity and longer radiation time will cause damage to the human body. The problem is expected to be solved by a superresolution algorithm, especially the image superresolution based on sparse reconstruction has good performance. Dictionary generation is a key issue that affects the performance of superresolution algorithms, and dictionary performance is affected by dictionary construction parameters: balance parameters, dictionary size, overlapping block size, and a number of training sample blocks. In response to this problem, we propose an optimal dictionary construction parameter search method through the experiment to find the optimal dictionary construction parameters on the MR image and compare them with the dictionary obtained by multiple sets of random dictionary construction parameters. The dictionary we searched for the optimal parameters of the dictionary construction training has more powerful feature expressions, which can improve the superresolution effect of MR images.

1. Introduction

Magnetic resonance imaging (MRI) becomes more and more widely used in medical clinical applications and plays an increasingly important role in the diagnosis of various diseases [1–5]. The mechanism of MR imaging is different from that of natural images. The hydrogen protons of human organs are magnetized under the action of external magnetic fields and generate a magnetic resonance phenomenon under the action of a magnetic field. The changing magnetic signals are converted into electrical signals by induction coils and fill the K space. Finally, an MR image is generated through the Fourier transform. The method of improving resolution relies heavily on increasing the magnetization of more free water in human tissues and organs [6], which will cause the increase of the radiation time and radiation intensity of the main magnetic field of the magnetic resonance imager and the loaded electromagnetic waves [7]; excessive radiation can lead to serious consequences, such as overheating of the human body and protein inactivation [8], causing harm to the human body and not suitable for clinical application. From the perspective of current imaging methods and technologies, the hardware imaging resolution reaches the limit value. To increase the resolution, software superresolution technology must be used to increase the image resolution.

Image superresolution methods are mainly based on interpolation, reconstruction, and learning. Li and Orchard [9] proposed a new edge-directed image interpolation (NEDI) method; Wang and Ling [10] proposed an Edge-Adaptive Interpolation Algorithm (EAIA), combined with bilinear and NEDI methods; Giachetti and Asuni [11] proposed an interpolation based on iterative curvature method based on the NEDI method. But the interpolation-based method does not essentially increase the image information. Irani and Peleg [12] proposed an iterative

backprojection method (IBP), Schultz and Stevenson [13] proposed a superresolution method based on maximum posterior probability (MAP), Patti et al. [14] proposed the convex set projection method, *and* in [15], the projection onto convex sets superresolution reconstruction method was used for the superresolution of cardiac valve MR images. POCS algorithm is not good at maintaining the image edge and can not restore the high-frequency information on the image. The superresolution method based on reconstruction regards low-resolution observation images as a constraint condition of the original high-resolution image, and a series of solution spaces satisfying the constraint condition can be obtained through the alternating iteration method. Most of the above algorithms use prior knowledge such as the edge characteristics of the image, the nonnegativity of pixels and local smoothing characteristics to construct constraints and then solve the optimization problem through an iterative algorithm. The reconstruction algorithms are computationally expensive, and the resulting image is reconstructed too smoothly. Learning-based superresolution methods mainly include dictionary learning and deep learning. Yang et al. [16] proposed a superresolution reconstruction algorithm based on sparse representation. This method effectively overcomes the problem of inaccurate representation caused by using a fixed number of neighbors. On the basis of the methods [17], adaptive sparse field selection and adaptive regularization are applied to superresolution. Yang et al. [18] proposed a double-geometric neighborhood embedding method (DGNE), which uses multiview features and local spatial neighbors of image blocks to find image features-spatial manifold embedding. Zhang et al. [19] combined subspace division and local regressor learning through the mixture of experts method to further improve the quality of image reconstruction. With the development of deep learning. Shi et al. [20] propose a novel image SR method that integrates both locals and global information for effective image recovery. This is achieved by, in addition to TV, low-rank regularization that enables utilization of information throughout the image. The reconstructed image will produce stripe distortion, and the texture and other details will be blurred to some extent. Dong et al. [21] used convolutional neural networks for image superresolution reconstruction for the first time. This method first uses bicubic interpolation to enlarge it to the target size and then passes through a three-layer convolution network doing the nonlinear mapping. The results obtained are output as high-resolution images, and experiments show that it has achieved good results. Since then, Residual Dense Network [22], SRGAN [23], and many deep networks [24, 25] are proposed and used. Deep learning methods are based on data driving, and Network performance is affected by the amount of data. However, due to the particularity and privacy of MR images, it is difficult to obtain large amounts of data. MR image superresolution task is more suitable for methods that weak dependence on data volume.

Our contributions are threefold. First, we propose a superresolution architecture based on joint dictionary learning suitable for a small number of MR images. Second, we analyze the effect of dictionary parameters on dictionary performance and find the optimal dictionary parameters through parameters learning. Third, experiments prove that our proposed method can achieve state-of-the-art performance, even if there are only a few image data.

2. Algorithm and Analysis

2.1. Algorithm. We propose a superresolution architecture based on joint dictionary learning suitable for a small number of MR images. The algorithm framework is shown in Figure 1. Through corresponding high and low-resolution image block training, learn high and low-resolution dictionary, low-resolution image blocks are sparsely represented by a low-resolution dictionary, and the sparse coefficients can be used for high-resolution image reconstruction. The performance of high and low-resolution dictionaries directly affects the image reconstruction effect.

Training a joint dictionary requires the use of high-resolution image block sets and low-resolution image block sets. Training image pair is represented by $P = \{Y^l, X^h\}$. $X^h = \{x_1, x_2, x_3, \ldots, x_n\}$ and $Y^l = \{y_1, y_2, y_3, \ldots, y_n\}$ represents features extracted from image blocks. The dictionary training of two image feature spaces is expressed as follows [16]:

$$D_h = \arg\min_{\{D_h, z\}} \left\| X^h - D_h Z \right\|_2^2 + \lambda \|Z\|_1, \tag{1}$$

$$D_l = \arg\min_{\{D_l, z\}} \left\| Y^l - D_l Z \right\|_2^2 + \lambda \|Z\|_1. \tag{2}$$

According to the idea of the joint training method, the image blocks corresponding to the two image feature block spaces are concatenated to form a new image feature block space, so formula (3) [16] can be obtained:

$$\min_{\{D_h, D_l, Z\}} \frac{1}{N} \left\| X^h - D_h Z \right\|_2^2 + \frac{1}{M} \left\| Y^l - D_l Z \right\|_2^2 + \lambda \left(\frac{1}{N} + \frac{1}{M} \right) \times \|Z\|_1, \tag{3}$$

M is the dimensions of the low-dimensional image feature block, and N is the dimensions of the high-dimensional image feature block. It can be seen from formula (3) that balance parameters, dictionary block size, overlapping block size, and the number of dictionary blocks have an important impact on the performance of the dictionary. We obtain the optimal parameters of dictionary construction through experimental analysis so as to achieve the improvement of dictionary performance and image reconstruction effect.

2.2. Parameter Set and Analysis. The parameter set is written as parameter = [λ, overlap, n, spn], where λ is balance parameter, the overlap is the size of the overlap block, n is the size of the dictionary block, and spn is the number of the exemplar patch.

From the mechanism of reconstruction perspective, changes of the parameter can cause changes in the structure and quantity of the data calculated by the dictionary, which has a great influence on the effect of reconstruction. The specific analysis is as follows.

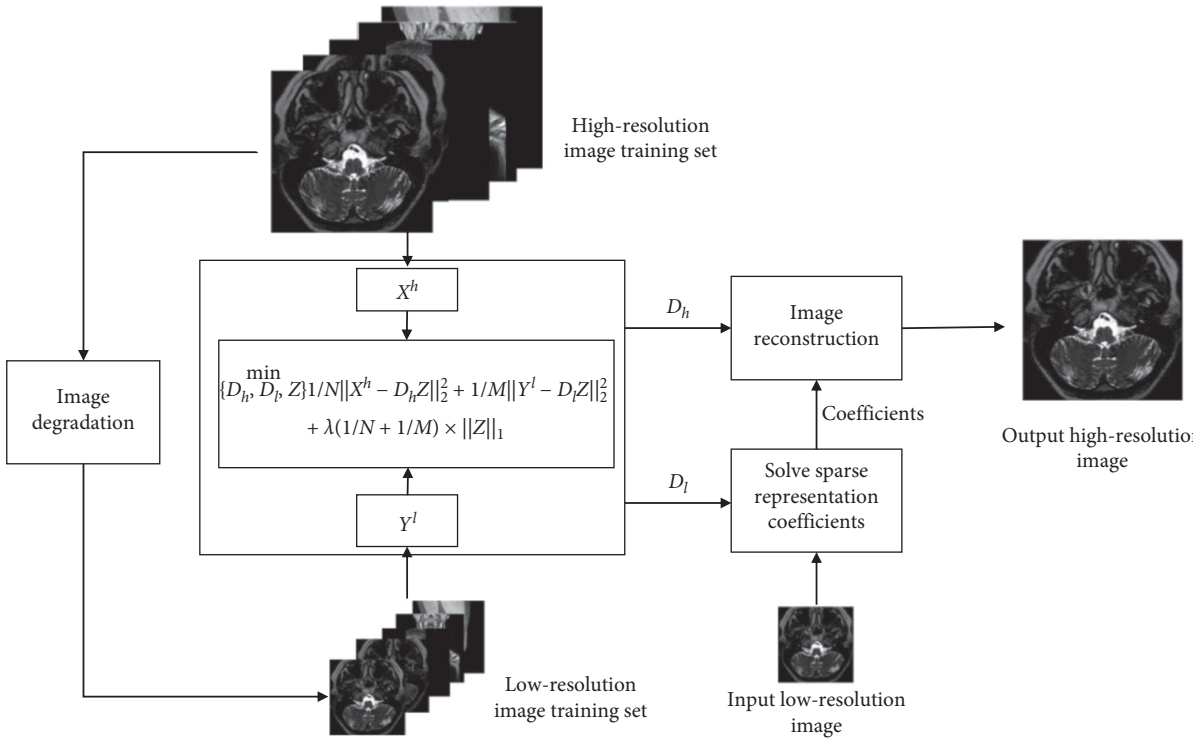


Figure 1: MR image superresolution architecture based on joint dictionary learning.

According to the formula in equation (3), the balance parameter λ is used to balance sparsity and low-resolution dictionary sparse, which represents image block errors. It can be analyzed from the equation that the sparsity is inversely proportional to the error. In order to find the minimum sparsity coefficient, the error will increase, and vice versa, so there must be an optimal value λ to minimize the value of equation (3).

The overlapped block overlap is the size of the overlapped portion among the image blocks, and is divided into the selection of sample blocks in dictionary training and overlap of test image blocks in reconstruction. In order to ensure that the detailed features of the sample block can be extracted in the dictionary training, the maximum overlapped block method is used to select samples, which makes 1 pixel gradually change between the training sample blocks of a localized image. In order to eliminate the boundary blur caused by the feature extraction of the test image block and the reliability of the connection between the reconstructed blocks, the overlap between the blocks is required. The larger the overlap block, the larger the constraint between the reconstructed blocks, and the better the reconstruction effect.

MR images are expected to show various tissue structures clearly, with high tissue resolution. However, the outline of the diseased tissue can not be seen clearly and separated from the surrounding structure, so it is important to clearly display the texture details, especially the boundaries between different organizational structures, and to some extent, reduce the chances of misjudgment of the diagnosis due to the blurring of the picture. Then, for the dictionary and the reconstructed image, the size of the dictionary block used to represent the feature will affect the number of effective features. The smaller the dictionary block, the fewer features are generated, which makes the reconstructed image have limitations and larger errors, such as the most extreme 1×1 and 2×2 blocks. The features they can describe are limited. However, the oversized block is also problematic. The image is too large, and the features described by the image block can be combined and described by several smaller feature blocks, then they lose the properties of the feature block, such as extreme cases; the test image itself is a large feature block; of course, this is unreasonable. Therefore, there must be an optimal value for image segmentation, which makes image reconstruction better.

We adopt the method of sparse representation and reconstruction of image blocks. When training the dictionary, a large number of sampling image blocks are required. The number of image sampling blocks has a certain influence on the reconstructed quality. If the image block sample is too little, it is not enough to complete the training of the dictionary. If there are too many image sample blocks, especially the features of some image sample blocks that are not obvious or typical, the characterization of the dictionary cannot be improved even if there is much training. Is there an optimal number of partitions? Since the selected image sampling blocks are randomly extracted, it is difficult to extract the required sample blocks in an accurate number of blocks, so there is no optimal number of sample blocks. Therefore, the selection of the sample block is as long as a certain amount. Too little or too much can both not change the image reconstruction effect.

3. Experimental Results and Analysis

3.1. Dataset. The experiments are all set as follows: The method adopts the experimental framework of Section 1 of Chapter 2. 81; representative pictures of different categories in the image library were selected as the training samples of the high-resolution dictionary. These MR images were obtained from Siemens 3T platforms using a 32-channel head coil. Low-resolution images are generated by the degradation of high-resolution images. LR images are generated following the steps: (1) the high-resolution images are transformed from image space to K space by FFT, (2) in the K space, the outer high frequency is truncated, (3) through the inverse Fourier transform, the truncated k space data are transformed into the image space to generate the corresponding low-resolution images. This mimics the actual acquisition of LR and HR images by MRI scanners. In the experiment, as shown in Figure 2, five images corresponding to different types of MR images are selected as test samples.

3.2. Joint Optimization of Parameter

3.2.1. Balance Parameter λ. The optimal value of the balance parameter is verified by the experiments below. The initial parameter configures are as follows: the dictionary size is 512, the balance parameter $\lambda = 0.1$, and the block size is 5×5, the overlap block is 4, the number of sample blocks is 100000, and the test samples are, respectively, selected from the head, ankle, carotid artery, knee, and neck, as shown in Figure 2.

It can be seen from Figure 4 that the value of PSNR decreases significantly with the increasing λ when $\lambda > 0.1$. On the contrary, when $\lambda < 0.1$, the value of PSNR decreases slowly with λ decreasing. As the balance parameter of the sparsity, λ exists optimal value, which makes PSNR maximum. For further verification, let $\lambda = 0.1$ as the optimal balance parameter. The super resolution ratio is 1 : 4.

The experiment used a superresolution ratio of 1 : 4, and other experimental parameters are the same as those in experiment ratio 1 : 2. The experimental results are shown in Table 2.

Figure 3 is obtained from Table 2. It can be seen from the figure that the extreme point is near $\lambda = 0.1$, and the experiment with a superresolution ratio of 1 : 4 has the same conclusion as the experiment with a superresolution ratio of 1 : 2.

3.2.2. Overlap Block. The relationship between the image reconstruction effect and overlapped blocks is verified by the following experiment. The initial parameters in the experiment are the same as those in Experiment Balance parameter, and the changed parameters are the size of the overlapped region. The overlapped blocks 1–4 are used to represent the superposed pixels. The experimental results can be seen in Table 3.

According to the experimental results, the images with superresolution ratios of 1 : 2 and 1 : 4 are demonstrated, respectively, in Figure 5, where the abscissa represents the number of overlapped blocks, and the ordinate is the corresponding PSNR. It can be seen from the figure that as the overlay area of the overlapped blocks decreases, the value of PSNR decreases. This is because the larger the overlay area where the image blocks selected for the reconstructed block, the larger the constraint between the reconstructed blocks, and this is easy to find the closest image block to be connected. The more the pixel points at the edge of the image block overlap, the easier it is to eliminate the truncation error caused by the feature extraction. It also has a certain inhibitory effect on noise.

3.2.3. Dictionary Blocking Size. The following experiments show the quality of the reconstructed image when having the different blocking conditions for the same test image, where the set of dictionaries is generated with the altering size of the blocked image.

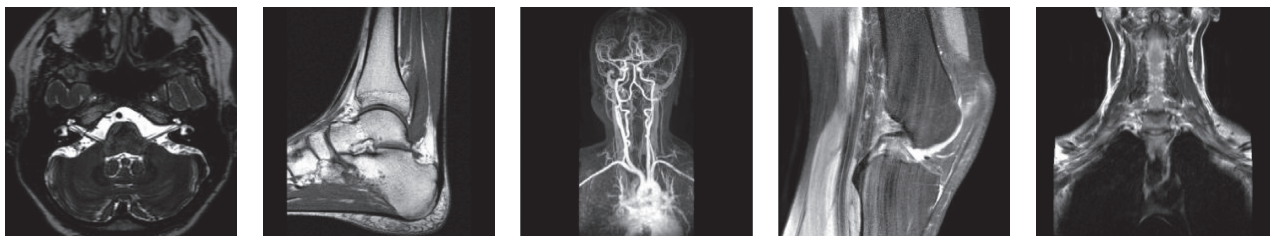
Other experimental parameters do not change, and the changed parameters are the block size of the image blocking. The image blocking of the dictionary has the same requirements as the image blocking of the test image.

The experimental results are as follows: when the superresolution ratio is 1 : 2, 8 high-definition dictionaries with the image block from 3×3 to 10×10 are generated. Three image blocks are shown in Figure 6. It can be seen that as the image block size increases, the dictionary block becomes more and more complicated. The resulting PSNR values are shown in Table 4.

Table 4 shows the value of the superresolution reconstruction PSNR corresponding to the different block training dictionaries of images. Since the reconstructed overlapped block is 4, which exceeds the block size of the sample block itself, the reconstructed samples are not correct when the dictionaries are, respectively, corresponding to block 3×3 and block 4×4. The two data sets are not analyzed. The other data corresponding to dictionary block and PSNR are shown in Figure 7.

The abscissa in Figure 7 only represents the block of the dictionary and the image. For example, the abscissa 5 indicates that the dictionary block is 5×5 and so on. Increasing the block size will reduce the value of PSNR when the overlap block size is unchanged. That is to say, the block is not bigger always better. When the block is large, the number of dictionary blocks that represent the image feature block will increase, and the reconstruction error will become larger. It can be seen that the preferred block value is 5×5 or 6×6 blocks, and the calculation efficiency 5×5 blocks is optimal.

When the superresolution ratio is 1 : 4, high-resolution dictionaries from 5×5 to 13×13 are generated through experiment. Three high-resolution dictionaries are shown in Figure 8, where the dictionary block becomes more and more complicated with the number of the blocks increasing. But the too-large block causes too many singular matrices when calculating the dictionary block, which causes the dictionary block information to be lost. The larger the block, the fewer the valid dictionary blocks. This will lead to a decrease in PSNR values, as shown in Figure 8(c). The



(a) (b) (c) (d) (e)

Figure 2: Image test sample. The comparison of reconstructed result with different balance parameter λ is shown in Table 1, where the superresolution is 1 : 2. The first column in Table 1 is the change of the balance parameter. The data in the table is the corresponding PSNR value. The results are plotted in Figure 3. (a) Head, (b) ankle, (c) carotid artery, (d) knee, and (e) neck.

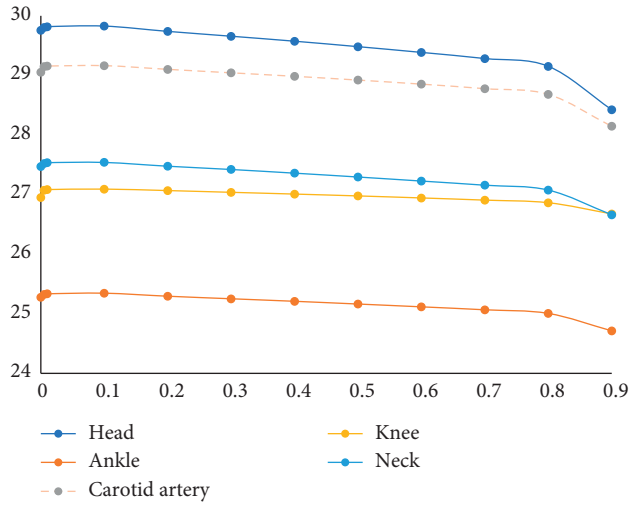


Figure 3: PSNR curve with the balance parameter λ variation (superresolution ratio 1 : 4).

Table 1: The PSNR values of the five groups of samples with parameters λ variation (superresolution ratio 1 : 2).

Balance parameter λ	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5
0.0001	32.1147	28.7875	32.0252	30.4382	30.1109
0.005	32.2643	28.9049	32.0483	30.6418	30.3621
0.01	32.2774	28.9341	32.0550	30.6463	30.3665
0.1	32.2431	28.8725	32.0357	30.5912	30.3487
0.2	32.1026	28.7828	31.9349	30.5253	30.2482
0.3	31.9502	28.6892	31.8094	30.4486	30.1392
0.4	31.7931	28.5935	31.6755	30.3667	30.0240
0.5	31.6271	28.4894	31.5267	30.2823	29.9142
0.6	31.4527	28.3712	31.3588	30.1950	29.7959
0.7	31.1624	28.1633	31.0738	30.0590	29.6033
0.8	29.1705	26.9047	29.2414	29.2739	28.1618
0.9	24.4864	23.1957	25.2566	26.5782	24.2847

parameters in the experiment only change is the superresolution ratio of 1 : 4, and the experimental results are shown in Table 5.

The data in Table 5 are the corresponding PSNR values generated by the superresolution reconstruction of the test sample with different block training dictionaries for the corresponding image. In order to intuitively distinguish the influence of the block on the reconstruction, the horizontal coordinate is the image block and the ordinate is the PSNR, as shown in Figure 9.

The abscissa in Figure 9 only shows the image blocking situation. The preferred PSNR corresponds to a 10×10 or 11×11 image blocking. Taking into account the calculation amount, 10×10 image blocking is the best. If the image

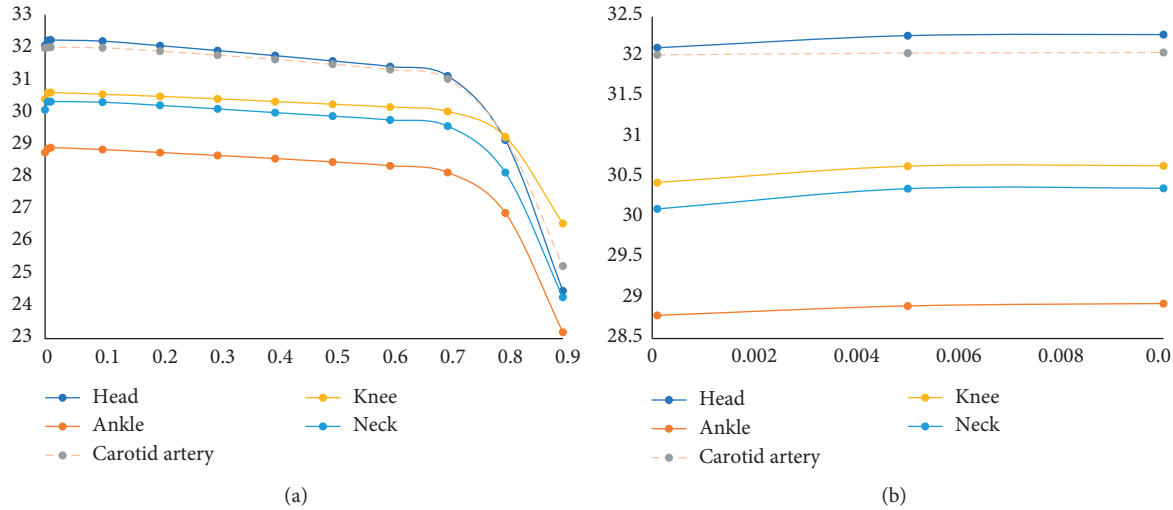


Figure 4: PSNR curve with the balance parameter λ variation (superresolution ratio 1 : 2).

Table 2: The PSNR values of the five groups of samples with parameters λ variation (superresolution ratio 1 : 4).

Balance parameter λ	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5
0.0001	29.7593	25.2764	29.0537	26.9532	27.4701
0.005	29.8049	25.3250	29.1497	27.0742	27.5213
0.01	29.8186	25.3349	29.1571	27.0843	27.5372
0.1	29.8281	25.3454	29.1645	27.0922	27.5412
0.2	29.7391	25.2944	29.1024	27.0675	27.4779
0.3	29.6576	25.2515	29.0441	27.0382	27.4229
0.4	29.5741	25.2092	28.9856	27.0088	27.3612
0.5	29.4826	25.1631	28.9223	26.9778	27.2964
0.6	29.3861	25.1159	28.8543	26.9439	27.2285
0.7	29.2832	25.0664	28.7783	26.9074	27.1586
0.8	29.1534	25.0057	28.6830	26.8639	27.0771
0.9	28.4246	24.7130	28.1476	26.6769	26.6622

Table 3: The PSNR data table varies with overlap blocks.

Superresolution ratio	Overlap block	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5
1 : 2	1	31.8015	28.5365	31.6296	30.3486	29.9244
1 : 2	2	31.8795	28.6155	31.7500	30.4587	30.0547
1 : 2	3	32.1207	28.8228	31.9781	30.5427	30.2708
1 : 2	4	32.2431	28.8725	32.0357	30.5912	30.3487
1 : 4	1	28.6725	24.7516	28.3602	26.6512	26.8626
1 : 4	2	29.0840	25.0284	28.6332	26.8210	27.0542
1 : 4	3	29.5786	25.2175	28.9978	27.0029	27.3708
1 : 4	4	29.8281	25.3454	29.1645	27.0922	27.5412

blocking is too small, it cannot represent features fully, and if the image blocking is too large, the algorithm itself has limitations.

Comparing results corresponding to the superresolution ratios of 1 : 4 and 1 : 2, they have their own best partitions. The image blocks with a superresolution ratio of 1 : 4 are approximately double that of 1 : 2. This is because the image local information required for 4 times superresolution becomes larger, and naturally, the image block needs to be correspondingly larger.

The above two experiments compared the results where the overlapped blocks are fixed as 4. But in the overlapped block experiment, the larger the overlapped blocks, the better the results. The experiment did not consider the best case of overlapped blocks. Next, we will consider that if the best block changes when blocking the different overlapped blocks corresponding maximum.

The experiment verified the effect of the maximum overlap block experiment on the superresolution performance. The parameters are the same as those before the experiments. The

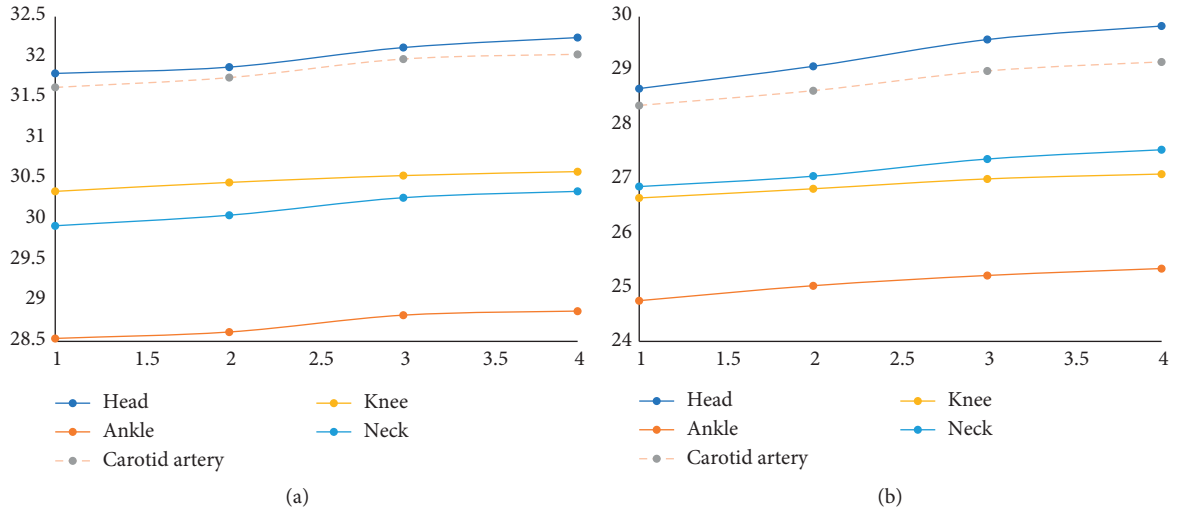


Figure 5: PSNR with overlap blocks changes (superresolution ratio 1 : 2 and 1 : 4). (a) Superresolution ratio 1 : 2, (b) superresolution ratio 1 : 4.

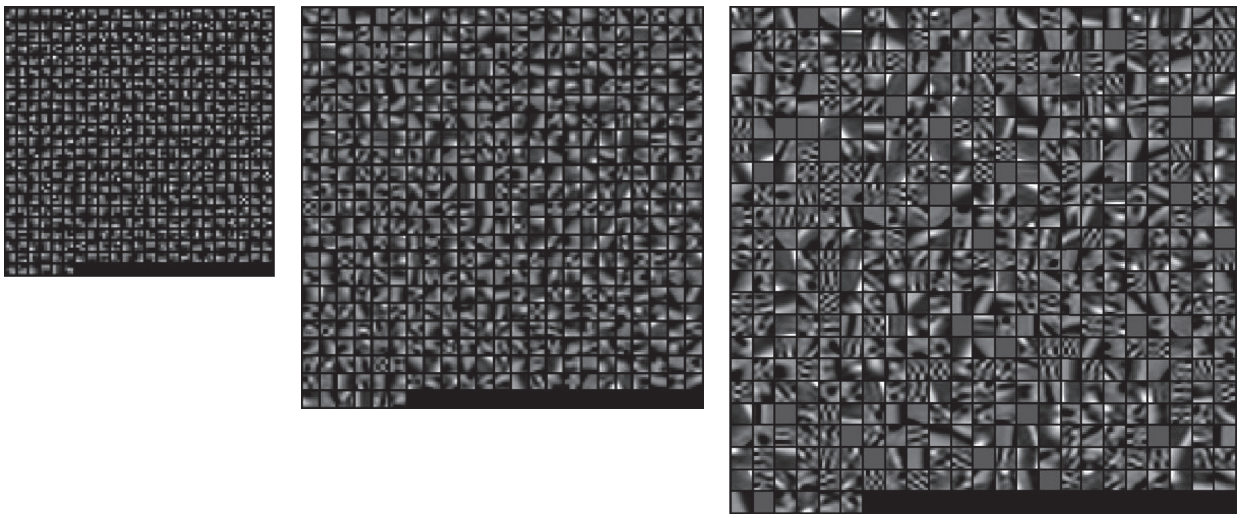


Figure 6: High-resolution dictionary blocks (superresolution 1 : 2). (a) Dictionary block 3×3. (b) Dictionary block $6 \times c$. (c) Dictionary block 10×10.

Table 4: PSNR data using different block constructed dictionaries (superresolution 1 : 2).

Dictionary blocking size	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5
5×5	32.24	28.87	32.04	30.59	30.35
6×6	32.30	28.87	32.06	30.53	30.31
7×7	31.99	28.75	31.95	30.32	30.14
8×8	31.62	28.44	31.65	30.09	29.89
9×9	31.20	28.21	31.33	29.80	29.52
10×10	30.92	27.95	31.18	29.60	29.27

changed parameters are only the block size and the overlapped block. For example, the block size is $n \times n$, and the overlapped block value is $n - 1$. When the superresolution ratio is 1 : 2, the experimental results can be seen in Table 6.

The data show the superresolution reconstruction, where each test sample corresponds to different blocks and overlap blocks. For comparison, the data are plotted as shown in Figure 10.

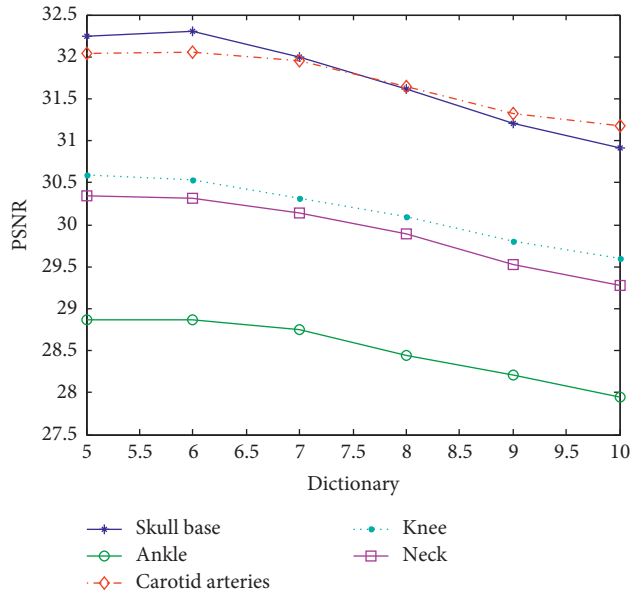
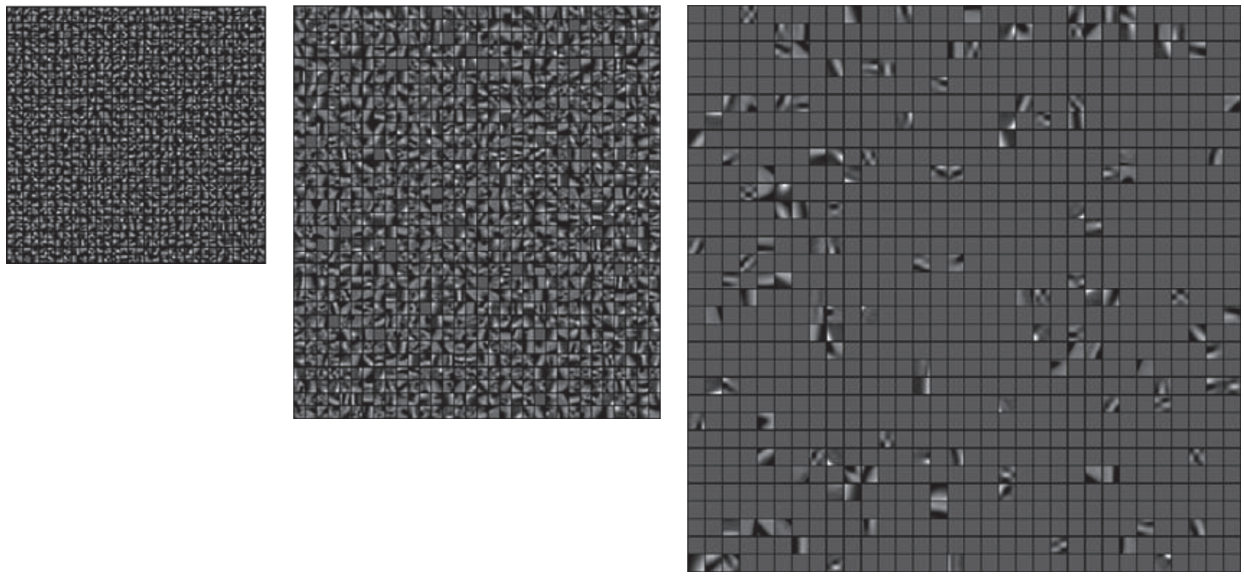


Figure 7: Superresolution PSNR values of dictionary images constructed with different blocks (superresolution ratio 1 : 2).



(a) (b) (c)

Figure 8: High-resolution dictionary blocks (superresolution 1 : 4). (a) Dictionary block 5×5. (b) Dictionary block 9×9. (c) Dictionary block 13×13.

The horizontal coordinate in Figure 11 only indicates the difference of the block. It can be clearly seen that the reconstructed effect is better from Figure 11 when the dictionary block is 5×5 or 6×6 blocks. The smaller blocks make the features of the blocks insufficient, and the too-large blocks need to increase the number of calculated pixels. The increase in the size of the dictionary representation block caused by the increase of the feature block makes the error larger and impacts the PSNR effect. In this experiment, the largest overlap block is used to make each component block reach the best reconstruction. It can be seen that the better block value is still 5×5 or 6×6, and it is best to select a 5×5 block for the calculation efficiency.

The above experiment verified the effect of the maximum overlap block experiment on the superresolution performance. The parameters are the same as those in other experiments. The changed parameters are only the block size and the overlapped block. When the superresolution ratio is 1 : 4, the experimental results are shown in Table 7. The obtained data is still plotted with the block size as the abscissa and PSNR values as the ordinate, as shown in Figure 11.

Table 5: Effects of dictionary blocks on PSNR (superresolution ratio 1 : 4).

Dictionary blocking size	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5
5 × 5	29.83	25.35	29.16	27.09	27.54
6 × 6	30.00	25.49	29.38	27.22	27.72
7 × 7	30.30	25.61	29.49	27.18	27.91
8 × 8	30.40	25.71	29.72	27.23	28.02
9 × 9	30.43	25.69	29.78	27.27	28.05
10 × 10	30.75	25.75	29.94	27.32	28.11
11 × 11	30.67	25.77	29.90	27.22	28.12
12 × 12	30.53	25.71	29.82	27.19	27.93
13 × 13	30.31	25.60	29.64	27.08	27.78

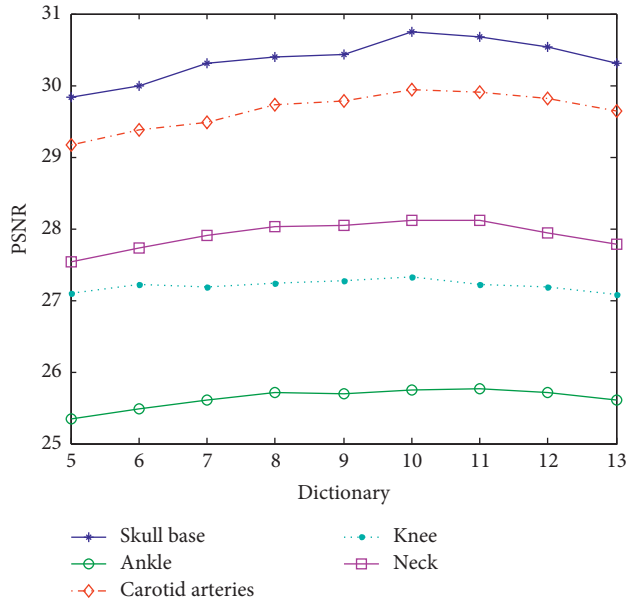


Figure 9: Superresolution PSNR values of dictionary images constructed with different blocks (superresolution ratio 1 : 4).

Table 6: Comparison of superresolution PSNR for different block and overlap block training dictionaries (superresolution ratio 1 : 2).

Block size	Overlap block	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5
3 × 3	2	31.30	28.21	31.22	30.26	29.70
4 × 4	3	31.96	28.70	31.73	30.54	30.19
5 × 5	4	32.24	28.87	32.04	30.59	30.35
6 × 6	5	32.29	28.90	32.07	30.55	30.32
7 × 7	6	32.11	28.83	32.03	30.40	30.26
8 × 8	7	31.90	28.71	31.93	30.26	30.08
9 × 9	8	31.71	28.60	31.80	30.10	29.93
10 × 10	9	31.43	28.42	31.65	29.95	29.70

It can be seen from Figure 11 that the 10 × 10 training dictionary has the best superresolution reconstruction when the superresolution ratio is 1 : 4. The above experiment shows that the block size has the highest value and is related to the superresolution ratio. The larger the superresolution ratio is, the larger the block is needed. The change of the overlap block does not influence the result of the optimal block. For the MR image, the optimal block with a superresolution ratio of 1 : 2 is 5 × 5, and the optimal block with a superresolution ratio of 1 : 4 is 10 × 10.

3.2.4. Number of Sampling Blocks. The experiment uses the same parameters as other experiments. The changed parameters are the sample amount of sample image blocks, and the data can be overlap extraction. The superresolution ratio

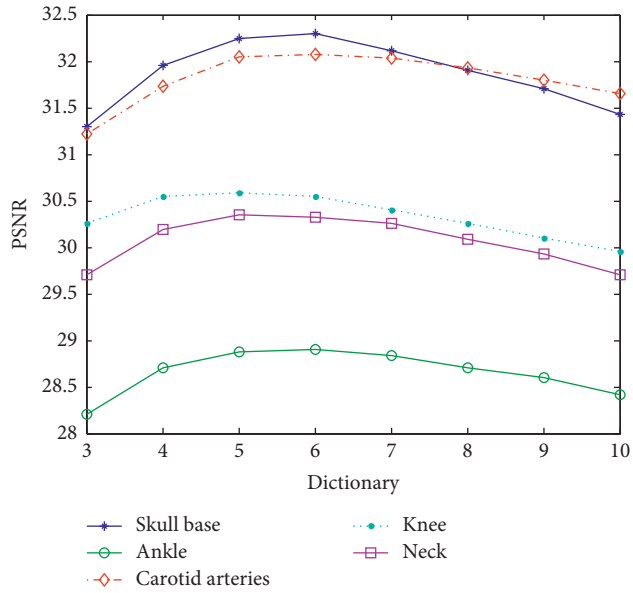


Figure 10: Comparison of superresolution PSNR for different blocks of a dictionary (superresolution ratio 1 : 2).

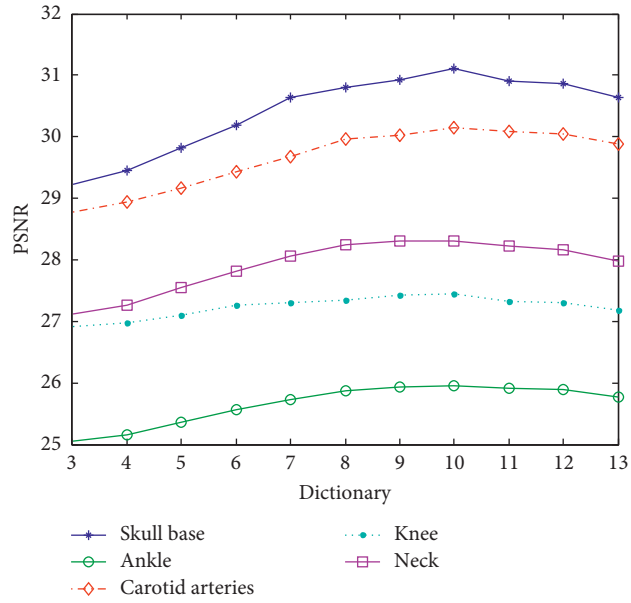


Figure 11: Comparison of superresolution PSNR for different blocks of a dictionary (superresolution ratio 1 : 4).

Table 7: Comparison of superresolution PSNR for different block and overlap block training dictionaries (superresolution ratio 1 : 4).

Block size	Overlap block	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5
3×3	2	29.22	25.05	28.78	26.91	27.11
4×4	3	29.45	25.16	28.93	26.97	27.26
5×5	4	29.83	25.35	29.16	27.09	27.54
6×6	5	30.19	25.57	29.43	27.26	27.81
7×7	6	30.63	25.72	29.68	27.30	28.06
8×8	7	30.80	25.88	29.96	27.35	28.25
9×9	8	30.92	25.94	30.03	27.43	28.30
10×10	9	31.11	25.96	30.15	27.44	28.31
11×11	10	30.91	25.91	30.09	27.32	28.22
12×12	11	30.87	25.90	30.05	27.30	28.17
13×13	12	30.63	25.76	29.88	27.19	27.98

Table 8: PSNR value corresponding to the number of sample blocks of different sampling images (superresolution ratio 1 : 2).

Number of blocks	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5
1000	31.98	28.64	31.74	30.45	30.10
5000	32.15	28.79	31.94	30.52	30.23
10000	32.20	28.81	31.96	30.56	30.28
50000	32.29	28.84	31.99	30.57	30.31
100000	32.24	28.87	32.04	30.59	30.35
150000	32.31	28.91	32.04	30.60	30.37
200000	32.26	28.84	32.04	30.58	30.35
500000	32.32	28.86	32.01	30.61	30.35

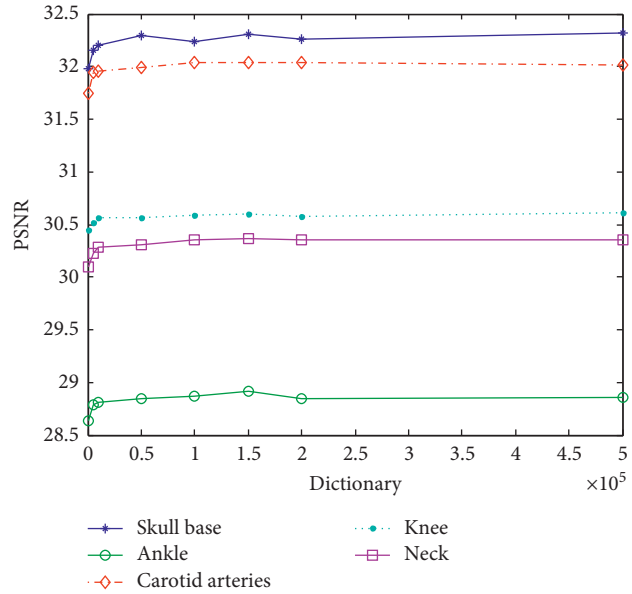


Figure 12: PSNR value corresponding to the training dictionary using different the number of samples (superresolution ratio 1 : 2).

Table 9: PSNR value corresponding to the number of sample blocks of different sampling images (superresolution ratio 1 : 4).

Number of sample image blocks	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5
1000	—	—	—	—	—
5000	29.78	25.29	29.11	27.04	27.43
10000	29.75	25.30	29.15	27.09	27.48
50000	29.77	25.35	29.16	27.11	27.53
100000	29.83	25.35	29.16	27.09	27.54
150000	29.84	25.38	29.18	27.08	27.54
200000	29.81	25.36	29.17	27.09	27.51
500000	29.85	25.39	29.16	27.10	27.53

is 1 : 2, and the experimental results are shown in Table 8. The data is taken as an abscissa in the image block sampling with different numbers of training dictionaries, and the image is plotted with PSNR values as the ordinate, as shown in Figure 12.

It can be seen from Figure 12 that the number of sample blocks below 10000 blocks is too small. Since the sample blocks that do not meet the requirements are removed in the algorithm, the MR images have many black or dark areas, and these gray scales are not changed much. Samples with little change in gray, all zeros, or near all zeros are rejected, which greatly reduces the number of blocks actually involved in the calculation. Therefore, as the training sample block, a training sample block with insufficient features reduces the value of PSNR when reconstructed. On the contrary, the large increase in the number of blocks does not cause a significant change in the PSNR, nor does it have a maximum value, showing a fluctuating change. All the training sample blocks participate in the training of the dictionary. Too many blocks will increase the training time, and there is no positive

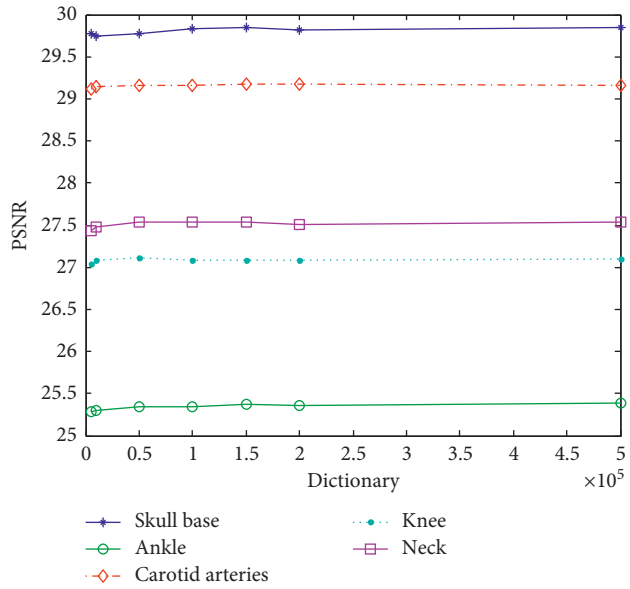


Figure 13: PSNR value corresponding to the training dictionary using different the number of samples (superresolution ratio 1 : 4).

Table 10: MR image superresolution optimal parameters.

Super resolution ratio	Balance parameter λ	Overlap block	Dictionary block size	Number of sampling blocks
1 : 2	0.1	4	5×5	150000
1 : 4	0.1	9	10×10	150000

Table 11: List of random group parameters.

Random group	Superresolution ratio	Balance parameter λ	Overlap block	Dictionary block size	Number of sampling blocks
First group	1 : 2	0.1	2	4×4	100000
Second group	1 : 2	0.2	7	8×8	50000
Third group	1 : 2	0.2	4	6×6	10000
Forth group	1 : 2	0.3	3	7×7	120000
Fifth group	1 : 4	0.4	6	9×9	30000
Sixth group	1 : 4	0.2	4	5×5	80000
Seventh group	1 : 4	0.3	9	12×12	90000
Eighth group	1 : 4	0.1	8	10×10	7000

significance for the generation of the HD dictionary. Therefore, it is better to select 150000 sampling blocks. The following experiment with a superresolution 1 : 4 is verified.

The data in Table 9 are taken as an abscissa in the image block sampling with different numbers of training dictionaries, and the image is plotted with PSNR values as the ordinate, as shown in Figure 13. As can be seen from Figure 13, the conclusion with the superresolution ratio of 1 : 4 and is the same as that with the superresolution ratio of 1 : 2 in block selection, while the dictionary cannot be trained with a superresolution ratio of 1 : 4 when the number of blocks is 1000. There are more requirements on the number of dictionaries. Considering the reduction of dictionary training time, it is better to select 150,000 blocks.

3.3. Experiment Simulation of Comprehensive Parameters. The previous section analyzes several parameters that affect the superresolution effect. The values of the optimal parameters of the superresolution MR image are shown in Table 10.

The validity of the optimal parameters is verified by the experiments below. The parameters select several sets of random parameters to form a random group training dictionary, which is compared with the dictionary of optimal parameter training, as shown in Table 11. The PSNR results obtained by experiments are shown in Table 12 below.

Comparing the data in Tables 12 and 13, superresolution PSNR data in the optimal group are higher than that in the random group, no matter the superresolution ratio is 1 : 2 or

Table 12: PSNR data of superresolution reconstruction in random group.

Random group	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5
First group	31.83	28.56	31.60	30.46	30.07
Second group	31.50	28.45	31.61	30.10	29.72
Third group	31.88	28.63	31.79	30.35	30.13
Forth group	30.78	27.85	30.87	29.74	29.17
Fifth group	30.23	25.53	29.61	27.17	27.83
Sixth group	29.80	25.34	29.14	27.07	27.50
Seventh group	30.27	25.63	29.64	27.04	27.81
Eighth group	30.66	25.81	29.85	27.28	28.06

Table 13: Superresolution reconstruction PSNR data in optimal group.

Optimal group	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5
Super resolution ratio 1 : 2	32.31	28.91	32.04	30.60	30.37
Super resolution ratio 1 : 4	31.10	26.00	30.14	27.37	28.36

1 : 4. This shows that the parameters of the optimal group are the best parameter values.

From the experimental results can be seen, the five human body parts of the superresolution effect have obvious differences. The head and carotid artery superresolution effect is best, and ankle superresolution effect is the worst. This is mainly because each part contains different water components. More water components can produce more hydrogen protons. Under the action of magnetic field and radio frequency pulse, high-frequency information will be generated, which can better generate image edge, texture, and other details.

4. Conclusion

We propose a joint dictionary learning framework for superresolution of MR images, in which changes in dictionary construction parameters will cause changes in the training dictionary and thus affect the performance of superresolution reconstructed images. We have learned the optimal dictionary construction parameters through a large number of experiments and verified that the automatically learned dictionary construction parameters could effectively improve the performance of the dictionary and enhance the expression ability of the image blocks, thereby achieving better MR image superresolution effects.

Data Availability

We have not used specific data from other sources for the simulation of the results. The two popular MRI datasets in this paper, fast MRI Dataset and IXI Dataset, can be freely downloaded from the website https://fastmri.org/and http://www.brain-development.org/.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors would like to thank Dr. Xie Haozhe and Dr. Wang Tingting for their assistance in the writing of this paper. This work was supported by National Science Foundation of China under Grant Nos. 61671170 and 61872085, Science and Technology Foundation of National Defense Key Laboratory of Science and Technology on Parallel and Distributed Processing Laboratory (PDL) under Grant no. 6142110180406, Science and Technology Foundation of ATR National Defense Key Laboratory under Grant no. 6142503180402, China Academy of Space Technology (CAST) Innovation Fund under Grant No. 2018CAST33, Joint Fund of China Electronics Technology Group Corporation, and Equipment Pre-Research under Grant No. 6141B08231109.

References

[1] J. A. López, F. Saez, J. A. Larena et al., "MRI diagnosis and follow-up of subcutaneous fat necrosis," *Journal of Magnetic Resonance Imaging*, vol. 7, no. 5, pp. 929–932, 2010.

[2] E.-S. A. El-Dahshan, H. M. Mohsen, K. Revett, and A.-B. M. Salem, "Computer-aided diagnosis of human brain tumor through MRI: a survey and a new algorithm," *Expert Systems with Applications*, vol. 41, no. 11, pp. 5526–5545, 2014.

[3] A. M. Venkatesan, R. J. Stafford, C. Duran et al., "Prostate MRI for brachytherapists: diagnosis, imaging pitfalls, and post-therapy assessment," *Brachytherapy*, vol. 16, no. 4, 2017.

[4] C. Lukas, C. Cyteval, M. Dougados et al., "MRI for diagnosis of axial spondyloarthritis: major advance with critical limitations "Not everything that glisters is gold (standard)," *RMD Open*, vol. 4, no. 1, Article ID e000586, 2018.

[5] F. Bruno, F. Arrigoni, P. Palumbo et al., "New advances in MRI diagnosis of degenerative osteoarthropathy of the peripheral joints," *La Radiologia Medica*, vol. 124, no. 11, pp. 1121–1127, 2019.

[6] M. O. Leach, "Principle of magnetic resonance," in *Physics for Medical Imaging Applications*, Springer, Dordrecht, Netherlands, 2007.

[7] R. Krug, C. Stehling, D. A. C. Kelley, S. Majumdar, and T. M. Link, "Imaging of the musculoskeletal system in vivo using ultra-high field magnetic resonance at 7 T," *Investigative Radiology*, vol. 44, no. 9, pp. 613–618, 2009.

[8] J. Karpowicz, K. Gryz, P. Politański et al., "Exposure to static magnetic field and health hazards during the operation of

magnetic resonance scanners," *Medycyna Pracy*, vol. 62, no. 3, pp. 309–321, 2011.

[9] X. Li and M. T. Orchard, "New edge-directed interpolation," *IEEE Transactions on Image Processing*, vol. 10, no. 10, pp. 1521–1527, 2001.

[10] X. F. Wang and H. F. Ling, "An edge-adaptive interpolation algorithm for superresolution reconstruction," in *Proceedings of the International Conference on Multimedia Information Networking & Security. IEEE Computer Society*, Nanjing, China, November 2010.

[11] A. Giachetti and N. Asuni, "Real-time artifact-free image upscaling," *IEEE Transactions on Image Processing*, vol. 20, no. 10, pp. 2760–2768, 2011.

[12] M. Irani and S. Peleg, "Improving resolution by image registration," *CVGIP: Graphical Models and Image Processing*, vol. 53, no. 3, pp. 231–239, 1991.

[13] R. R. Schultz and R. L. Stevenson, "Extraction of high-resolution frames from video sequences," *IEEE Transactions on Image Processing*, vol. 5, no. 6, pp. 996–1011, 1996.

[14] A. J. Patti, M. I. Sezan, and A. Murat Tekalp, "Superresolution video reconstruction with arbitrary sampling lattices and nonzero aperture time," *IEEE Transactions on Image Processing*, vol. 6, no. 8, pp. 1064–1076, 1997.

[15] A. W. Dowsey, J. Keegan, M. Lerotic, S. Thom, D. Firmin, and G.-Z. Yang, "Motion-compensated MR valve imaging with COMB tag tracking and superresolution enhancement," *Medical Image Analysis*, vol. 11, no. 5, pp. 478–491, 2007.

[16] J. Yang, J. Wright, T. S. Huang et al., "Image superresolution via sparse representation," *IEEE Transactions on Image Processing*, vol. 19, no. 11, pp. 2861–2873, 2010.

[17] W. Dong, L. Zhang, G. Shi et al., "Image deblurring and superresolution by adaptive sparse domain selection and adaptive regularization," *IEEE Transactions on Image Processing*, vol. 20, no. 7, pp. 1838–1857, 2010.

[18] S. Yang, Z. Wang, L. Zhang, and M. Wang, "Dual-geometric neighbor embedding for image super resolution with sparse tensor," *IEEE Transactions on Image Processing*, vol. 23, no. 7, pp. 2793–2803, 2014.

[19] K. Zhang, D. Tao, X. Gao et al., "Learning multiple linear mappings for efficient single image superresolution," *IEEE Transactions on Image Processing*, vol. 24, no. 3, pp. 846–861, 2015.

[20] F. Shi, J. Cheng, L. Wang et al., "LRTV: MR image superresolution with low-rank and total variation regularizations," *IEEE Transactions on Medical Imaging*, vol. 34, no. 12, p. 1, 2015.

[21] C. Dong, C. C. Loy, K. He et al., "Learning a deep convolutional network for image superresolution," in *Proceedings of the European Conference on Computer Vision*, Zurich, Switzerland, September 2014.

[22] Y. Zhang, Y. Tian, Y. Kong, B. Zhong, and Y. Fu, "Residual dense network for image restoration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018.

[23] C. Ledig, L. Theis, F. Huszar et al., "Photo-realistic single image superresolution using a generative adversarial network," in *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, July 2016.

[24] X. Hu, H. Mu, X. Zhang, Z. Wang, T. Tan, and J. Sun, "Meta-SR: a magnification-arbitrary network for superresolution," in *Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, CA, USA, June 2019.

[25] Z. Wang, J. Chen, and C. H. Hoi Steven, "Deep learning for image superresolution: a survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 99, p. 1, 2020.

Hindawi
Journal of Healthcare Engineering
Volume 2021, Article ID 9958017, 8 pages
https://doi.org/10.1155/2021/9958017

Research Article

A Medical Image Fusion Method Based on SIFT and Deep Convolutional Neural Network in the SIST Domain

Lei Wang[1], **Chunhong Chang**[1], **Zhouqi Liu**[1], **Jin Huang**[1], **Cong Liu**[1] and **Chunxiang Liu**[2]

[1]*School of Computer Science and Technology, Shandong University of Technology, Zibo 255000, China*
[2]*Anhui Key Laboratory of Plant Resources and Plant Biology, Huaibei Normal University, Huaibei 235000, China*

Correspondence should be addressed to Lei Wang; wanglei0511@sdut.edu.cn and Chunxiang Liu; xiaoguo6655@126.com

Received 4 March 2021; Revised 24 March 2021; Accepted 12 April 2021; Published 21 April 2021

Academic Editor: Chi-Hua Chen

The traditional medical image fusion methods, such as the famous multi-scale decomposition-based methods, usually suffer from the bad sparse representations of the salient features and the low ability of the fusion rules to transfer the captured feature information. In order to deal with this problem, a medical image fusion method based on the scale invariant feature transformation (SIFT) descriptor and the deep convolutional neural network (CNN) in the shift-invariant shearlet transform (SIST) domain is proposed. Firstly, the images to be fused are decomposed into the high-pass and the low-pass coefficients. Then, the fusion of the high-pass components is implemented under the rule based on the pre-trained CNN model, which mainly consists of four steps: feature detection, initial segmentation, consistency verification, and the final fusion; the fusion of the low-pass subbands is based on the matching degree computed by the SIFT descriptor to capture the features of the low frequency components. Finally, the fusion results are obtained by inversion of the SIST. Taking the typical standard deviation, $Q^{AB/F}$, entropy, and mutual information as the objective measurements, the experimental results demonstrate that the detailed information without artifacts and distortions can be well preserved by the proposed method, and better quantitative performance can be also obtained.

1. Introduction

The pathology information displayed by medical imaging of different modalities plays a key role in modern medical diagnosis. Unfortunately, it is difficult to synchronously get the full-information images by one imaging device at the same time due to their different imaging principles [1]. Therefore, doctors have to spend more time and energy to read the medical information they want from different devices. A common method to deal with this problem is to fuse the multi-modal images from the same location of the body into one image, which is called the medical image fusion and has been widely used in medical image analysis, precision radiotherapy surgery, and computer-aided medical diagnosis [2].

Nowadays, various medical image fusion methods have been proposed, all of which can be roughly classified into two categories: methods in the spatial domain and in the transformed domain. Different from the former directly using some algebraic operations or filtering, the latter methods, capturing more features in different scales and directions, are the research hotspot. Such scheme usually contains three steps: decomposition, combination, and reconstruction [3].

From the fusion procedure, it is clear that the fusion performance is highly determined by the decomposition tools and the fusion rules. The tools play the role of providing the sparse representations of the features and the fusion rules play the role of transferring the features into the final fusion results. For the decomposition tools, the Laplace Pyramid transform cannot provide directional information; the typical wavelet transform only can decompose the images into three high-pass subbands in each level, so it is limited by the number of directions. The contourlet can get more directional subbands in each level, but the loss of shift-invariance is easy to result in the pseudo-Gibbs

phenomenon [4]. For the fusion rules, the active level measurement-based rule [5] is popular; however, it is easy to produce the artifacts. Though some other fusion rules have been proposed, such as the SVM [6], PCA [7], ICA [8], etc., the fusion results are still unsatisfactory. It is important to consider the feature information during the implementation of the fusion rules [9]. Recently, there has been some good work to improve the fusion performance from these aspects. For example, in literature [10], it proposes a multi-modality image fusion method in the non-subsampled contourlet transform (NSCT), in which the high-pass subbands are integrated by the phase congruency-based rule and the low-pass subbands are combined by the local Laplacian energy-based rule. In literature [11], it proposes an image fusion framework, which integrates NSCT into sparse representation and a principal component analysis (PCA) is implemented in dictionary training to reduce the dimension of learned dictionary. The low- and high-pass coefficients are fused by the sparse representation and Sum Modified-Laplacian, respectively. In literature [12], the source multi-modality images are decomposed into cartoon and texture components. The cartoon components are combined by an energy-based fusion rule for morphological structure preservation and the texture components are combined by the dictionary training. In addition, some similar fusion schemes can be found in the literatures [13, 14]. Such schemes provide good fusion results for they have made full use the of the good mathematical properties of NSCT and the learning abilities of dictionary learning to capture the important features. The main disadvantage, however, is the time cost. The NSCT, the shift-invariance version of the contourlet transform, is too time-consuming because it has to use the non-subsampled band-pass filters to produce the shift-invariance. And the dictionary training usually suffers from the number and the dimension of the dictionary. It is easy to result in the dimensional disaster in the fusion.

Another research hotspot is the neural networks-based medical image fusion methods. Some good results have been reported, such as the artificial neural networks-based method, the Pulse Coupled Neural Network- (PCNN-) based method [15]. However, the fusion performance is limited by how to tune the parameters and the number of the layers in the traditional neural networks models. Very recently, the deep learning technologies, such as the deep convolutional neural network, have achieved great success in the areas of image classification and target recognition, as well as in image fusion. For example, literature [16] proposes a novel image fusion algorithm based on deep support value convolutional neural network, literature [17] proposes the medical image fusion with the all convolutional neural network, and literature [18] proposes a general image fusion framework based on convolutional neural network, which is called IF-CNN. Literatures [19, 20] review the recent advances and future prospects about deep learning for pixel-level image fusion. In the above methods, the good results are obtained for their better learning ability than the traditional neural network models. However, such methods are directly learned in the pixel level, losing the important feature information.

To deal with the above problems, a medical image fusion method based on the SIFT and CNN in the SIST domain is proposed. Different from other transformation tools, such as the wavelet and the contourlet, the SIST can decompose the images into high-pass and low-pass subbands to extract more useful features in different scales and directions. Besides, with the same shift-invariance as the NSCT, the calculation efficiency of SIST is higher. To make full use of the features in the source images, the fusion rule for the low-pass subbands is based on matching degree of the SIFT descriptor. The SIFT feature is based on the local points of interest on the object and is independent of the size and rotation of the image. So, its tolerance for changes in noise and micro-viewpoints is quite high [21]. From this point of view, it is more suitable than the structure features for medical image fusion. The fusion of the high-pass subband is based on the CNN-based scheme to employ the good learning ability of the CNN model.

The rest of the framework is organized as follows. The details of the proposed method are shown in Section 2. Experimental results with important discussions are shown in Section 3. Finally, the conclusion is shown in Section 4.

2. Methodology

The whole procedure of the proposed medical image fusion method is described in Figure 1. After the decomposition and directional partition in different scales, the coefficients of the source medical images are obtained. Then, the high-pass and low-pass coefficients of the fused image are produced by the corresponding fusion rules. Finally, the fusion results are obtained.

The principle of the proposed method can be specifically explained from three aspects to understand: firstly, from the tools of sparse representation in medical image fusion, the SIST has better mathematical properties to provide the good representations of the important features; secondly, the traditional fusion rules are easy to lose the captured feature information during the procedure of transferring them into the final results; thirdly, the transferred feature information is only in the low level and it is not abstract enough to do the feature fusion. Therefore, considering the above needs, a CNN model is pre-trained the get the deep and abstract features in the SIST domain and the SIFT-based fusion rule is developed.

2.1. The Shift-Invariance Shearlet Transform. The discretion of SIST mainly consists of two steps [21]: the multi-scale partition and the directional localization. To provide the shift-invariance, the former step is done by the non-subsampled pyramid filters, and the latter step is implemented by using the filters of shearing. Let j be the scale of image decomposition, $j = 1, 2, \ldots, M$; the whole process can be summarized as the following steps.

(1) The image f^j is decomposed into low-pass image f^{j+1} and high-pass image g^{j+1} using the non-subsampled pyramid f.

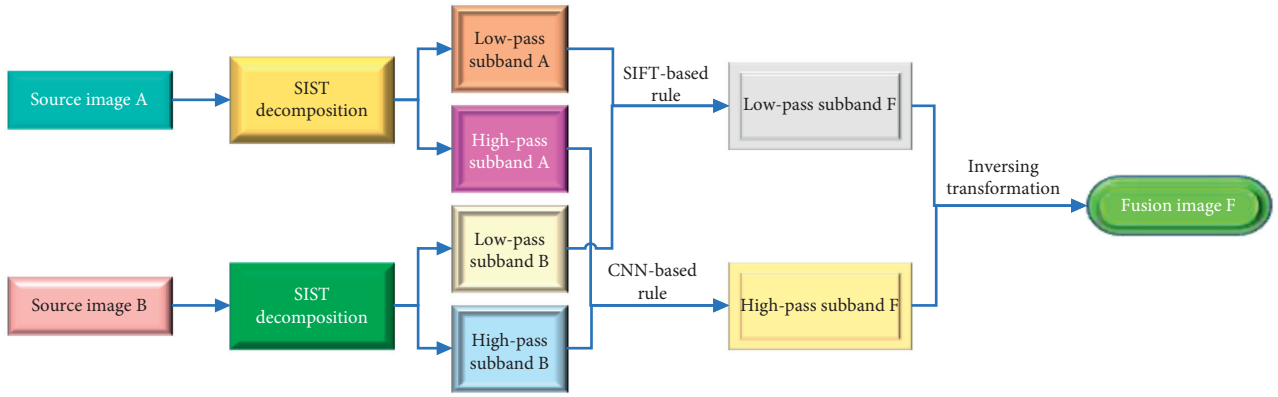


Figure 1: The architecture of the proposed method.

(2) Construct the Meyer Window for the high-pass image g^{j+1}:

Generate shearing filter window W in pseudo-polarization grid;
Map W from pseudo-polarized grid system to Cartesian coordinate system to generate a new shearing filter W_{new};
Compute the 2D discrete Fast Fourier Transform (FFT) of g^{j+1} to generate the matrix Fg^{j+1};
Apply band-pass filtering to the matrix Fg^{j+1} to compute different directional components.

(3) Directly re-assemble the Cartesian sampled values and apply the inversing 2-D FFT to produce the SIST coefficients.

The inversing transformation is the opposite process of the forward transformation. Since there is no need to use directional band-pass banks to get different directions like the NSCT; the SIST is more efficient. More details about the implementation can be found in literatures [22, 23].

2.2. The Fusion of the High-Pass Subbands. The procedure of the high-pass fusion is shown in Figure 2. Before the fusion, a CNN model is trained by the pre-fused images. The whole fusion process mainly consists of four steps: feature detection, initial segmentation, consistency verification [24–26], and the final fusion. In the first step, the high-pass subbands are input into the CNN model to output the score map, which contains the feature information of each high-pass subband. Each coefficient in the score map represents the feature attribute of a pair of corresponding blocks from two high-pass subbands. Then, by averaging the overlapping regions, a feature map of the same size is obtained from the score map. Furthermore, the feature map is segmented into a binary map with the threshold. In the third step, the consistency verification is implemented to refine the binary segmentation mapping to generate the decision map. Finally, the fused image is obtained by applying the pixel-weighted scheme on the decision map.

2.2.1. Train the CNN Model. For a pair of medical image patches $\{A, B\}$ of the pre-fuse images, the goal is to learn a CNN model whose output is a scalar ranging from 0 to 1. Specifically, when the feature is almost from A but not B, the output value should be close to 1, or the value should be close to 0. In other words, the output represents the feature degree of the pair of the image patches. Therefore, a large number of example pairs are used to be the training examples.

In Figure 3, the structure of the trained CNN model is shown. It has two identical architecture branches, each of which takes the medical image blocks as the input. According to [27], it is suitable to set the size of image block to 16×16. There are three convolution layers and a maximum pooling layer in each branch of the network. The size of neuron perception is determined by the core size of the convolution layer. In this paper, the core size is set to 3×3, the step size is set to 1, the scaling factor of the pooling layer is set to 2×2, and the span space is set to 2.

2.2.2. The Feature Detection. Let A_H and B_H be the two high-pass subbands; a score map can be obtained once A_H and B_H are input into the constructed CNN model. The value of each coefficient in the score map ranges from 0 to 1, indicating the feature degree of a pair of 16×16 blocks. The closer that the value is to 1, the more concentrated patches are from image A_H, and vice versa. In order to generate a feature map (represented as M here) of the same size, it assigns the value of each coefficient in the score map to all the coefficients of the corresponding block in M and averages the overlapping pixels.

2.2.3. Initial Segmentation. In order to retain as much useful information as possible, the feature map needs to be applied on the maximum strategy. According to the experience, a d threshold of 0.5 is applied to the feature map to generate the binary map; that is, the focus map is divided by the following formula:

$$T(x, y) = \begin{cases} 1, & M(x, y) > \tau, \\ 0, & \text{otherwise}, \end{cases} \quad (1)$$

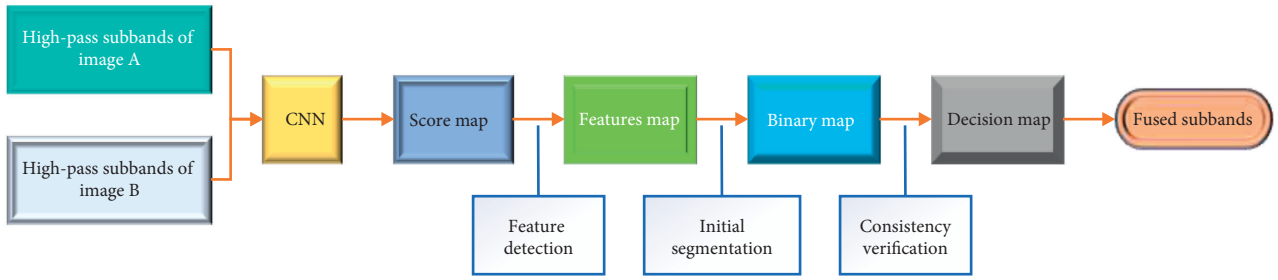


Figure 2: The fusion procedure of the high-pass subbands.

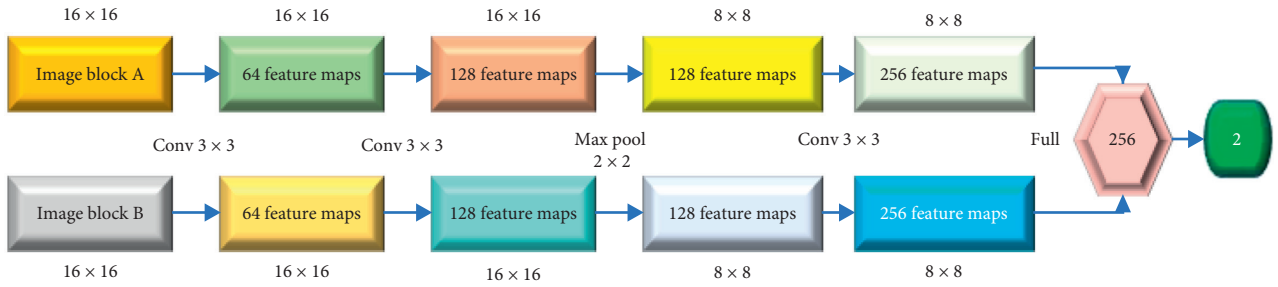


Figure 3: The trained CNN model in this paper.

where is M is the focus map, T is the binary map, is the τ threshold. According to our experience in the experiments, it is found that when the threshold equals 0.5, it is good enough for medical image fusion and it is suggested to be in the range from 0.4 to 0.7 for the other multi-modal image fusion.

2.2.4. The Consistency Verification and Fusion. There may be some misclassified pixels in the above binary map, so it is necessary to remove these mistakes. The traditional method to deal with this problem is to use the threshold scheme, but it is easy to result in some unexpected artifacts around the boundary between the focused and defocused regions. Therefore, the guided filter [28], which is the effective edge-preserving filter to retain the structural information, is employed. There are two free parameters in the guided filtering algorithm: the local window radius r and the regularization parameter ε. In this paper, r is set to 8 and ε is set to 0.1. More details about its implementation can be found in [29]. Finally, the fused high-pass subbands can be obtained by the following weighted formula:

$$F_H(x,y) = D(x,y)A_H(x,y) + (1 - D(x,y))B_H(x,y), \tag{2}$$

where F_{H} is the high-pass subband of the fused image, D is the decision map, and A_H and B_H are the corresponding high-pass subbands of the image to be fused, respectively.

2.3. The Fusion of the Low-Pass Subband. The fusion of the low-pass subband is based on the matching degree of the SIFT [29, 30]. Suppose $\mathrm{fdesc}_1(i)$ and $\mathrm{fdesc}_2(j)$ are the SIFT descriptor from the low-pass subbands of the two images to be fused, where $i \in (1,m)$, $j \in (1,n)$, and m and n are the total number of the SIFT descriptor, respectively. Then, compute the distance $\mathrm{dist}(i,j)$ between $\mathrm{fdesc}_1(i)$ and $\mathrm{fdesc}_2(j)$, and sort all the distances. Let $2\mathrm{nd\,BigDist}(i,j)$ be the second largest value; if $\mathrm{dist}(i,j) < 2\mathrm{nd\,BigDist}(i,j)$, the two SIFTs are called matched.

If $\mathrm{fdesc}_1(i)$ and $\mathrm{fdesc}_2(j)$ are matched, record their location, respectively. If the locations are also the same, it means that both of the content and the location of the region that computed the SIFT descriptors are the same [20]. Finally, the SIFT descriptors that meet the above conditions are recorded to generate a matching degree map match_map, where $1 <= i <= n$. The low-pass subband of the fusion result can be obtained by the following formula:

$$F_L(x,y) = \mathrm{match\_map} * A_L(x,y) \sim \mathrm{match\_}map * B_L(x,y), \tag{3}$$

where "~" means the negation, F_L is the low-pass subband of the fused image, and A_L and B_L are the corresponding low-pass subbands of the image to be fused, respectively.

3. Results and Discussion

In this section, experiments in six groups are carefully done to show the performance of the proposed fusion method. Before the experiments, a CNN model is firstly trained under the public medical data set LIDC, Whole Brain Atlas, and the nature data set ImageNet. All the data sets are downloaded and pre-processed to be the same size of 256 × 256. To get the parameters of the CNN model, 2000 medical images from LIDC, 3000 nature images from the ImageNet, and 200 medical images from the Whole Brain Atlas are, respectively, used to produce the sub-model and the final model is integrated based on the three sub-models. The experimental

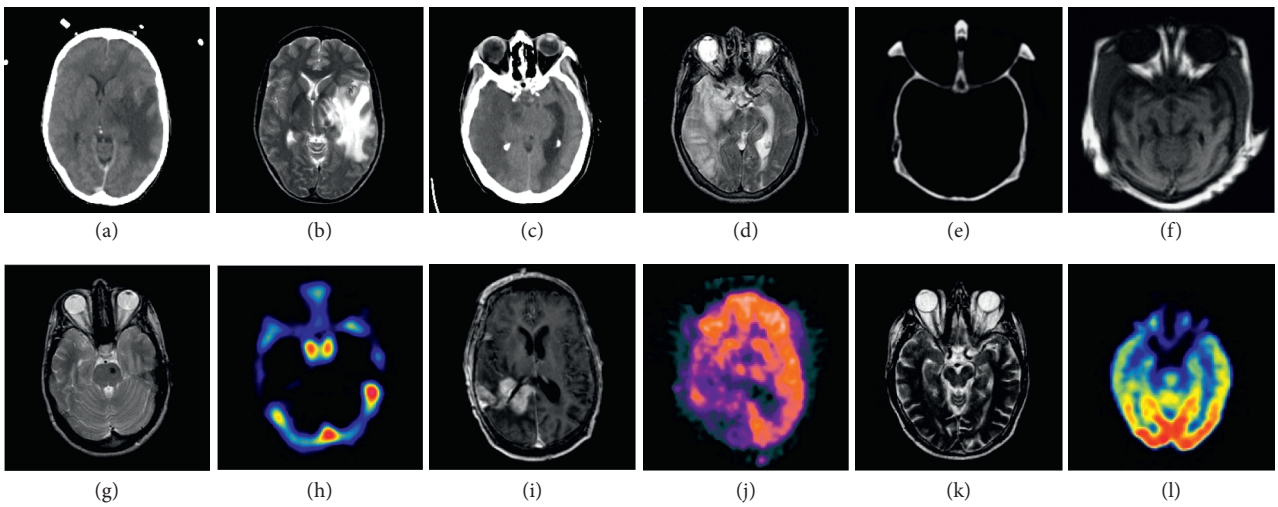


Figure 4: The six groups of source images. Every two of them are captured in the same location with different modalities,CT-MRI in the first row and MRI-PET in the second row.

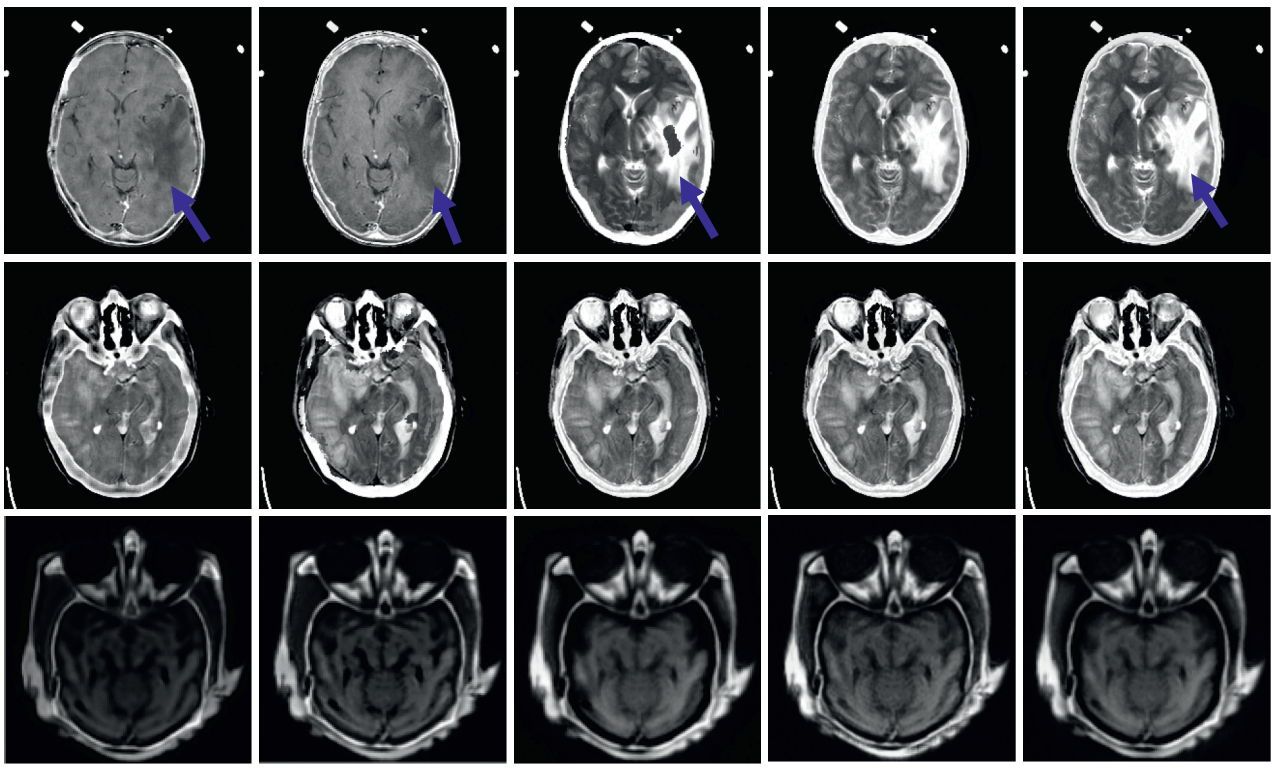


Figure 5: The fusion results of the three groups of CT-MRI. (a) PCNN. (b) CSR. (c) Shearlet. (d) DCNN. (e) Proposed.

platform is the INSPUR big data processing server NF5280M5, Intel Xeon CPU, 128 GB RAM.

Four famous medical image fusion methods, i.e. the Pulse Coupled Neural Network-based method (noted as PCNN) [31], the convolutional sparse representation based method (noted as CSR) [32], the Shearlet based method (noted as Shearlet) [33], and the Deep Convolutional Neural Network-based method (noted as DCNN) [34], are employed to prove the efficiency of the proposed fusion method (Proposed for short). All the parameters are set as the same as what they are reported in the corresponding literature. The decomposition level of SIST is set to 4 and the filters are all "maxflat." After decomposition of each level, 32, 32, 16, and 16 high-pass subbands are obtained.

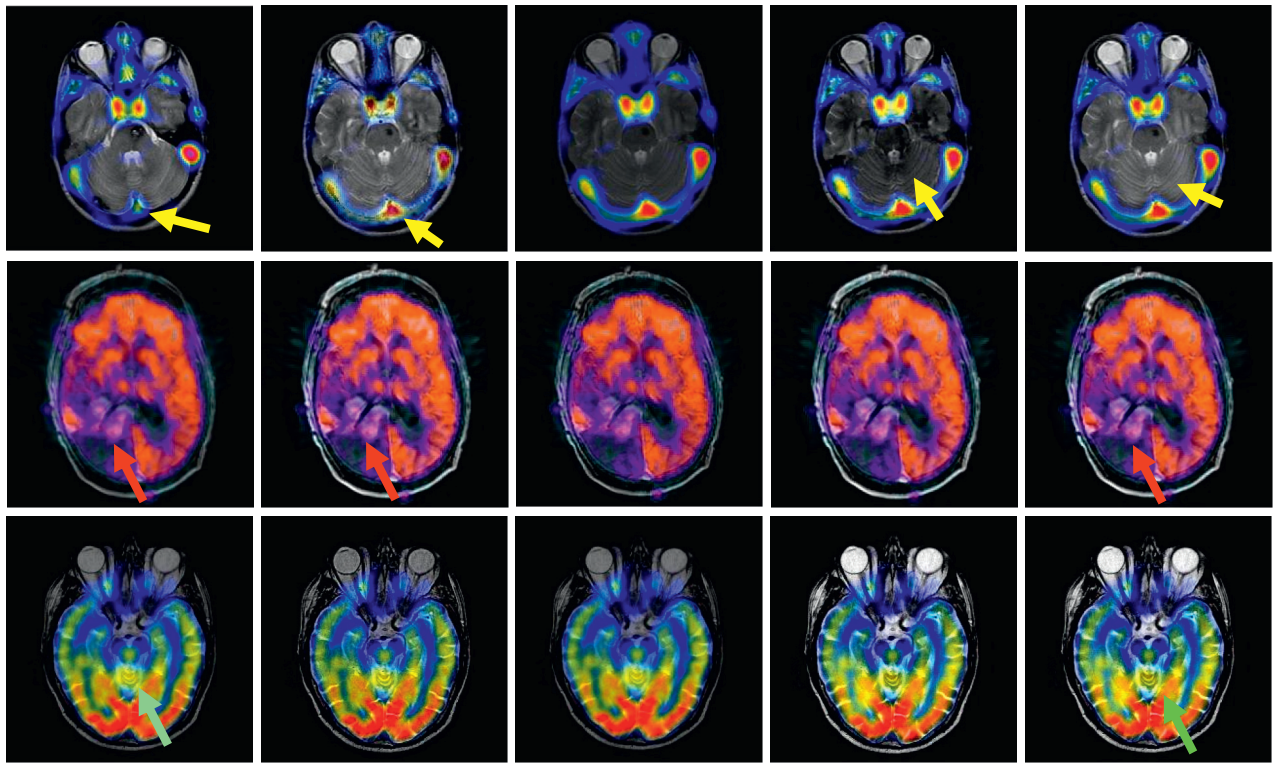
There is no gold standard for evaluating image fusion at present. The usual approach is to use subjectively visual comparisons and objectively quantitative comparisons. This convention is also followed in our paper. Standard deviation (SD for short), entropy (En for short), mutual information (MI for short), and $Q^{AB/F}$ are used to be the

Table 1: The objective evaluations of CT-MRI fusion in Figure 5.

Method	Group 1				Group 2				Group 3			
	SD	$Q^{AB/F}$	En	MI	SD	$Q^{AB/F}$	En	MI	SD	$Q^{AB/F}$	En	MI
PCNN	20.63	0.50	2.12	0.65	20.36	0.55	2.00	0.58	20.87	0.39	1.91	0.45
CSR	20.86	0.55	2.20	0.77	21.56	0.60	2.11	0.62	23.96	0.46	2.00	0.62
Shearlet	21.44	0.59	2.28	0.78	22.43	0.62	2.18	0.69	23.55	0.51	2.10	0.69
DCNN	22.35	0.71	2.31	0.82	23.58	0.68	2.26	0.76	24.20	0.55	2.19	0.78
Proposed	23.22	0.83	2.45	0.88	24.43	0.79	2.33	0.82	24.55	0.68	2.28	0.85

Table 2: The objective evaluations of MRI-PET fusion in Figure 6.

Method	Group 4				Group 5				Group 6			
	SD	$Q^{AB/F}$	En	MI	SD	$Q^{AB/F}$	En	MI	SD	$Q^{AB/F}$	En	MI
PCNN	35.62	0.59	3.76	0.46	41.05	0.62	4.11	0.38	31.12	0.58	3.89	0.62
CSR	37.21	0.65	3.88	0.50	43.56	0.72	4.25	0.42	33.93	0.65	3.96	0.71
Shearlet	37.64	0.67	4.13	0.59	44.95	0.78	4.35	0.48	34.52	0.69	4.31	0.76
DCNN	38.19	0.73	4.50	0.65	45.36	0.81	4.44	0.62	34.91	0.76	4.55	0.81
Proposed	39.42	0.84	4.65	0.70	47.21	0.86	4.63	0.63	35.68	0.83	4.87	0.85



(a) (b) (c) (d) (e)

Figure 6: The fusion results of the three groups of MRI-PET. (a) PCNN. (b) CSR. (c) Shearlet. (d) DCNN. (e) Proposed.

objective evaluation measurements. SD measures the degree of single pixel value relative to the mean value. En shows how much information the image itself contains. MI shows how much information the fused image captured from the source images; $Q^{AB/F}$ measures the edge information transferred from source image to fusion image. The greater the value of these measurements, the better the fusion results [35].

In Figure 4, the three groups in the first row are the gray CT and MRI images, and the three groups in the second row are the color CT and PET images from the patients of anaplastic astrocytoma and mild Alzheimer's disease, respectively. In each data set, the number of slices is 20, 15, and 21, respectively. To save the space, parts of the fusion results are shown in Figures 5 and 6, and the average of the objective evaluation is shown in Tables 1 and 2, respectively.

Though the information expressed by the above fusion images is better than the source image, the fusion results are different. By comparing the arrows of different colors in Figures 5 and 6, it shows that the edge of PCNN method is obviously blurred, and the individual character details (labeled by the blue arrows) and contour features (labeled by the yellow arrows) have been lost. For the Shearlet and DCNN method, the result is clear enough for its good learning ability, but the detail and texture are not good (labeled by the red and green arrows). The main reason is that it is directly learned in the pixel level. In contrast, the detail and texture structures in the fusion results obtained by the proposed method are much clearer, and the ghosting phenomenon can be effectively eliminated. It can be seen that the method proposed is better than CSR and Shearlet method in detail processing. Particularly, comparing the yellow arrow in Figure 6 for the DCNN and the proposed method, it obviously indicates that the information in the results obtained by the proposed method can be kept as what they are like in the source data. In addition, from the objective evaluation in Tables 1 and 2, it can be seen that the objective value of the proposed method is much higher than other methods under the four indicators, which further verifies that more feature information can be effectively captured and fully transferred into the fusion results by the proposed method, showing better visual sensing. All the results prove that more detail features from the source images can be captured and transformed well into the final results by the proposed method.

4. Conclusion

Based on the SIST and the CNN, this paper proposes a medical image fusion method, which makes full use of the multi-resolution and multi-directional characteristics of SIST, and also combines the self-learning advantages of CNN. According to the careful objective analysis and subjective comparison, experiments show that the target information and contour features can be well displayed in the final results. Besides, the artifacts and distortions can be effectively suppressed. Compared with other famous fusion methods such as the PCNN-based method, DCNN-based method, sparse representation-based method, etc., the proposed method can get better fusion results.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by a Project of Shandong Province Higher Educational Science and Technology Program (J18KA362), the National Natural Science Foundation of China (61502282, 61902222), the Natural Science Foundation of Shandong Province (ZR2015FQ005), and the Taishan Scholars Program of Shandong Province (tsqn201909109).

References

[1] B. Meher, S. Agrawal, R. Panda, and A. Abraham, "A survey on region based image fusion methods," *Information Fusion*, vol. 48, pp. 119–132, 2019.

[2] B. Lu, Y. Hu, L. Lin et al., "Using ensemble deep learning method to integrate multi-source data to develop national visibility grid data," *Advances in Meteorological Science and Technology*, vol. 8, pp. 77–82, 2018.

[3] J. Yang, "Medical image fusion algorithm based on quaternion discrete Fourier transform," *Journal of Southwest Normal University*, vol. 45, pp. 1–39, 2020.

[4] H. Ghassemian, "A review of remote sensing image fusion methods," *Information Fusion*, vol. 32, pp. 75–89, 2016.

[5] C. Deng, X. Liu, J. Chanussot, Y. Xu, and B. Zhao, "Towards perceptual image fusion: a novel two-layer framework," *Information Fusion*, vol. 57, pp. 102–114, 2020.

[6] Y. Wang, Xu Chang, and F. Shu, "Analysis of two image classification methods based on SVM algorithm," *Computer and Information Technology*, vol. 27, pp. 18–20, 2019.

[7] V. P. S. Naidu, "Hybrid DDCT-PCA based multi sensor image fusion," *Journal of Optics*, vol. 43, no. 1, pp. 48–61, 2014.

[8] D. Carone, G. W. J. Harston, J. Garrard et al., "ICA-based denoising for ASL perfusion imaging," *NeuroImage*, vol. 200, pp. 363–372, 2019.

[9] J. Tian, G. Liu, and J. Liu, "Multi-focus image fusion based on edges and focused region extraction," *Optik*, vol. 171, pp. 611–624, 2018.

[10] Z. Zhu, M. Zheng, G. Qi, D. Wang, and Y. Xiang, "A phase congruency and local laplacian energy based multi-modality medical image fusion method in NSCT domain," *IEEE Access*, vol. 7, pp. 20811–20824, 2019.

[11] Y. Li, Y. Sun, X. Huang, G. Qi, M. Zheng, and Z. Zhu, "An image fusion method based on sparse representation and sum modified-laplacian in NSCT domain," *Entropy*, vol. 20, no. 7, p. 522, 2018.

[12] Z. Zhu, H. Yin, Y. Chai, Y. Li, and G. Qi, "A novel multi-modality image fusion method based on image decomposition and sparse representation," *Information Sciences*, vol. 432, pp. 516–529, 2018.

[13] Y. Liao, W. Huang, L. Shang et al., "Image fusion based on shearlet and improved PCNN," *Computer Engineering and Application*, vol. 50, pp. 142–146, 2014.

[14] L. Niu and F. Gaofeng, "Multi focus image fusion method based on shearlet and PCNN," *Fire and Command Control*, vol. 2, pp. 41–46, 2016.

[15] Y. Li and T. Xiang, "Infrared and visible image fusion combining edge features and adaptive PCNN in NSCT domain," *Acta Electronica Sinica*, vol. 44, pp. 761–766, 2016.

[16] C. Du, S. Gao, Y. Liu, and B. Gao, "Multi-focus image fusion using deep support value convolutional neural network," *Optik*, vol. 176, pp. 567–578, 2019.

[17] C.-B. Du and S.-S. Gao, "Multi-focus image fusion with the all convolutional neural network," *Optoelectronics Letters*, vol. 14, no. 1, pp. 71–75, 2018.

[18] Y. Zhang, Y. Liu, P. Sun, H. Yan, X. Zhao, and L. Zhang, "IFCNN: a general image fusion framework based on convolutional neural network," *Information Fusion*, vol. 54, pp. 99–118, 2020.

[19] Y. Liu, X. Chen, Z. Wang, Z. J. Wang, R. K. Ward, and X. Wang, "Deep Learning for pixel-level image fusion: recent advances and future prospects," *Information Fusion*, vol. 42, pp. 158–173, 2018.

[20] S. C. Kulkarni and P. P. Rege, "Pixel level fusion techniques for SAR and optical images: a review," *Information Fusion*, vol. 59, pp. 13–29, 2020.

[21] N. Hayat and M. Imran, "Ghost-free multi exposure image fusion technique using dense SIFT descriptor and guided filter," *Journal of Visual Communication and Image Representation*, vol. 62, pp. 295–308, 2019.

[22] W. Zheng, X. Sun, H. Dongmei, and S. Wu, "Thyroid image fusion based on shearlet transform and sparse representation," *Optoelectronic Engineering*, vol. 42, no. 1, pp. 77–83, 2015.

[23] F. Song, Z. Miao, and Z. Zhang, "Multimodal medical image fusion algorithm based on MSVD and MPCNN," *China Digital Medicine*, vol. 14, pp. 9–12, 2019.

[24] L.-L. Kong, Z.-Y. Han, H.-L. Qi, and M.-Y. Yang, "Source retrieval model focused on aggregation for plagiarism detection," *Information Sciences*, vol. 503, pp. 336–350, 2019.

[25] K. Worapan, Q. Wu, R. Panrasee, and J. Zhang, "Hard exudates segmentation based on learned initial seeds and iterative graph cut," *Computer Methods and Programs in Biomedicine*, vol. 158, pp. 173–183, 2018.

[26] J. Dou, Q. Qin, and Z. Tu, "Image fusion based on wavelet transform with genetic algorithms and human visual system," *Multimedia Tools and Applications*, vol. 78, no. 9, pp. 12491–12517, 2019.

[27] Y. Yang, Z. Nie, S. Huang, P. Lin, and J. Wu, "Multilevel features convolutional neural network for multifocus image fusion," *IEEE Transactions on Computational Imaging*, vol. 5, no. 2, pp. 262–273, 2019.

[28] H. Tang, B. Xiao, W. Li, and G. Wang, "Pixel convolutional neural network for multi-Focus image fusion," *Information Sciences*, vol. 433-434, pp. 125–141, 2018.

[29] S. Li, X. Kang, and J. Hu, "Image fusion with guided filtering," *IEEE Transactions on Image Processing: A Publication of the IEEE Signal Processing Society*, vol. 22, no. 7, pp. 2864–2875, 2013.

[30] Z. Yang, J. Lian, S. Li, Y. Guo, and Y. Ma, "A study of sine-cosine oscillation heterogeneous PCNN for image quantization," *Soft Computing*, vol. 23, no. 22, pp. 11967–11978, 2019.

[31] X. Jin, D. Zhou, S. Yao et al., "Multi-focus image fusion method using S-PCNN optimized by particle swarm optimization," *Soft Computing*, vol. 22, no. 19, pp. 6395–6407, 2018.

[32] Y. Liu, X. Chen, R. K. Ward, and Z. Jane Wang, "Image fusion with convolutional sparse representation," *IEEE Signal Processing Letters*, vol. 23, no. 12, pp. 1882–1886, 2016.

[33] Y. Wang, L. fan, and Z. Chen, "Image fusion algorithm based on improved weighting method in shearlet domain and adaptive PCNN," *Computer Science*, vol. 46, pp. 261–267, 2019.

[34] Y. Liu, X. Chen, H. Peng, and Z. Wang, "Multi-focus image fusion with a deep convolutional neural network," *Information Fusion*, vol. 36, pp. 191–207, 2017.

[35] S. Lin, Z. Han, D. Li et al., "Integrating model- and data-driven methods for synchronous adaptive multi-band image fusion," *Information Fusion*, vol. 54, pp. 145–160, 2020.

Hindawi
Journal of Healthcare Engineering
Volume 2021, Article ID 6618818, 10 pages
https://doi.org/10.1155/2021/6618818

Research Article

Using DEMATEL Technique to Identify the Key Success Factors of Shared Decision-Making Based on Influential Network Relationship Perspective

Wen-Yi Liu,[1,2,3] Tao-Hsin Tung,[4] Yen-Ching Chuang,[5] and Ching-Wen Chien[2]

[1]*Department of Health Policy Management, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA*
[2]*Institute for Hospital Management, Tsing Hua University, Shenzhen Campus, Shenzhen, China*
[3]*Shanghai Bluecross Medical Science Institute, Shanghai, China*
[4]*Evidence-Based Medicine Center, Taizhou Hospital of Zhejiang Province Affiliated to Wenzhou Medical University, Linhai 317000, China*
[5]*Institute of Public Health & Emergency Management, Taizhou University, Taizhou, China*

Correspondence should be addressed to Yen-Ching Chuang; yenching.chuang@gmail.com and Ching-Wen Chien; ihhca@sz.tsinghua.edu.cn

Received 24 November 2020; Revised 4 March 2021; Accepted 29 March 2021; Published 21 April 2021

Academic Editor: Chi-Hua Chen

In the field of medicine, shared decision-making (SDM) is an important issue primarily aimed at resolving the problem of information asymmetry between clinicians and patients in the selection of treatment options and follow-up nursing plans. Most previous studies on this topic have focused on key elements and the development and implementation of SDM scales. This study used the decision-making trial and evaluation laboratory (DEMATEL) method to establish a network of influence relationships among factors that are keys to the success of the SDM process. Survey data were obtained from a well-known brain hospital in China. The key factors of success included tailor information, flexibility approach, check understanding patient, document (discussion about) decision, present evidence, make or explicitly defer decision, and patient values and preferences. We determined that clinicians should provide a series of treatment options and follow-up care plans based on a patient's conditions and preferences. Clinicians should also actively communicate with patients and their families to ensure a thorough understanding of the entire treatment and nursing process. This study also highlights the academic value of the cross-disciplinary integration of medical decision issues and multiple attribute decision-making methodologies.

1. Introduction

Advances in medical technology have increased the number of treatment options, thereby necessitating the involvement of patients in shared decision-making (SDM) [1, 2]. SDM is a patient-oriented approach to the selection of treatment options, emphasizing the need to consider the professional knowledge of clinicians within the context of patient values [3–5]. Considered best practice in terms of medical ethics, SDM is meant to reduce the problem of information asymmetry between clinicians and patients [6]. By bridging the gap between the priorities of clinicians and patients, SDM is also expected to improve drug compliance and patient satisfaction, well-being, and quality of life [7–10]. Clinicians tend to base decisions on medical knowledge and evidence, whereas patients consider the rigors of recovery and their quality of life.

Previous research on SDM can be divided into (1) key elements/features [11–13], (2) barriers and facilitators [14–16], and (3) SDM scales [17–21]. Identifying key factors of success is a multiple attribute decision-making (MADM) problem dealing with multiple criteria and the preferences of different decision-makers at the same time. Dolan [22] used the analytic hierarchy process (AHP) to improve communication between clinicians and patients in SDM. Dolan et al. [23] used an AHP-based decision model to analyse the

screening priorities of patients. Those studies demonstrated the academic value of MADM in the context of medical care.

The AHP method is based on the assumption that relationships between attributes are independent; however, this is not necessarily consistent with actual operations [24, 25]. Two methods have been developed to explore the relationship between attributes: interactive structural modelling (ISM) [26] and the decision-making trial and evaluation laboratory (DEMATEL) [27]. ISM uses 0 and 1 scales to construct an influence relationship network among attributes. Under this scheme, computation is simple; however, the information generated is sparce [28]. By contrast, DEMATEL illustrates the degree of influence among key elements as well as the interdependency of relationships between attributes using an influential network relationship map (INRM) based on graph theory [29].

This study proposes a hybrid shared decision-making and multiattribute decision-making model (called SDM-MADM model). First of all, the SDM index system is established on the basis of previous literature review of key elements of SDM. Then, the DEMATEL method calculates the total influence degree among the key elements and shows their influence network relationship map (INRM). Finally, based on INRM, clinicians can understand the influence relationship structure among key elements and identify a few key success factors in SDM. The contributions of this research are as follows: (i) we present an SDM index system based on a review of the literature pertaining to SDM and practical clinical experience; (ii) this study demonstrates the potential of applying MADM methods to the issue of medical management; and (iii) the proposed SDM-MADM model can elucidate for clinicians the factors that are key to the success of SDM in order to facilitate a meaningful exchange with patients.

Empirical data in this study were obtained from actual clinical experience in a major hospital in China. The remainder of this study is organized as follows. Section 2 presents a review of the literature pertaining to SDM. Section 3 presents the proposed SDM-MADM model. Section 4 illustrates the clinical application of the SDM-MADM model in a real-world setting. Conclusions and future research directions are presented in Section 5.

2. Literature Review

SDM is an important factor in treatment counselling [30]. Researchers have long sought to develop a core definition of codecision-making and consistent measures related to patient outcomes [31, 32]. The most widely accepted SDM model draws on the work of Charles et al., which outlined a number of key features [33, 34]: (i) the patient and clinician are both involved in the treatment decision-making process, (ii) the patients and clinician share information, (iii) the patient and clinician express treatment preferences, and (iv) the patient and clinician jointly make treatment decisions and agree on follow-up treatment. Subsequent research can be divided into (1) the development and application of SDM scales and (2) key elements/features in the SDM process.

2.1. The Development and Application of SDM Scales. Braddock et al. [17] proposed six key elements of informed decision-making (IDM) for routine practice. Braddock et al. [18] then expanded the IDM model to include seven key elements to explore the nature and integrity of informed decisions made by primary care doctors and surgeons in routine outpatient clinics. Simon et al. [35] compiled and verified a theory-driven questionnaire to assess the shared decision process, referred to as the shared decision-making questionnaire (SDM-Q). Kriston et al. [19] revised the original SDM-Q as a new 9-item instrument referred to as the SDM-Q-9. Kheirkhah et al. [36] applied the SDM-Q-Doc version to explore the psychological predictors of SDM between residents and patients. Based on the SDM-Q-9, Kriston et al. [37] formulated a network model to assess the competency of clinicians in applying SDM and predicting outcomes. Song and Sandra [38] used a 10-item IDM scale to investigate the experience of elderly patients in postdialysis decision-making. Rothberg et al. [39] used a 7-item IDM scale to assess IDM and its correlation with patient decision-making. The IDM and SDM scales can be used alone or in combination with other scales to broaden analysis to other issues, such as quality of life.

2.2. Composition of Key Elements in the SDM Process. Using content analysis, DuBenske et al. [12] identified 3 key components and 13 corresponding key elements. Makoul and Clayman [31] identified 3 key components and 23 corresponding elements in the SDM process. Legare and Witteman [40] identified 3 key factors in routine clinical practice. Bomhof-Roordink et al. [11] identified 53 elements and classified them within 24 categories. Dy and Purnell [41] conducted a literature review on the key concepts related to SDM in healthcare. The studies above provided a framework to develop SDM index systems in subsequent MADM models.

2.3. Research Gaps. The above research focused on the application of SDM scales and key elements in SDM. The former examined the use of SDM, alone and when implemented with other clinical scales. The latter demonstrated the importance of key elements in SDM; however, there was a failure to identify the factors that are keys to the success of such endeavours. In this study, we used the MADM methodology to overcome this problem.

3. Methodology

In this study, we developed an integrated SDM-MADM model, which began with the development of an index system based on a review of the literature pertaining to key elements in SDM. We then applied the DEMATEL method to the proposed index system to establish the structure of influence among key elements.

3.1. The SDM Index System. We began by examining changes in the shared decision-making environment and the views of

decision-makers over time. Bomhof-Roordink [11] performed a systematic search of SDM-related articles published prior to September 2, 2019, from seven electronic databases (i.e., Academic Search Premier, Cochrane, Embase, Emcare, PsycINFO, PubMed, and Web of Science). From these databases, they identified 53 elements in SDM model descriptions and classified them within 24 components. Note that key elements apply to three types of SDM models: (1) models containing more than half of the components, (2) models focusing on a single actor within components (i.e., healthcare professional), and (3) models in which the components vary over time (without a clear pattern). In the current study, we based our index system on the research of Bomhof-Roordink [11] and corresponding references.

Our primary objective in this study was to explore the interdependent relationships among key elements of SDM and identify the factors that are keys to its success. We created a model covering more than half of the components with top ranking main components (i.e., recurring in research with the highest frequency) as our framework. The top 4 components appeared in more than 60% of the studies, as follows: "Describe treatment options" (80%), "Make the decision" (75%), "Patient preferences" (65%), and "Tailor information" (65%). Table 1 lists the references corresponding to the key elements of SDM models.

3.2. The DEMATEL Method. The DEMATEL method was developed by the Battelle Memorial Institute in 1972 to resolve interdependent structure problems in the real world [27]. This graph theory-based method can be used to analyse and solve problems involving relationships of mutual influence via visualization [50, 51]. Structural modelling makes it easier for a researcher to understand interdependent relationships among elements in a complex system as well as resolve complex problems based on an understanding of cause-effect relationships. DEMATEL has been applied with considerable success in green supply chain management [52, 53], cultural products [54], shelter sites [55], company performance [56], and water management [57]. The computation steps are outlined in the following [58–60]:

Step 1. Initial influence relation matrix **A** based on domain knowledge.

The practical experience of domain experts can be leveraged to establish a direct influence relation matrix, in which the numerical value represents the degree of influence between attribute/element i and attribute/element j. Note that the values are derived from a set of 5-value Likert scales (from no influence (0) to very high influence (4)). Direct influence relation matrices are integrated into matrix **A** via averaging in order to assess the opinions of a group of experts pertaining to the interactions among attributes, as follows:

$$\mathbf{A} = \left[a_{ij}\right]_{n\times n} = \left[\frac{\left(\sum_{\eta=1}^{k} e_{ij}^{\eta}\right)}{k}\right]_{n\times n}, \tag{1}$$

where the diagonal values in the matrix are 0, the other values are between 0 and 4, and n indicates the total number of attributes/elements in the system.

The level of significant confidence of matrix **A** can be obtained using equation (2). Here, the threshold must exceed 95% (i.e., average ratio gap less than 5%), where a higher ratio indicates results of higher stability/confidence.

$$\frac{1}{n(n-1)} \sum_{i=1}^{k} \sum_{j=1}^{k} \frac{\left|a_{ij}^{k} - a_{ij}^{k-1}\right|}{a_{ij}^{k}} \times 100\%. \tag{2}$$

Step 2. Normalized influence relation matrix **B** transferred from matrix **A**.

The range of the influence relationship depicted in matrix **A** requires a series of conversion actions to establish the preliminary state of the total influence relationship between attributes/elements. This normalization process is implemented as follows:

$$\mathbf{B} = \rho \times \mathbf{A}, \tag{3}$$

$$\rho = \min\left\{\frac{1}{\max_i \sum_{j=1}^{n} a_{ij}}, \frac{1}{\max_j \sum_{i=1}^{n} a_{ij}}\right\}, \quad i, j \in \{1, 2, \ldots, n\}, \tag{4}$$

where ρ is the adjustment coefficient, and the maximum value of the sum of rows or columns of matrix **B** is 1.

Step 3. Total influence relation matrix **O** calculated from matrix.

In accordance with the Markov chain, the total influence degree (direct, indirect, and subsequent influences) between attributes/elements can be estimated under the condition of matrix **B**. Total influence relation matrix **O** is derived as follows:

$$\begin{aligned}\mathbf{O} &= \mathbf{B} + \mathbf{B}^2 + \mathbf{B}^3 + \cdots + \mathbf{B}^{\tau} \\ &= \mathbf{B}\left(\mathbf{I} + \mathbf{B} + \mathbf{B}^2 + \cdots + \mathbf{B}^{\tau-1}\right)\left[(\mathbf{I} - \mathbf{B})(\mathbf{I} - \mathbf{B})^{-1}\right] \\ &= \mathbf{B}\left(\mathbf{I} - \mathbf{B}^{\tau}\right)(\mathbf{I} - \mathbf{B})^{-1} \\ &= \mathbf{B}(\mathbf{I} - \mathbf{B})^{-1)}, \quad \text{when } \tau \longrightarrow \infty,\ \mathbf{B}^{\tau} = [0]_{n\times n},\end{aligned} \tag{5}$$

where matrix **I** refers to the identity matrix.

Step 4. Influence network relationship map (INRM).

This map is used to visualize the net mutual influence among attributes in order to identify key elements and elucidate the means by which they interact. First, four indices are extended from matrix **O** as main components of the network diagram: "influence given (r_i)," "influence received (u_i)," "prominence/centrality ($r_i + u_i$)," and "relation/causal effect ($r_i - u_i$)," such as equations (6)–(9). In constructing the space of the network diagram, prominence/centrality ($r_i + u_i$) and the relation/causal-effect ($r_i - u_i$) are, respectively, listed along the x-axis and y-axis. Finally, the direction of the main influence between attributes/elements

Table 1: Key elements of the shared decision-making model.

Aspect	Attribute	Description	Ref
Describe treatment options (C_1)	Benefits/risk (C_{11})	Doctors and patients should discuss the advantages and disadvantages of the proposed treatment options because they may have different views on the relative importance of the benefits, disadvantages, risks, and costs of treatment.	[11, 31, 42–44]
	Feasibility of options (C_{12})	Doctors and patients should review the feasibility of treatment options based on a variety of reference information, including the patient's physical condition, treatment methods, and costs.	[11, 31, 45]
	List options (C_{13})	Doctors should make a clear list of treatment options based on medical and clinical knowledge. Patients should also propose other treatment options of which they may know.	[4, 5, 11, 31, 44, 46]
	Present evidence (C_{14})	Doctors should base decisions on existing medical evidence.	[11, 31]
Make the decision (C_2)	Document (discussion about) decision (C_{21})	It is important to record the issues and decisions related to the treatment options discussed by doctors and patients, including patient concerns, preferences, and information needs as well as treatment and nursing methods.	[11, 43]
	Make or explicitly defer decision (C_{22})	Patients do not always make decisions when they first discuss problems, postponing them to a later time (e.g., waiting for discussions with families and/or medical team members). Doctors and patients should therefore arrange follow-up visits to track outstanding decisions and reach a resolution.	[11, 21, 31, 33, 34, 46, 47]
	Patient retains ultimate authority over decision (C_{23})	Patients should maintain decision-making power pertaining to final treatment options.	[11, 45, 48]
	Revisiting decision (C_{24})	In cases where the available treatment options cannot produce the expected health results, the patient may reconsider the decision.	[11, 44, 45]
Patient preferences (C_3)	Patient concerns (C_{31})	Clinicians should consider the concerns (fears) of patients pertaining to treatment options, including postoperative recovery, quality of life, and medical expenses.	[11, 21, 32]
	Patient goals of care (C_{32})	Clinicians should consider the expectations of patients in terms of nursing objectives (preoperative and postoperative).	[4, 11, 21]
	Patient values and preferences (C_{33})	Clinicians should consider the values of patients (e.g., outlook on life, concepts, spirit, and culture).	[4, 11, 21, 31]
Tailor information (C_4)	Ascertain preferred information (C_{41})	Clinicians should provide patients with all important information.	[11, 47]
	Check understanding patient (C_{42})	Clinicians should confirm the overall situation of the patient, including illness, concerns, and preferences.	[11, 31, 44, 47]
	Flexibility approach (C_{43})	As situations change, doctors and patients should be prepared to reconsider their decision(s). The decision-making process must remain flexible.	[11, 31]
	Use clear language (C_{44})	Clinicians should use simple language in the provision of information to patients.	[5, 11, 49]

is adopted as the net influence relationship for subsequent evaluations.

Influence given is

$$r_i = (r_i)_{n\times 1} = (r_1, \ldots, r_i, \ldots, r_n) = \left[\sum_{j=1}^{n} o_{ij}\right]_{n\times 1}. \quad (6)$$

Influence received is

$$u_i = (u_i)_{n\times 1} = (u_j)_{1\times n}{}' = (u_1, \ldots, u_j, \ldots, u_n)' = \left[\sum_{i=1}^{n} o_{ij}\right]_{1\times n}{}'. \quad (7)$$

Prominence/centrality is

$$r_i + u_i, \quad i \in \{1, 2, \ldots, n\}. \quad (8)$$

Relation/causal effect is

$$r_i - u_i, \quad i \in \{1, 2, \ldots, n\}. \tag{9}$$

where $'$ represents the transposed action; r_i represents the total impact of attributes/elements on other attributes/elements (including itself) within the system; u_i represents the total impact of attributes/elements affected by other attributes/elements (including itself) within the system; $r_i + u_i$ indicates the influence strength of attributes/elements in the system, where a higher value indicates higher correlation between itself and all other attributes; and $r_i - u_i$ indicates the net influence of attributes/elements in the system. The values can be divided into three scenarios: (1) values exceeding 0 mainly affect other attributes/elements and may be included in the reason group; (2) values lower than 0 are affected mainly by other attributes/elements and may be included in the results group; and (3) values close to 0 do not have a significant net influence.

4. Empirical Case Study, Results, and Discussion

In this section, we illustrate an SDM-MADM model based on the experience of 11 clinicians. We also identify the most influential factors and discuss the degree to which these insights many inform management decisions.

4.1. Background Description and Data Sources. The hospital in this case study was a three-level brain hospital overseen by the Shanghai Municipal Health Committee. In establishing this hospital, administrators sought to build a "Chinese Mayo Clinic." At this hospital, the process of diagnosis and treatment begins with clinicians seeking to gain a comprehensive understanding of a patient's condition and economic situation. They accordingly draw up a suitable medical plan and recommendations for the patient. Finally, clinicians and patients jointly formulate a final medical plan. Essentially, shared decision-making is used as the basis of clinical decision-making within the context of patient concerns and preferences. Discussions between the two sides are meant to reduce information asymmetry in the decision-making process. For these reasons, the concept and management mode of case hospital is very suitable as a demonstration case for this study.

There are 11 domain experts in this study, all of whom had more than 10–15 years of practical clinical experience as attending doctors ($n = 5$) and director physicians ($n = 6$). As outlined above, a 5-point Likert scale (no influence (0) to very high influence (4)) is routinely used to transform clinical experience into a direct influence relation matrix. The 11 direct influence relation matrices were integrated into a matrix **A** via averaging to create the initial direct influence relation matrix, as shown in Table 2. The confidence level of the matrix was 97.41% (i.e., average rat).

4.2. Influence Structure among Attributes. Using equations (3) and (4), initial influence relation matrix **A** was converted into normalized influence relation matrix **B** with values between 0 and 1. Using equation (5), the total influence degree between attributes was calculated to obtain total influence relation matrix **O** (Table 3).

Using equations (6)–(9), we then identified the 4 influential indicators for each attribute (Table 4). The top attributes of the aspect level in the prominence/centrality ($r_i + u_i$) index were ranked from high to low as follows: make the decision (C_2) ≻ describe treatment options (C_1) ≻ patient preferences (C_3) ≻ tailor information (C_4). The top attributes in the relation/causal effect ($r_i - u_i$) were ranked from high to low as follows: tailor information (C_4) ≻ patient preferences (C_3) ≻ describe treatment options (C_1) ≻ make the decision (C_2). At the attribute level, the top attributes in the prominence/centrality ($r_i + u_i$) were ranked from high to low as follows: patient retains ultimate authority over decision (C_{23}) ≻ patient concerns (C_{31}) ≻ make or explicitly defer decision (C_{22}). The top attribute in the relation/causal effect ($r_i - u_i$) was patient values and preferences (C_{33}).

Based on prominence/centrality ($r_i + u_i$) and the relation/causal effect ($r_i - u_i$) in Table 4, we established an influence network-relation map (INRM) of the SDM index system, as shown in Figure 1. Aspect analysis (Figure 1) identified tailor information (C_4) as the most important aspect based on to its effects on patient preferences (C_3), describe treatment options (C_1), and make the decision (C_2). This means that when communicating with patients, clinicians should seek to tailor information specifically to the needs of the patient. When we looked more closely at the issue of tailor information (C_4), we found that the most important attribute was check understanding patient (C_{42}). Intuitively, this appears correct as it helps to ensure that clinicians explain procedures using simple language. The most important attributes under patient preferences (C_3) were patient values and preferences (C_{33}). Naturally, considering the values and preferences of patients is the only way to allay their concerns. The most important attribute under describe treatment options (C_1) was present evidence (C_{14}). Obviously, treatment options must be based on practical clinical research and evidence. The most important attribute under make the decision (C_2) was make or explicitly defer decision (C_{22}). Generally, patients feel the need to confer with friends and family before making important medical decisions. Note also that patients are expected to make the final decisions.

4.3. Analysis and Discussion. Check understanding patient (C_{42}), patient values and preferences (C_{33}), make or explicitly defer decision (C_{22}), and present evidence (C_{14}) were identified as the most influential factors. These may also be regarded as the key factors determining success in overcoming information asymmetry via joint decision-making. SDM is meant to be used by clinicians and patients jointly when making medical decisions. It is based on the sharing of information for which evidence exists in order to satisfy the needs of the patient based on their values and preferences [3–5, 61]. Thus, the first task in SDM is to ensure that patients make decisions with a full understanding of key issues, rather than in the face of avoidable ignorance [62].

Table 2: Initial influence relation matrix **A**.

	C_{11}	C_{12}	C_{13}	C_{14}	C_{21}	C_{22}	C_{23}	C_{24}	C_{31}	C_{32}	C_{33}	C_{41}	C_{42}	C_{43}	C_{44}
C_{11}	0.00	3.18	1.82	2.00	2.36	2.64	2.45	2.82	2.45	2.55	1.27	2.09	2.00	1.55	1.82
C_{12}	3.00	0.00	2.27	2.91	2.45	2.36	2.91	2.64	2.82	2.45	1.91	2.00	1.82	1.82	2.00
C_{13}	2.45	2.27	0.00	2.18	2.45	2.27	2.73	2.82	2.27	2.18	1.91	2.18	1.82	1.73	1.91
C_{14}	2.73	3.18	2.27	0.00	2.91	2.73	2.82	2.73	2.55	2.09	1.82	2.27	1.55	1.91	2.00
C_{21}	2.73	2.64	2.73	2.91	0.00	2.73	2.64	2.73	2.55	2.18	1.82	2.64	1.55	2.00	2.18
C_{22}	3.00	3.09	3.09	2.55	2.82	0.00	2.64	2.82	2.64	2.27	1.73	2.18	1.82	1.82	2.00
C_{23}	2.55	3.00	2.82	2.18	3.18	2.55	0.00	2.73	2.45	2.55	2.00	2.55	2.09	2.00	2.36
C_{24}	2.27	2.55	2.45	2.45	2.36	2.64	2.64	0.00	2.27	2.36	1.55	2.45	1.91	1.91	2.18
C_{31}	2.55	2.55	2.64	2.45	2.18	2.64	2.82	2.55	0.00	2.73	2.55	2.36	2.36	1.91	2.36
C_{32}	2.18	2.64	2.18	2.09	2.36	2.64	3.00	2.55	2.55	0.00	1.73	2.09	2.09	1.64	2.00
C_{33}	2.00	2.27	2.09	1.91	2.18	2.45	2.64	2.36	2.82	2.73	0.00	2.00	2.45	2.00	2.36
C_{41}	2.36	2.36	2.45	2.64	2.64	2.82	2.82	2.73	3.00	2.45	2.36	0.00	2.09	2.27	2.55
C_{42}	2.00	2.36	2.64	2.18	2.73	2.45	2.36	2.45	2.36	2.18	2.27	2.18	0.00	2.27	2.27
C_{43}	2.00	2.18	2.27	1.82	2.45	2.18	2.45	2.64	2.55	2.18	1.82	2.27	2.00	0.00	2.18
C_{44}	1.82	2.18	2.36	2.00	2.45	2.55	2.73	2.73	2.82	2.36	2.27	2.55	2.18	2.00	0.00

Note: the value is obtained by equation (1), and the significant confidence level of the matrix is 97.41% (i.e., average ratio gap is 2.59%).

Table 3: Total influence relation matrix **O**.

	C_{11}	C_{12}	C_{13}	C_{14}	C_{21}	C_{22}	C_{23}	C_{24}	C_{31}	C_{32}	C_{33}	C_{41}	C_{42}	C_{43}	C_{44}
C_{11}	0.42	0.52	0.46	0.45	0.49	0.50	0.52	0.52	0.50	0.47	0.37	0.44	0.39	0.37	0.42
C_{12}	0.52	0.48	0.50	0.50	0.53	0.52	0.56	0.55	0.54	0.50	0.40	0.47	0.41	0.40	0.45
C_{13}	0.48	0.51	0.42	0.46	0.50	0.49	0.53	0.52	0.50	0.46	0.38	0.45	0.39	0.38	0.42
C_{14}	0.52	0.56	0.51	0.43	0.54	0.54	0.56	0.56	0.53	0.49	0.40	0.48	0.41	0.40	0.45
C_{21}	0.52	0.55	0.52	0.51	0.47	0.54	0.56	0.56	0.54	0.50	0.41	0.49	0.41	0.41	0.46
C_{22}	0.53	0.57	0.54	0.50	0.55	0.48	0.57	0.57	0.55	0.51	0.41	0.49	0.42	0.41	0.46
C_{23}	0.53	0.57	0.54	0.50	0.56	0.55	0.51	0.57	0.55	0.52	0.42	0.50	0.43	0.42	0.47
C_{24}	0.48	0.52	0.49	0.47	0.51	0.51	0.54	0.47	0.51	0.48	0.38	0.46	0.40	0.39	0.44
C_{31}	0.52	0.56	0.53	0.50	0.54	0.55	0.58	0.57	0.49	0.52	0.43	0.49	0.44	0.41	0.47
C_{32}	0.48	0.52	0.48	0.46	0.50	0.51	0.54	0.53	0.51	0.42	0.38	0.45	0.40	0.38	0.43
C_{33}	0.48	0.52	0.49	0.46	0.50	0.51	0.54	0.53	0.52	0.49	0.35	0.46	0.42	0.39	0.44
C_{41}	0.53	0.57	0.54	0.52	0.56	0.56	0.59	0.58	0.57	0.52	0.44	0.44	0.44	0.43	0.48
C_{42}	0.49	0.53	0.50	0.47	0.52	0.52	0.54	0.54	0.52	0.48	0.41	0.47	0.36	0.40	0.45
C_{43}	0.46	0.50	0.47	0.44	0.49	0.49	0.52	0.52	0.50	0.46	0.38	0.45	0.39	0.33	0.42
C_{44}	0.49	0.53	0.50	0.47	0.52	0.52	0.55	0.55	0.53	0.49	0.41	0.48	0.42	0.40	0.39

Note: the value is obtained by equations (3)–(5).

Table 4: The four influence indicators for each attribute/element.

	r_i	u_i	$r_i + u_i$	$r_i - u_i$		r_i	u_i	$r_i + u_i$	$r_i - u_i$
C_1	1.891 (4)	2.006 (2)	3.897 (2)	−0.115 (3)	C_{11}	6.844 (14)	7.444 (8)	14.288 (11)	−0.601 (12)
					C_{12}	7.323 (7)	8.003 (3)	15.326 (4)	−0.679 (14)
					C_{13}	6.873 (13)	7.497 (7)	14.370 (9)	−0.624 (13)
					C_{14}	7.370 (6)	7.148 (10)	14.519 (8)	0.222 (6)
C_2	1.977 (1)	2.128 (1)	4.105 (1)	−0.151 (4)	C_{21}	7.459 (5)	7.780 (6)	15.239 (5)	−0.320 (10)
					C_{22}	7.542 (4)	7.799 (5)	15.341 (3)	−0.258 (7)
					C_{23}	7.659 (2)	8.210 (1)	15.870 (1)	−0.551 (11)
					C_{24}	7.049 (11)	8.135 (2)	15.184 (6)	−1.086 (15)
C_3	1.920 (3)	1.875 (3)	3.796 (3)	0.045 (2)	C_{31}	7.581 (3)	7.862 (4)	15.443 (2)	−0.281 (8)
					C_{32}	6.998 (12)	7.305 (9)	14.302 (10)	−0.307 (9)
					C_{33}	7.099 (10)	5.969 (14)	13.068 (14)	1.129 (1)
C_4	1.935 (2)	1.714 (4)	3.650 (4)	0.221 (1)	C_{41}	7.780 (1)	7.012 (11)	14.792 (7)	0.767 (4)
					C_{42}	7.187 (9)	6.118 (13)	13.306 (13)	1.069 (2)
					C_{43}	6.834 (15)	5.930 (15)	12.764 (15)	0.904 (3)
					C_{44}	7.259 (8)	6.644 (12)	13.903 (12)	0.615 (5)

Note: the value is obtained by equations (6)–(9), and the value in () is ranking.

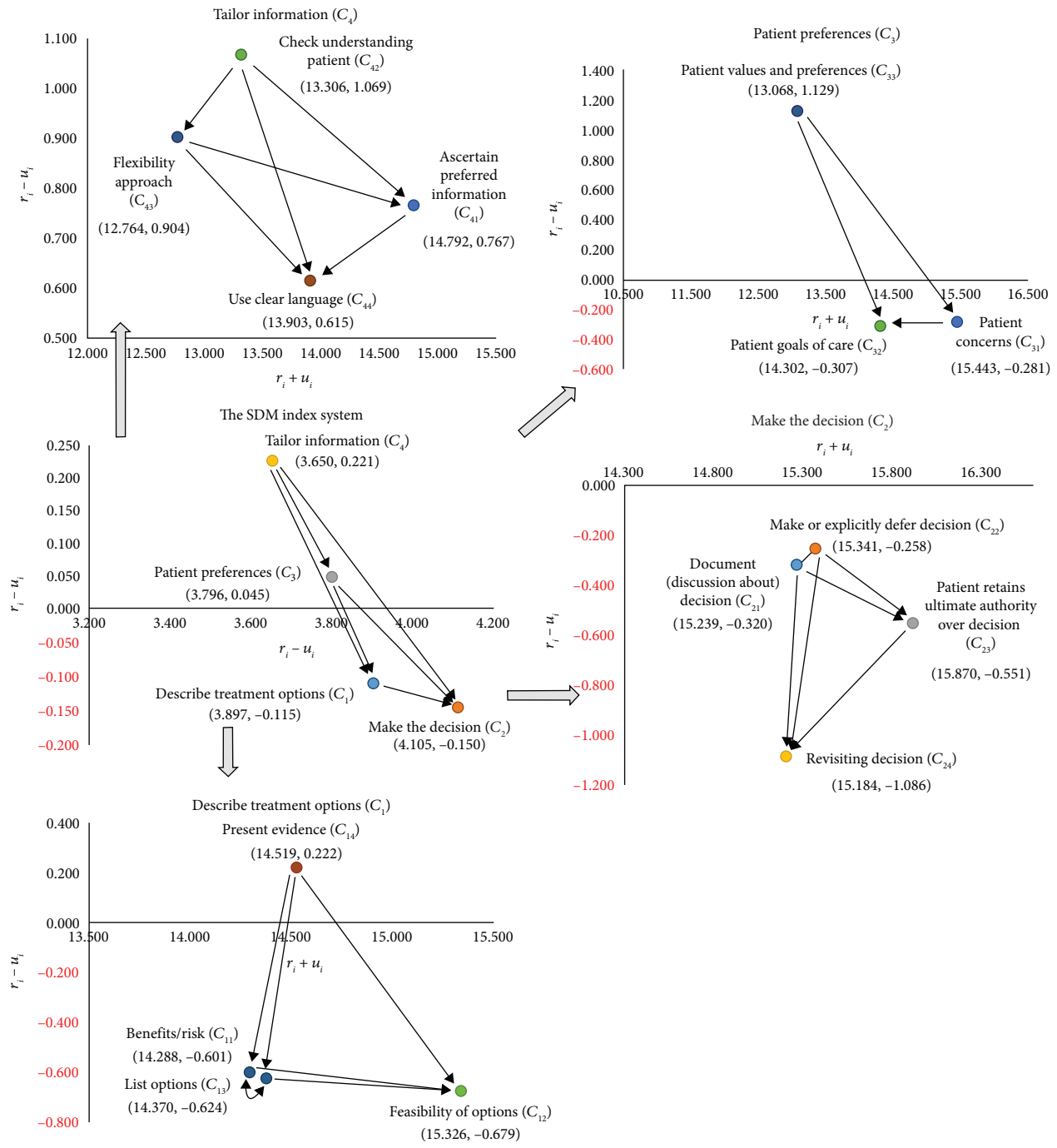


Figure 1: Influential network-relation map (INRM).

One randomized trial study revealed that patients who possessed more knowledge were more confident in making decisions [63]. Only when clinicians make an effort to confirm that patients fully comprehend all of the information that they have received can patients then make informed treatment decisions [64]. From the perspective of clinical management, SDM should focus on communication between clinicians and patients through a patient-centred exchange of information. Clinicians should strive to explain all options, advantages, and disadvantages of any given treatment plan as well as the subsequent nursing plan. Likewise, patients and their families should feel free to outline their concerns, values, and preferences to facilitate the establishment of further options in line with their wishes and expectations. Finally, clinicians should learn to bestow information using simple language and clear illustrations.

5. Conclusions

SDM aimed at reducing information asymmetry between clinicians and patients should be based on clinical evidence

within the context of patient preferences. In this study, we employed the DEMATEL technique to construct a network of influential relationships among key elements pertaining to SDM. Our analysis revealed four issues that are crucial to the success of all such endeavours: check understanding patient (C_{42}), patient values and preferences (C_{33}), make or explicitly defer decision (C_{22}), and present evidence (C_{14}). Clinicians should seek to consider the concerns and preferences of patients in the presentation of medical options and follow-up care plans. Clinicians should also strive to ensure that the patient fully comprehends the entire treatment process as well as its advantages and disadvantages. At the same time, patients should be encouraged to express their concerns and thoughts and feel free to ask questions. Different clinical conditions may have different results, so the results of this study are limited to the experience of the hospital in this case. In addition, this study does not consider various ambiguities in the decision-making environment. The future MADM model can be combined with fuzzy theory or the grey system to discuss the shared decision-making problem more deeply.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Wen-Yi Liu and Tao-Hsin Tung contributed equally in this study.

Acknowledgments

The authors are very grateful to the brain doctors of the case hospital for assisting in the questionnaire investigation and the reviewers for their valuable opinions.

References

[1] T. Treffers and P. M. Putora, "Emotions as social information in shared decision-making in oncology," *Oncology*, vol. 98, pp. 430–437, 2020.

[2] N. Singh Ospina, F. J. K. Toloza, F. Barrera, C. L. Bylund, P. J. Erwin, and V. Montori, "Educational programs to teach shared decision making to medical trainees: a systematic review," *Patient Education and Counseling*, vol. 103, pp. 1082–1094, 2020.

[3] F. Legare, D. Stacey, S. Pouliot et al., "Interprofessionalism and shared decision-making in primary care: a stepwise approach towards a new model," *Journal of Interprofessional Care*, vol. 25, pp. 18–25, 2011.

[4] M. H. van de Pol, C. R. Fluit, J. Lagro, Y. H. Slaats, M. G. Olde Rikkert, and A. L. Lagro-Janssen, "Expert and patient consensus on a dynamic model for shared decision-making in frail older patients," *Patient Education and Counseling*, vol. 99, pp. 1069–1077, 2016.

[5] A. Waschwill, A. Bittner, and S. Harendza, "Assessment of medical students' shared decision-making skills in simulated physician-patient encounters," *Patient Education and Counseling*, vol. 103, pp. 500–504, 2020.

[6] M. Tai-Seale, G. Elwyn, C. J. Wilson et al., "Enhancing shared decision making through carefully designed interventions that target patient and provider behavior," *Health Affairs*, vol. 35, pp. 605–612, 2016.

[7] M. A. Stewart, "Effective physician-patient communication and health outcomes: a review," *Canadian Medical Association Journal*, vol. 152, pp. 1423–1433, 1995.

[8] E. A. Joosten, L. DeFuentes-Merillas, G. H. de Weert, T. Sensky, C. P. van der Staak, and C. A. de Jong, "Systematic review of the effects of shared decision-making on patient satisfaction, treatment adherence and health status," *Psychotherapy and Psychosomatics*, vol. 77, pp. 219–226, 2008.

[9] C. Rathert, M. D. Wyrwich, and S. A. Boren, "Patient-centered care and outcomes: a systematic review of the literature," *Medical Care Research and Review*, vol. 70, pp. 351–379, 2013.

[10] L. A. Shay and J. E. Lafata, "Where is the evidence? a systematic review of shared decision making and patient outcomes," *Medical Decision Making*, vol. 35, pp. 114–131, 2015.

[11] H. Bomhof-Roordink, F. R. Gartner, A. M. Stiggelbout, and A. H. Pieterse, "Key components of shared decision making models: a systematic review," *BMJ Open*, vol. 9, Article ID e031763, 2019.

[12] L. L. DuBenske, S. B. Schrager, M. E. Hitchcock et al., "Key elements of mammography shared decision-making: a scoping review of the literature," *Journal of General Internal Medicine*, vol. 33, pp. 1805–1814, 2018.

[13] G. Darien, C. Wilson, A. Balch, and A. Rebekah, "Talking about the cost of care: a critical component of shared decision making: patient and provider perspectives," *Journal of the National Medical Association*, vol. 112, pp. 225–228, 2020.

[14] L. Boland, I. D. Graham, F. Legare et al., "Barriers and facilitators of pediatric shared decision-making: a systematic review," *Implementation Science*, vol. 14, p. 7, 2019.

[15] K. Gravel, F. Legare, and I. D. Graham, "Barriers and facilitators to implementing shared decision-making in clinical practice: a systematic review of health professionals' perceptions," *Implementation Science*, vol. 1, p. 16, 2006.

[16] F. Legare, S. Ratte, K. Gravel, and I. D. Graham, "Barriers and facilitators to implementing shared decision-making in clinical practice: update of a systematic review of health professionals' perceptions," *Patient Education and Counseling*, vol. 73, pp. 526–535, 2008.

[17] C. H. Braddock 3rd, S. D. Fihn, W. Levinson, A. R. Jonsen, and R. A. Pearlman, "How doctors and patients discuss routine clinical decisions. Informed decision making in the outpatient setting," *Journal of General Internal Medicine*, vol. 12, pp. 339–345, 1997.

[18] C. H. Braddock 3rd, K. A. Edwards, N. M. Hasenberg, T. L. Laidley, and W. Levinson, "Informed decision making in outpatient practice: time to get back to basics," *Journal of the American Medical Association*, vol. 282, pp. 2313–2320, 1999.

[19] L. Kriston, I. Scholl, L. Holzel, D. Simon, A. Loh, and M. Harter, "The 9-item Shared Decision Making Questionnaire (SDM-Q-9). Development and psychometric properties in a primary care sample," *Patient Education and Counseling*, vol. 80, pp. 94–99, 2010.

[20] E. M. Driever, A. M. Stiggelbout, and P. L. P. Brand, "Shared decision making: physicians' preferred role, usual role and

their perception of its key components," *Patient Education and Counseling*, vol. 103, pp. 77–82, 2020.

[21] G. Elwyn, M. Tsulukidze, A. Edwards, F. Legare, and R. Newcombe, "Using a 'talk' model of shared decision making to propose an observation-based measure: observer OPTION 5 Item," *Patient Education and Counseling*, vol. 93, pp. 265–271, 2013.

[22] J. G. Dolan, "Shared decision-making--transferring research into practice: the analytic hierarchy process (AHP)," *Patient Education and Counseling*, vol. 73, pp. 418–425, 2008.

[23] J. G. Dolan, E. Boohaker, J. Allison, and T. F. Imperiale, "Can streamlined multicriteria decision analysis be used to implement shared decision making for colorectal cancer screening?," *Medical Decision Making*, vol. 34, pp. 746–755, 2014.

[24] Y.-C. Chuang, S.-K. Hu, J. J. H. Liou, and H.-W. Lo, "Building a decision dashboard for improving green supply chain management," *International Journal of Information Technology & Decision Making*, vol. 17, pp. 1363–1398, 2018.

[25] S. S. Weng, Y. Liu, and Y. C. Chuang, "Reform of Chinese universities in the context of sustainable development: teacher evaluation and improvement based on hybrid multiple criteria decision-making model," *Sustainability*, vol. 11, 2019.

[26] A. Vafadarnikjoo, H. B. Ahmadi, B. T. Hazen, and J. J. H. Liou, "Understanding interdependencies among social sustainability evaluation criteria in an emerging economy," *Sustainability*, vol. 12, 2020.

[27] A. Gabus and E. Fontela, *World Problems, an Invitation to Further Thought within the Framework of DEMATEL*, Battelle Institute, Geneva Research Center, Geneva, Switzerland, 1972.

[28] W. C. J. Hsu, J. J. H. Liou, and H. W. Lo, "A group decision-making approach for exploring trends in the development of the healthcare industry in Taiwan," *Decision Support Systems*, vol. 141, 2021.

[29] H. W. Lo, J. J. H. Liou, C. N. Huang, Y. C. Chuang, and G. H. Tzeng, "A new soft computing approach for analyzing the influential relationships of critical infrastructures," *International Journal of Critical Infrastructure Protection*, vol. 28, Article ID 100336, 2020.

[30] A. N. Siriwardena, A. G. Edwards, P. Campion, A. Freeman, and G. Elwyn, "Involve the patient and pass the MRCGP: investigating shared decision making in a consulting skills examination using a validated instrument," *British Journal of General Practic*, vol. 56, pp. 857–862, 2006.

[31] G. Makoul and M. L. Clayman, "An integrative model of shared decision making in medical encounters," *Patient Education and Counseling*, vol. 60, pp. 301–312, 2006.

[32] M. C. Weiss and T. J. Peters, "Measuring shared decision making in the consultation: a comparison of the OPTION and Informed Decision Making instruments," *Patient Education and Counseling*, vol. 70, pp. 79–86, 2008.

[33] C. Charles, A. Gafni, and T. Whelan, "Shared decision-making in the medical encounter: what does it mean? (or it takes at least two to tango)," *Social Science & Medicine*, vol. 44, pp. 681–692, 1997.

[34] C. Charles, A. Gafni, and T. Whelan, "Decision-making in the physician-patient encounter: revisiting the shared treatment decision-making model," *Social Science & Medicine*, vol. 49, pp. 651–661, 1999.

[35] D. Simon, G. Schorr, M. Wirtz et al., "Development and first validation of the shared decision-making questionnaire (SDM-Q)," *Patient Education and Counseling*, vol. 63, pp. 319–327, 2006.

[36] F. Kheirkhah, R. M. Larijani, M. Faramarzi, M. H. Yadollahpour, and S. Khafri, "Psychological predictors of medical residents' perspectives on shared decision-making with patients: a cross-sectional study," *BMC Medical Education*, vol. 20, p. 251, 2020.

[37] L. Kriston, P. Hahlweg, M. Harter, and I. Scholl, "A skills network approach to physicians' competence in shared decision making," *Health Expectations*, vol. 23, pp. 1466–1476, 2020.

[38] M. K. Song and S. E. Ward, "The extent of informed decision-making about starting dialysis: does patients' age matter?," *Journal of Nephrology*, vol. 27, pp. 571–576, 2014.

[39] M. B. Rothberg, S. K. Sivalingam, R. Kleppel, M. Schweiger, B. Hu, and K. R. Sepucha, "Informed decision making for percutaneous coronary intervention for stable coronary disease," *JAMA Internal Medicine*, vol. 175, pp. 1199–1206, 2015.

[40] F. Legare and H. O. Witteman, "Shared decision making: examining key elements and barriers to adoption into routine clinical practice," *Health Affairs*, vol. 32, pp. 276–284, 2013.

[41] S. M. Dy and T. S. Purnell, "Key concepts relevant to quality of complex and shared decision-making in health care: a literature review," *Social Science & Medicine*, vol. 74, pp. 582–587, 2012.

[42] S. Turcotte, H. Robitaille, L. Blair, and F. Legare, "The actor-partner interdependence model in shared decision-making: an illustrative example of its application to the physician-patient dyad in primary care consultations," *Journal of Clinical Epidemiology*, vol. 108, pp. 132–139, 2019.

[43] M. A. Schonberg, M. B. Hamel, R. B. Davis et al., "Development and evaluation of a decision aid on mammography screening for women 75 years and older," *JAMA Internal Medicine*, vol. 174, pp. 417–424, 2014.

[44] S. Rennke, P. Yuan, B. Monash et al., "The SDM 3 circle model: a literature synthesis and adaptation for shared decision making in the hospital," *Journal of Hospital Medicine*, vol. 12, pp. 1001–1008, 2017.

[45] F. Legare, D. Stacey, S. Gagnon et al., "Validating a conceptual model for an inter-professional approach to shared decision making: a mixed methods study," *Journal of Evaluation in Clinical Practice*, vol. 17, pp. 554–564, 2011.

[46] G. Elwyn, A. Edwards, P. Kinnersley, and R. Grol, "Shared decision making and the concept of equipoise: the competences of involving patients in healthcare choices," *British Journal of General Practic*, vol. 50, pp. 892–897, 2000.

[47] M. Truglio-Londrigan, J. T. Slyer, J. K. Singleton, and P. Worral, "A qualitative systematic review of internal and external influences on shared decision-making in all health care settings," *JBI Library of Systematic Reviews*, vol. 10, pp. 4633–4646, 2012.

[48] C. Y. Lin, L. Renwick, and K. Lovell, "Patients' perspectives on shared decision making in secondary mental healthcare in Taiwan: a qualitative study," *Patient Education and Counseling*, vol. 103, pp. 2565–2570, 2020.

[49] E. J. Kim and S. H. Kim, "Simplification improves understanding of informed consent information in clinical trials regardless of health literacy level," *Clinical Trials*, vol. 12, pp. 232–236, 2015.

[50] E. Fontela and A. Gabus, *The DEMATEL Observer*, Battelle Geneva Research Center, Geneva, Switzerland, 1976.

[51] C. Bai and J. Sarkis, "A grey-based DEMATEL model for evaluating business process management critical success factors," *International Journal of Production Economics*, vol. 146, pp. 281–292, 2013.

[52] H. H. Wu and S. Y. Chang, "A case study of using DEMATEL method to identify critical factors in green supply chain

management," *Applied Mathematics and Computation*, vol. 256, pp. 394–403, 2015.

[53] Z. Wang, K. Mathiyazhagan, L. Xu, and A. Diabat, "A decision making trial and evaluation laboratory approach to analyze the barriers to Green Supply Chain Management adoption in a food packaging company," *Journal of Cleaner Production*, vol. 117, pp. 19–28, 2016.

[54] C. T. Lin, C. Lee, S. W. Wang, and Y. H. Chen, "Using the DEMATEL method to explore the critical factors that influence visitors to purchase museum cultural products," *Journal of Testing and Evaluation*, vol. 46, 2018.

[55] A. Trivedi, "A multi-criteria decision approach based on DEMATEL to assess determinants of shelter site selection in disaster response," *International Journal of Disaster Risk Reduction*, vol. 31, pp. 722–728, 2018.

[56] E. Yadegaridehkordi, M. Hourmand, M. Nilashi, L. Shuib, A. Ahani, and O. Ibrahim, "Influence of big data adoption on manufacturing companies' performance: an integrated DEMATEL-ANFIS approach," *Technological Forecasting and Social Change*, vol. 137, pp. 199–210, 2018.

[57] R. X. Nie, Z. P. Tian, J. Q. Wang, H. Y. Zhang, and T. L. Wang, "Water security sustainability evaluation: applying a multi-stage decision support framework in industrial region," *Journal of Cleaner Production*, vol. 196, pp. 1681–1704, 2018.

[58] Y. C. Chuang, S. K. Hu, J. J. H. Liou, and G. H. Tzeng, "A data-driven madm model for personnel selection and improvement," *Technological and Economic Development of Economy*, vol. 26, pp. 751–784, 2020.

[59] J. J. H. Liou, Y. C. Chuang, E. K. Zavadskas, and G. H. Tzeng, "Data-driven hybrid multiple attribute decision-making model for green supplier evaluation and performance improvement," *Journal of Cleaner Production*, vol. 241, 2019.

[60] R. D. Raut, V. S. Yadav, N. Cheikhrouhou, V. S. Narwane, and B. E. Narkhede, "Big data analytics: implementation challenges in Indian manufacturing supply chains," *Computers in Industry*, p. 125, 2021.

[61] G. Elwyn, S. Laitner, A. Coulter, E. Walker, P. Watson, and R. Thomson, "Implementing shared decision making in the NHS," *BMJ*, vol. 341, Article ID c5146, 2010.

[62] G. Elwyn, D. Frosch, R. Thomson et al., "Shared decision making: a model for clinical practice," *Journal of General Internal Medicine*, vol. 27, pp. 1361–1367, 2012.

[63] A. M. O'Connor, D. Stacey, V. Entwistle et al., "Decision aids for people facing health treatment or screening decisions," *Cochrane Database of Systematic Reviews*, vol. 12, no. 4, Article ID CD001431, 2003.

[64] A. Schattner, A. Bronstein, and N. Jellin, "Information and shared decision-making are top patients' priorities," *BMC Health Services Research*, vol. 6, p. 21, 2006.

Hindawi
Journal of Healthcare Engineering
Volume 2021, Article ID 6654600, 13 pages
https://doi.org/10.1155/2021/6654600

Research Article

A DEA-Based Decision Support Framework for Organizations' Performance Evaluation considering TQM and Knowledge Management

Azin Karami,[1] Hadi Shirouyehzad ,[1] and Milad Asadpour [2,3]

[1]*Department of Industrial Engineering, Najafabad Branch, Islamic Azad University, Najafabad, Iran*
[2]*Department of Information Systems and Operations Management, Business School, The University of Auckland, Auckland, New Zealand*
[3]*Young Researchers and Elite Club, Najafabad Branch, Islamic Azad University, Najafabad, Iran*

Correspondence should be addressed to Hadi Shirouyehzad; hadi.shirouyehzad@gmail.com

Received 13 November 2020; Revised 16 February 2021; Accepted 8 March 2021; Published 14 April 2021

Academic Editor: Chi-Hua Chen

Total quality management (TQM) and knowledge management (KM) play a significant role in improving the situation of the organization and creating and maintaining competitive advantage (CA). Therefore, harmonizing TQM and KM with the organizational goals and performance leads to the organization's greater capabilities. The current paper presents a decision support framework to evaluate the performance of organizations based on the success factors (SFs) of TQM and KM using the BCC model of data envelopment analysis (DEA). The SFs of KM are considered as input and the SFs of TQM are considered as output of the BCC model, and the efficiency scores of five pharmaceutical companies in Esfahan, Iran, are obtained. The results show that, among the aforementioned companies, three companies were recognized as efficient, while two of them were introduced as inefficient companies. Furthermore, in order to provide more recommendations for the managers of these companies, two area charts according to SFs of TQM and KM were drawn, and the results were analyzed and discussed.

1. Introduction

TQM is the art of managing an entire organization to get the best out of it. In the TQM approach, the main focus is on quality, particularly the quality of work and processes. Therefore, TQM is contrasted with result-oriented management, which only emphasizes result and more production. In general, the important principles governing TQM are the senior management's commitment, customer orientation, evaluation, and decision-making based on facts, participation and collaboration, training, and continuous improvement [1]. TQM is a smart, smooth, and continuous action that has a synergistic effect on achieving the goals of the organization and ultimately leads to customer satisfaction, increasing efficiency, and boosting the ability of competing in the market [2]. One of the main features of TQM is its ability of integrating techniques and methods which are related to management and quality issues comprehensively. Additionally, it has a close relationship with KM and performance management. Consequently, TQM's principles and attitudes underlie the gradual evolution of the learning organization, and its attitudes radically change it [3].

On the other hand, knowledge is assumed to be a strategic and intangible asset that can help organizations maintain their competitiveness in a turbulent environment [4]. KM is the success secret of organizations in the 21st century. Organizational managers can rely on superior knowledge to make more sensible decisions on important issues and improve knowledge-based performance [5]. In fact, KM is an organizational issue that seeks to control the spiritual and intangible capacities, experiences, and skills of employees. Therefore, in the increasing and accelerating enlargements era, organizations should employ novel managerial approaches to gain CA and survive in the

market. However, creating new knowledge in the organization is required for most quality enhancement programs [6–10].

Companies need a TQM procedure that presents knowledge as a potential source of CA. TQM and KM play an important role in improving the situation of the organization, its long-term survival, as well as creating a CA for it. Accordingly, how to institutionalize these two important issues (TQM and KM) and harmonizing them with organizational goals and performance is important. Hence, the managers of organizations should try to create a mechanism for making a special situation for TQM and KM as fundamental rules which are accepted by employees, because successful organizations have the capacity of attracting TQM and KM as an organizational culture. Nonetheless, based on a review on previous research in the areas of TQM and KM, although the combination of DEA and TQM or DEA and KM has been used in the previous studies, any research which simultaneously investigates TQM and KM factors in performance evaluation applying a DEA approach was not observed. Also, among conducted research in pharmaceutical companies, any research which considered TQM and KM in performance evaluation concurrently was not observed. Meanwhile, since TQM and KM play an important role in providing continuous improvement and CA to all industries including pharmaceutical sector, it is vital to examine their simultaneous impact on the performance of pharmaceutical companies. Therefore, in the present study, the BCC model of DEA is applied to measure the performance of five pharmaceutical companies in Esfahan, Iran, over the period of March 2018 to September 2018, so that the SFs of KM and the SFs of TQM are considered as the inputs and outputs of the model, respectively. Besides, managerial directions based on the results of two area charts have been provided.

For this purpose, firstly, by reviewing previous research and consulting university experts, the SFs of TQM and KM were determined and, using distributed questionnaires within pharmaceutical companies in Esfahan, the importance and performance of SFs of TQM and KM were measured. The reliability and validity of the questionnaires were also investigated using Cronbach's alpha, Exploratory Factor Analysis, and Lawshe's method, respectively. Then, using data mining algorithms, the SFs of TQM and KM were clustered. Considering the clustered SFs of KM and TQM as input and output of a variable Return to Scale (RTS) model of DEA, namely, BCC model, the relative efficiency scores for aforementioned pharmaceutical companies were calculated. To sum up, the innovative contributions of this research are as follows:

(i) Identifying and clustering SFs of TQM in pharmaceutical companies

(ii) Identifying and clustering SFs of KM in pharmaceutical companies

(iii) Simultaneous usage of clustered SFs of TQM and KM for performance evaluation by deploying data envelopment analysis

(iv) Applying the proposed approach within real pharmaceutical companies

(v) Providing practical suggestions and managerial insights with respect to results of implementing the proposed approach within the pharmaceutical companies

The rest of the current paper is organized as follows.

To begin, the previous related studies are reviewed, the research gap, and the research innovation is highlighted. Next, the concepts and tools are elaborated. In the methodology section, the steps of the present research are explained. After that, in the data collecting and analysis section, the collected data are prepared to enter the BCC model and performance evaluation of the pharmaceutical companies. Following this and in the managerial directions section, two area charts are presented and analyzed for the SF of TQM and KM. Finally, in the conclusion section, the most important results along with the limitations of the current paper are stated while some suggestions for future research are provided as well.

2. Literature Review

It seems that, among the previous studies, the most contribution is related to simultaneous consideration of TQM and KM and their relationships. Hsu and Shen [11] discussed the resemblances and differences between KM and TQM and suggested that both of them can be considered as a complement for another one. Molina et al. [12] evaluated how quality management (QM) affects knowledge transfers. The results obtained from a case study in Spain approved that different QM practices can be influenced on both internal and external knowledge transfers. Chong et al. [13] integrated TQM and KM approaches into a model to consider whether the hybrid procedure can improve the adoption of collaborative commerce or not. Daud and Yusoff [14] consider the effect of TQM criteria on KM processes in public firms. Results show that soft criteria of TQM such as human capital have more influence on KM and lead to better performance of organization. Ooi [15] investigates the connection between TQM and KM in manufacturing and nonmanufacturing companies in Malaysia from a multidimensional perspective. Findings reveal the relationship between TQM dimensions and KM dimensions. Alshatnawi and Abd Ghani [16] study the comprehension of staff within Jordan universities towards the application of TQM and KM rules. In particular, they indicate how comprehension of the connection between TQM and KM can affect performance of organization. Obeidat et al. [17] investigate the effect of KM criteria on TQM practices within Jordan banks. Results state that most KM criteria have a sharp effect on TQM practices. Alkhazali et al. [18] consider the influence of TQM, on human resource management (HRM) practices and KM strategies and their integrated impact on performance of Jordan banks simultaneously. Results show that TQM as a moderating factor along with HRM practices and KM strategies can lead to improvement of Jordan banks' performance. Abbas [19] examines the connection of KM, TQM, and corporate sustainability (CS) within

manufacturing and nonmanufacturing companies in Pakistan. Results indicate that TQM affects CS dramatically while KM somehow mediates this relationship.

On the one hand, the innovation has been the subject of some hybrid research on TQM and KM. Hung et al. [20] studied the effect of organizational learning and TQM on innovative performance within a case study in Taiwan. Results show that both TQM and organizational learning affect innovative performance significantly. Honarpour et al. [21] illustrate the relation between KM and TQM and their integrated impact on innovation performing the Joint Variance Analysis method. In another study, Honarpour et al. [22] concentrate on R&D units in 190 firms of Malaysia and indicate how KM and TQM connection can affect the performance of R&D units from the innovation point of view.

On the other hand, the combination of DEA with either KM or TQM approaches has been noticed in the related literature. Kuah and Wong [23] applied a two-level DEA model for measuring the performance of Malaysian universities with regard to KM factors while Shirouyehzad et al. [24] considered the criteria of KM and safety management as inputs and customer satisfaction and accident indicators as outputs of a DEA model for performance evaluation of 12 companies in the car industry in Esfahan. In addition, Salhieh and Abu-Doleh [25] indicated the impact of TQM on banks efficiency applying a DEA-based approach. However, some research can be found which applied DEA for selecting, improving, or evaluating six sigma projects as well. For example, Hadi-Vencheh and Yousefi [26] and Wen et al. [27] applied a hybrid DEA technique to select six sigma projects whereas Azadeh et al. [28] performed a PCA-DEA approach to investigate the effect of six sigma extension on key job features in car industry.

Apart from TQM and KM, different methods and approaches have been employed by researchers for performance evaluation of the pharmaceutical companies. Kamath [29] studied the connection between intellectual capital (IC) criteria, profitability, productivity, and market valuation within pharmaceutical sector in India applying VAIC method. Also, Sharabati et al. [30] considered the connection between IC criteria and business performance using real data from Jordan pharmaceutical industry performing the structural equation modelling method. Mehralian et al. [31] studied the connection between IC indicators and profitability, productivity, and market valuation using real data from Iranian pharmaceutical industry and applying a hybrid approach including correlation, simple linear multiple regression, and ANN. Enekwe et al. [32] studied the impact of financial leverage on financial performance within Nigeria pharmaceutical industry during a determined time period. They have used descriptive statistics, Pearson's correlation, and regression for analyzing collected data.

Nevertheless, DEA technique has been used in several studies for performance evaluation of the pharmaceutical companies. Tavana et al. [33] studied public pharmaceutical firms and proposed a three-stage hybrid approach including MCDM, balanced scorecard, and two fuzzy DEA models. Fuzzy DEA models calculate the efficiency scores of the firms using operating budget, cost of goods as inputs, and market share, earnings per share, P/E ratio, sales growth, rank of liquidity, and volume of exports as outputs. Varmaghani et al. [34] used Malmquist index to examine the productivity status within Iranian pharmaceutical companies during 2000–2013. They considered total assets, capital stock as inputs, and net sales and net profit as outputs of the DEA model. Gascón et al. [35] evaluated the efficiency of 37 large pharmaceutical laboratories during 2008–2013 applying DEA approach. They considered employees, total assets, investment in R&D, and number of days discounted until December 13 as inputs and net income basic, market capitalization, net sales, and number of days discounted until December 13 as outputs. Al-Refaie et al. [36] applied DEA for performance evaluation of blistering lines. The planned production quantity in units, defect quantity in units, and idle time in units were considered as inputs while the actual produced quantity was the output of DEA model. Alam and Rastgi [37] performed DEA for performance evaluation of five pharmaceutical firms in India. Net block, cash and bank balance, share capital, reserve and surplus, secure loan, and unsecured loan were considered as input and investments, loans, and advances were outputs of the DEA model. Liu and Lyu [38] performed a dynamic network DEA model for evaluating the innovation efficiency of the pharmaceutical firms in China considering knowledge innovation and commercialization.

According to the above-mentioned studies, whilst the hybridization of DEA and TQM or DEA and KM has been used in the previous works, any research which simultaneously investigates TQM and KM factors in performance evaluation applying a DEA perspective was not observed. Also, among conducted research in the pharmaceutical companies, any research which considered TQM and KM in performance evaluation was not observed concurrently. Consequently, due to the existing research gap and the importance of this issue, in the present study, a DEA-based decision support framework has been presented to evaluate the performance of organizations based on the SF of TQM considering KM approach within pharmaceutical firms in Esfahan, Iran. The proposed decision support framework can provide invaluable suggestions to managers to improve their organizations' performance.

3. Concepts and Methods

In this section, the concepts and methods which have been applied in the present study are explained briefly.

3.1. TQM. TQM is a systematic structure that emphasizes the continuous improvement of all activities within an organization. In fact, TQM focuses on improving the quality of products and services by improving human resources, processes, and existing equipment, as well as reducing operating costs [39]. In other words, TQM emphasizes design, technology, and appropriate production processes selection, quality training, improving employees' involvement, considering customer requirements, and the necessity to measure the work. TQM is not just about the product; rather, it is a comprehensive view towards both the organization and the product and includes all the activities,

processes, and details of the work [40]. There are different definitions for TQM which are partly presented in Table 1.

3.2. KM. KM deals with the management of individual and organizational knowledge in organizations to gain CA. Knowledge is a source of CA because it provides intangible assets which are unique and cannot be imitated [45]. Achieving this CA depends on the organization's ability of effectively using existing KM for creating new knowledge assets and acting based on them. Although KM seems to be a business approach, every organization should develop its own strategies to acquire the potential values of KM [46]. KM has been defined from different perspectives which are partly presented in Table 2.

3.3. DEA. DEA is a nonparametric performance evaluation method that was firstly introduced by Charnes et al. [52]. In fact, they generalized Farrell's [53] single-input single-output technical efficiency measure to the multiple-input multiple-output case to evaluate the relative efficiency of peer units with respect to multiple performance measures [54–57]. The units under evaluation in DEA are called decision-making units (DMUs) and a DMU is considered to be efficient when no other DMU can produce more outputs using an equal or lower amount of inputs [58–60]. What makes the difference in DEA models is the concept of RTS. In the case of constant RTS, the CCR model will be used. Meanwhile, in terms of variable RTS, the BCC model should be used. Since recognizing the RTS is a long-term procedure, there is a need to collect data from different periods and analyze them [61], and regarding the limitations on data collection (we only were allowed to collect data in one period) we were not able to determine the RTS. Consequently, by consulting with related experts, both university experts and specialists in pharmaceutical companies, we assumed that RTS is variable. As a result, in this research, under this assumption that the RTS is variable, a BCC model is performed. It should be noticed that, in both input-oriented and output-oriented DEA models as well as envelopment and multiplier DEA models, efficient DMUs would be the same. In other words, there is no difference in the recognized efficient DMUs in all aforementioned modes of BCC model. Thus, we have used the envelopment input-oriented BCC model in this research. The applied model is as follows [62–64]:

$$
\begin{aligned}
&\text{Min}\theta \\
&s.t \\
&\sum_{j=1}^{n} \lambda_j x_{ij} \le \theta x_{ip} \quad i = 1, 2, \ldots m \\
&\sum_{j=1}^{n} \lambda_j y_{ij} \ge y_{rp} \quad i = 1, 2, \ldots m \\
&\sum_{j=1}^{n} \lambda_j = 1 \\
&\lambda_j \ge 0 \quad j = 1, 2, \ldots, n \\
&\theta \text{Free.}
\end{aligned}
\tag{1}
$$

4. Methodology

First of all, the SFs of TQM and KM are extracted from reputable scientific references and library studies and will be selected according to the opinions of university experts who are faculty members. Accordingly, academic experts selected the 14 factors identified by Talib et al. [65] as SFs of TQM and the 12 factors recognized by Valmohammadi and Ahmadi [66] as SFs of KM. In the next step, based on identified indicators in the previous step, questionnaires are distributed among experts of pharmaceutical companies to collect required data. In the DEA, if the number of DMUs is considerably less than the number of inputs and outputs, then the number of efficient units will increase and consequently the obtained efficiency scores will not be valid. Since, in this study, there are five DMUs (pharmaceutical companies) which are lower than the total number of inputs and outputs (12 factors for KM and 14 factors for TQM), therefore, an appropriate data mining method is employed to reduce the number of inputs and output criteria. Next, to calculate the efficiency scores for pharmaceutical companies, an appropriate DEA model (with respect to variable RTS) is applied and companies are categorized as efficient and inefficient companies. Also, complementary explanation and analysis are provided through two area charts. It should be mentioned that, among six pharmaceutical companies in Esfahan, Reyhaneh Pharmacy did not agree to share its data in this study. Therefore, this study was conducted using collected data from five other pharmaceutical companies in Esfahan.

The steps of this research are summarized as follows:

(1) Identifying the SFs of TQM and KM

The SFs of TQM and KM are identified according to a comprehensive review on related previous works and consultation with university experts.

(2) Measuring the importance of SFs of TQM and KM

The importance of the mentioned factors is measured using the relevant questions in the researcher-made questionnaire of TQM and the researcher-made questionnaire of KM which are filled by experts of pharmaceutical companies.

(3) Measuring the performance of SFs of TQM and KM

The performance of the mentioned factors is measured using the relevant questions in the researcher-made questionnaire of TQM and the researcher-made questionnaire of KM which are filled by experts of pharmaceutical companies.

(4) Clustering the key SFs of TQM and KM

For this purpose, the K-means algorithm and the SPSS software are applied.

(5) Calculating the relative efficiency of pharmaceutical companies

Clustered SFs of KM and TQM are considered as inputs and outputs of the BCC model and the relative efficiency is calculated for pharmaceutical companies.

Table 1: Definitions of the TQM.

Author(s)	Definition
Dubey and Gunasekaran [40]	TQM is an effective cost management system for continuous efforts to improve people at all levels.
Kanji et al. [41]	It is a process that constantly satisfies the customer's needs with a reasonable cost using each person's abilities.
Harnesk and Abrahamsson [42]	TQM is a higher credit symbol.
Taddese and Osada [43]	It is a philosophy which develops the market and improves the success of the company's business.
Kristianto et al. [44]	It is a strategy to increase customer satisfaction.

Table 2: Definitions of the KM.

Author(s)	Definition
Davenport [47]	It is an effort to discover the hidden assets in people's minds and turn them into the organizational assets, so that this asset be accessible for all employees to exploit it.
O'Dell et al. [48]	It is a systematic procedure to finding, comprehending, and applying knowledge for value creation.
Choy and Suk [49]	It is a framework for integrating interactions at all activity levels of an organization so that the firm can use its knowledge and, if it is required, gain new knowledge for value creation for its customers and stakeholders. This managerial structure integrates individuals, processes, and technology for sustainable performance development.
Nevo and Chan [50]	It is an integrated procedure to recognize, collect, retrieve, share, and measure an organization's information capital. The information assets may include databases, documents, mechanisms, methods, or even the experiences of managers.
Karkoulian et al. [51]	It is a way for establishing an organization whose members can gain, share, and create knowledge or employ it for decision-making.

(6) Providing managerial directions

This step is done by drawing two area charts for SFs of TQM and KM.

Figure 1 depicts the steps of the present study.

5. Data Collection and Analysis

In this section, the procedure of data collection, the selection of SFs of TQM and KM, the method of clustering SFs of TQM and KM, and ultimately the efficiency calculation of all pharmaceutical companies are elaborated in detail.

5.1. Identifying the SFs of TQM and KM. After a detailed literature review and with respect to the academic experts' suggestions, 14 SFs identified by Talib et al. [65] for TQM and 12 SFs recognized by Valmohammadi and Ahmadi [66] for KM were selected. Table 3 contains SFs of TQM identified by Talib et al. [65].

Also, Table 4 includes SFs of KM identified by Valmohammadi and Ahmadi [66].

5.2. Measuring the Importance of SFs of TQM and KM. Using a researcher-made questionnaire, including 71 questions related to the SFs of TQM, the importance of 14 identified SFs of TQM is measured by experts of pharmaceutical companies in Esfahan, including managers, supervisors, and specialists based on the five-point Likert scale. Finally, according to 14 questions related to the importance coefficient of these factors in organizations and using averaging, the importance of each of the SFs of TQM is determined for each firm. Table 5 contains the results of the completion of the questionnaire by 58 experts of pharmaceutical companies in Esfahan for the importance of the SFs of TQM.

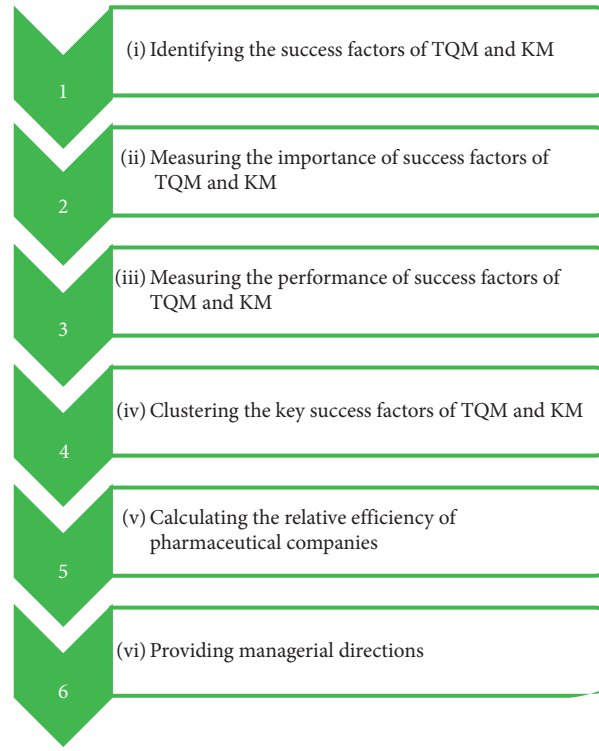


Figure 1: Research steps.

Also, through performing a researcher-made questionnaire, including 69 questions related to the SFs of KM, the

Table 3: SFs of TQM.

Code	Success factor
T1	Leadership and top-management support
T2	Organizational culture
T3	TQM strategy
T4	Benchmarking
T5	Organizational structure
T6	Process management, control, and improvement
T7	Data analysis
T8	Customer relationship management
T9	Supplier relationship management
T10	Reward and encouragement
T11	Resource innovation
T12	Training and education
T13	Continuous improvement
T14	Human resource management

Table 4: SFs of KM.

Code	Success factor
K1	Leadership and top-management support
K2	Organizational culture
K3	Information technology
K4	KM strategy
K5	Performance measurement
K6	Organizational infrastructure management
K7	Processes and activities
K8	Reward and encouragement
K9	Removing resource limitations
K10	Training and education
K11	Human resource management
K12	Benchmarking

Table 5: Importance values of SFs of TQM.

Code	Name of pharmaceutical company				
	Raha	Goldaru	Jey Pharmed Spandana	Amin	Farabi
T1	3.917	3.684	4.667	3.875	3.846
T2	3.667	3.921	3.417	3.938	3.846
T3	3.306	3.421	3.556	3.833	3.974
T4	3.750	3.579	4.000	3.250	4.538
T5	3.500	3.474	3.750	3.563	4.038
T6	3.467	3.263	3.567	3.450	4.200
T7	3.139	2.982	3.889	3.167	4.205
T8	3.583	3.526	3.833	3.250	3.846
T9	3.583	3.789	4.333	4.125	4.308
T10	3.467	3.589	3.833	3.525	4.376
T11	3.333	2.579	3.167	3.500	3.385
T12	3.167	3.211	3.167	3.375	3.692
T13	3.250	3.263	3.833	3.375	4.308
T14	3.750	3.579	4.000	3.250	4.538

importance of 12 identified SFs of KM is measured by experts of pharmaceutical companies in Esfahan, including managers, supervisors, and specialists based on the five-point Likert scale. Finally, according to 12 questions related to the importance coefficient of these factors in organizations and using averaging, the importance of each of the SFs of KM is determined for each firm. Table 6 depicts the results

Table 6: Importance values of SFs of KM.

Code	Name of pharmaceutical company				
	Raha	Goldaru	Jey Pharmed Spandana	Amin	Farabi
K1	3.917	3.737	4.500	3.857	3.923
K2	3.528	3.404	4.056	3.762	4.048
K3	3.500	3.526	4.333	3.429	4.143
K4	3.500	3.158	4.000	3.857	3.643
K5	3.583	3.421	3.833	3.571	4.429
K6	3.333	3.263	3.500	3.429	4.143
K7	3.125	3.421	3.500	3.500	3.964
K8	3.000	3.368	4.167	3.143	3.643
K9	3.083	3.368	3.833	3.571	4.214
K10	3.333	3.368	3.833	3.143	4.000
K11	3.125	3.316	4.000	3.500	4.143
K12	3.167	3.368	3.667	3.857	3.857

of the completion of the questionnaire by 58 experts of pharmaceutical companies in Esfahan for the importance of the SF of KM.

5.3. Measuring the Performance of SFs of TQM and KM. To measure the performance of each of the SFs of TQM, 57 remaining questions in the researcher-made questionnaire for TQM are answered by experts of pharmaceutical companies based on the five-point Likert scale. Consequently, according to these questions and using averaging, the performance of each of the SFs of TQM is determined for each firm. Table 7 illustrates the results of the completion of the questionnaire by 58 experts of pharmaceutical companies in Esfahan for the performance of the SFs of TQM.

Furthermore, to measure the performance of each of the SFs of KM, 57 remaining questions in the researcher-made questionnaire for KM are answered by experts of pharmaceutical companies based on the five-point Likert scale. Consequently, according to these questions and using averaging, the performance of each of the SFs of KM is determined for each firm. Table 8 reveals the results of the completion of the questionnaire by 58 experts of pharmaceutical companies in Esfahan for the performance of the SFs of KM.

5.4. Clustering the Key SFs of TQM and KM. In DEA models, when the number of DMUs is lower than the number of inputs and outputs noticeably, the number of efficient DMUs increases irrationally and therefore the results of the DEA model will not be reliable. In such a situation, it is necessary to somehow reduce the number of inputs and outputs. To do so, in the current study, 12 SFs of KM and 14 SFs of TQM, which are the inputs and outputs of the model, are clustered using the SPSS software and the K-means algorithm to decrease the number of inputs and outputs. During the execution of the algorithm, the factors of the same cluster will have the most similarity to each other while the factors in different clusters will not resemble.

For clustering, firstly, the number of clusters must be determined for the K-means algorithm. Hence, the SPSS software and the Hierarchical Clustering method are used.

Table 7: Performance values of SFs of TQM.

Code	Name of pharmaceutical company				
	Raha	Goldaru	Jey Pharmed Spandana	Amin	Farabi
T1	3.194	2.934	3.722	3.762	3.964
T2	2.885	2.789	3.271	3.232	3.786
T3	3.194	2.711	3.194	3.69	4.071
T4	2.854	2.566	2.917	3.429	3.629
T5	3.467	2.979	3.667	4.086	3.839
T6	3.144	2.719	3.370	3.603	4.548
T7	3.194	3.211	3.222	3.952	4.022
T8	2.813	3.276	2.667	3.536	4.357
T9	3.028	3.316	3.056	3.714	4.238
T10	2.686	2.184	2.625	3.071	3.971
T11	3.367	3.358	3.933	3.343	3.625
T12	3.167	2.684	3.611	3.500	3.690
T13	3.000	3.105	3.000	3.571	3.723
T14	2.569	2.281	3.222	2.81	3.661

Table 8: Performance values of SFs of KM.

Code	Name of pharmaceutical company				
	Raha	Goldaru	Jey Pharmed Spandana	Amin	Farabi
K1	3.190	2.677	3.690	2.714	3.796
K2	2.906	2.704	3.104	2.500	3.848
K3	3.125	2.82	3.333	2.714	4.095
K4	2.917	2.693	3.167	2.629	3.762
K5	2.625	2.776	2.542	2.314	3.929
K6	2.883	2.842	3.067	2.971	3.971
K7	3.092	2.539	3.250	2.343	4.314
K8	2.467	1.947	2.200	2.542	3.686
K9	2.983	3.211	3.367	2.833	3.600
K10	2.681	2.658	3.167	2.595	3.905
K11	2.792	2.421	3.333	2.839	4.268
K12	3.111	2.544	3.000	3.000	3.833

Figures 2 and 3 show the dendrogram diagram which determines the number of clusters of the SFs of TQM and KM, respectively.

In fact, after determining the number of clusters using the Hierarchical Clustering method in the SPSS software, the K-means algorithm is executed and the main SFs of TQM and the main SFs of KM are clustered into three clusters. Thus, with the implementation of this algorithm, firstly, some points are selected as the initial centralization in each dimension for TQM and KM randomly. The number of these points is selected according to the number of clusters that were determined in the previous step by the Hierarchical Clustering method, which were three clusters. Then, the closest points to each center form a cluster, and the algorithm will be repeated for other dimensions. Besides, each time the algorithm is repeated, the centrality of each cluster is iterated by averaging the factors in each cluster, and the algorithm is replicated until placing all the factors in the appropriate clusters according to their similarity. Tables 9 and 10 show the clustering results of the SFs of TQM and KM, respectively.

Ultimately, the average score of the factors related to each cluster indicates the final score of the cluster. Tables 11 and 12 show these values for SFs of TQM and SFs of KM in each organization, respectively.

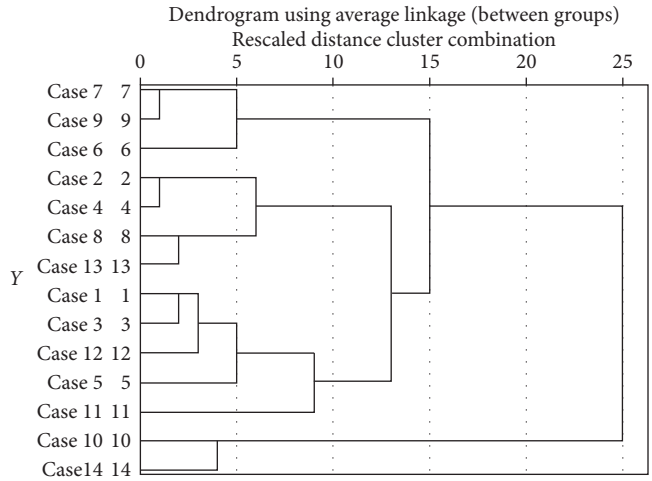


Figure 2: Dendrogram diagram for determining the number of clusters of TQM success factors.

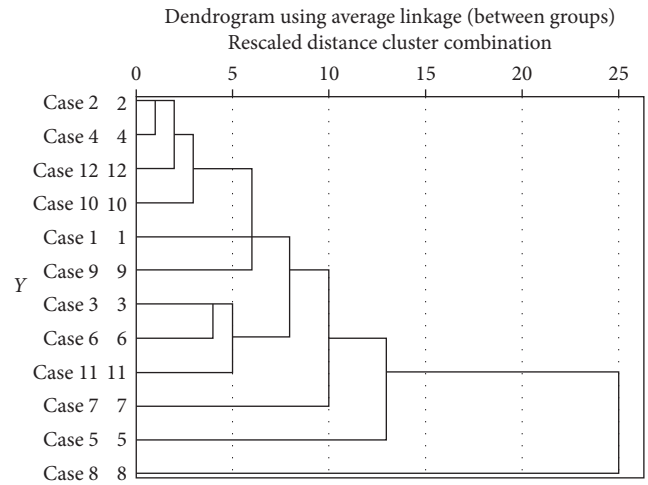


Figure 3: Dendrogram diagram for determining the number of clusters of KM success factors.

5.5. Calculating the Relative Efficiency of Pharmaceutical Companies. To calculate the efficiency of pharmaceutical companies in Esfahan, the BCC input-oriented model is used. The clustered data of KM are considered as input of the model (three inputs) and the average of the clustered data of TQM is considered as the output of the model (one output) and the efficiency values for DMUs (pharmaceutical companies) are obtained. The inputs and outputs of the model as well as the calculated efficiency for each DMU applying the BCC model are presented in Table 13.

It should be noted that organizations with an efficiency score of 1 have the best performance and will be efficient. Accordingly, organizations are divided into two categories, namely, efficient companies and inefficient companies.

Table 9: Clustering of SFs of TQM.

Code	Name of pharmaceutical company					Cluster number
	Raha	Goldaru	Jey Pharmed Spandana	Amin	Farabi	
	Performance values of success factors of TQM					
T1	3.194	2.934	3.722	3.762	3.964	3
T2	2.885	2.789	3.271	3.232	3.786	1
T3	3.194	2.711	3.194	3.69	4.071	2
T4	2.854	2.566	2.917	3.429	3.629	1
T5	3.467	2.979	3.667	4.086	3.839	3
T6	3.144	2.719	3.37	3.603	4.548	2
T7	3.194	3.211	3.222	3.952	4.022	2
T8	2.813	3.276	2.667	3.536	4.357	2
T9	3.028	3.316	3.056	3.714	4.238	2
T10	2.686	2.184	2.625	3.071	3.971	1
T11	3.367	3.358	3.933	3.343	3.625	3
T12	3.167	2.684	3.611	3.5	3.69	3
T13	3	3.105	3	3.571	3.723	2
T14	2.569	2.281	3.222	2.81	3.661	1

Table 10: Clustering of SFs of KM.

Code	Name of pharmaceutical company					Cluster number
	Raha	Goldaru	Jey Pharmed Spandana	Amin	Farabi	
	Performance values of success factors of KM					
K1	3.190	2.677	3.690	2.714	3.796	3
K2	2.906	2.704	3.104	2.500	3.848	2
K3	3.125	2.82	3.333	2.714	4.095	3
K4	2.917	2.693	3.167	2.629	3.762	2
K5	2.625	2.776	2.542	2.314	3.929	2
K6	2.883	2.842	3.067	2.971	3.971	3
K7	3.092	2.539	3.250	2.343	4.314	3
K8	2.467	1.947	2.200	2.542	3.686	1
K9	2.983	3.211	3.367	2.833	3.600	3
K10	2.681	2.658	3.167	2.595	3.905	2
K11	2.792	2.421	3.333	2.839	4.268	3
K12	3.111	2.544	3.000	3.000	3.833	2

Table 11: Average of each cluster for SFs of TQM.

Name of pharmaceutical company	Cluster 1 T2, T4, T10, and T14	Cluster 2 T3, T6, T7, T8, T9, and T13	Cluster 3 T1, T5, T11, and T12
Raha	2.749	3.065	3.299
Goldaru	2.749	2.868	2.99
Jey Pharmed Spandana	3.336	3.014	3.598
Amin	3.203	3.571	3.82
Farabi	3.735	4.055	3.833

Table 12: Average of each cluster for SFs of KM.

Name of pharmaceutical company	Cluster 1 K8	Cluster 2 K2, K4, K5, K10, and K12	Cluster 3 K1, K3, K6, K7, K9, and K11
Raha	3.011	2.848	2.467
Goldaru	2.750	2.675	1.947
Jey Pharmed Spandana	3.456	2.996	2.200
Amin	2.673	2.849	2.314
Farabi	4.007	3.855	4.296

Table 13: Efficiency scores of the BCC model.

Name of pharmaceutical company (DMU)	Input (KM)			Output (TQM)	Efficiency score
Raha	3.011	2.848	2.467	3.037	0.955
Goldaru	2.75	2.675	1.947	2.869	1
Jey Pharmed Spandana	3.456	2.996	2.2	3.316	0.998
Amin	2.673	2.849	2.314	3.531	1
Farabi	4.007	3.855	4.296	3.874	1

5.6. Managerial Directions. In this section, with the aim of explaining how the proposed decision support framework of the current paper can be helpful for managers of pharmaceutical companies, some recommendations are provided. To do so, based on measuring the value of importance and performance and the average score of SFs of TQM with KM approach, two area charts are drawn and analyzed.

In the area chart of TQM, the vertical axis is the importance average of the SFs of TQM for the five pharmaceutical companies in Esfahan (average values of each row of Table 5) and the horizontal axis is the performance average of the SFs of TQM for the aforementioned pharmaceutical companies (average values of each row in Table 7). Table 14 contains the required data for drawing the area chart of TQM. In addition, Figure 4 presents the area chart of TQM.

In the same way, the area chart of KM is drawn for the SFs of KM. In the area chart of KM, the vertical axis is the importance average of the SFs of KM for the five pharmaceutical companies in Esfahan (average values of each row of Table 6) and the horizontal axis is the performance average of the SFs of KM for the aforementioned pharmaceutical companies (average values of each row in Table 8). Table 15 contains the required data for drawing the area chart of KM. In addition, Figure 5 presents the area chart of KM.

As it can be seen in the area chart of TQM, SFs of TQM in the five pharmaceutical companies in Esfahan are located in the four areas of Figure 4.

Area 1: area chart of TQM.

Four factors including organizational culture, benchmarking, supplier relationship management, and human resource management are located in area 1. In this area, the performance values of these factors within the five pharmaceutical companies in Esfahan are lower than average whereas their importance values are greater than average. Therefore, considering the importance of TQM, five pharmaceutical companies should pay more attention to these factors. In other words, these factors have a high priority for immediate improvement.

Area 2: area chart of TQM.

Three factors including leadership and top-management support, organizational structure, and reward and encouragement are located in area 2. In this area, both performance values and importance values of these factors within the five pharmaceutical companies in Esfahan are higher than average. Hence, these factors are in a desirable situation in pharmaceutical companies compared to other factors. In fact, these factors provide a CA for pharmaceutical companies, and the right solution is to continue to act in the same way as before.

Table 14: Required data of area chart of TQM.

Code	Average of importance values	Average of performance values
T1	3.998	3.455
T2	3.758	3.180
T3	3.618	3.296
T4	3.823	3.078
T5	4.040	3.641
T6	3.589	3.498
T7	3.476	3.587
T8	3.608	3.263
T9	4.028	2.950
T10	3.758	3.390
T11	3.193	3.531
T12	3.322	3.406
T13	3.606	3.292
T14	3.823	2.969
Average	3.689	3.324

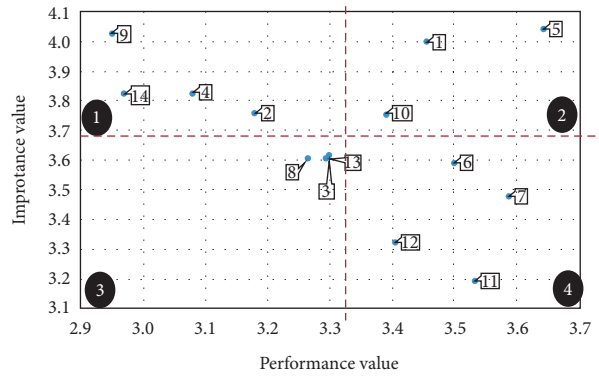


Figure 4: Area chart of TQM.

Area 3: area chart of TQM.

Three factors including TQM strategy, customer relationship management, and continuous improvement are located in area 3. In this area, both the performance values and the importance values of these factors are lower than average. Accordingly, these factors do not require any additional investment, since these factors are neither critical nor threatening for the five pharmaceutical companies.

Table 15: Required data of area chart of KM.

Code	Average of importance values	Average of performance values
K1	3.987	3.271
K2	3.759	3.080
K3	3.786	3.194
K4	3.632	3.074
K5	3.768	3.003
K6	3.534	3.160
K7	3.502	3.233
K8	3.464	2.645
K9	3.614	3.158
K10	3.536	3.027
K11	3.617	3.065
K12	3.583	3.041
Average	3.648	3.079

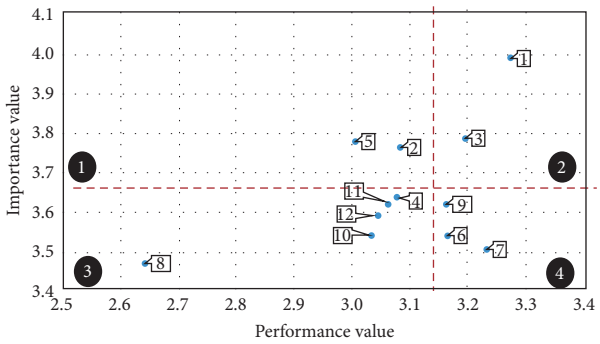


Figure 5: Area chart of KM.

Area 4: area chart of TQM.

Four factors including process management, control, and improvement, data analysis, resource innovation, and training and education are located in area 4. In this area, the performance values of the SFs of TQM within the five pharmaceutical companies in Esfahan are higher than the average while the importance values are lower than the average. Thus, pharmaceutical companies should modify the allocated resources to these factors and focus on other factors, especially those located in area 1.

Similar to the area chart of TQM, SFs of KM in the five pharmaceutical companies in Esfahan are located in the four areas of Figure 5 as well.

Area 1: area chart of KM.

Two factors including organizational culture and performance measurement are located in area 1. In this area, the performance values of these factors within the five pharmaceutical companies in Esfahan are lower than average whereas their importance values are greater than average. Therefore, considering the importance of KM, five pharmaceutical companies should pay more attention to these factors. In other words, these factors have a high priority for immediate improvement.

Area 2: area chart of KM.

Two factors including leadership and top-management support and information technology are located in area 2. In this area, both performance values and importance values of these factors within the five pharmaceutical companies in Esfahan are higher than average. Hence, these factors are in a desirable situation in pharmaceutical companies compared to other factors. In fact, these factors provide a CA for pharmaceutical companies, and the right solution is to continue to act in the same way as before.

Area 3: area chart of KM.

Five factors including KM strategy, reward and encouragement, training and education, human resource management, and benchmarking are located in area 3. In this area, both the performance values and the importance values of these factors are lower than average. Accordingly, these factors do not require any additional investment, since these factors are neither critical nor threatening for the five pharmaceutical companies.

Area 4: area chart of KM.

Three factors including organizational infrastructure management, processes and activities, and removing resource limitations are located in area 4. In this area, the performance values of the SFs of KM within the five pharmaceutical companies in Esfahan are higher than the average while the importance values are lower than the average. Thus, pharmaceutical companies should modify the allocated resources to these factors and focus on other factors, especially those located in area 1.

6. Concluding Remarks and Future Research

Simultaneous investigation of TQM and KM in the performance of organizations can help organizations in the continuous improvement and benefiting from the CA as much as possible, because this concurrent investigation compares organizations from TQM and KM perspectives and determines the aspects that each organization can focus on to achieve better results. Accordingly, in a long-term horizon, organizations will be able to institutionalize TQM and KM as a culture in their organization and exploit its advantages in improving the organization's performance.

In this study, we evaluated the performance of pharmaceutical companies in Esfahan considering TQM and KM using a DEA model. Initially, the SFs of TQM and KM were selected by reviewing previous studies and consulting academic experts. Then, two researcher-made questionnaires were distributed among the experts of pharmaceutical companies in Esfahan to examine the SFs of TQM and KM, respectively. Within these researcher-made questionnaires, the importance values of the SFs of TQM and KM were determined. Next, in order to evaluate the performance of organizations, the BCC model was used, so that the clustered SFs of KM and TQM were considered as inputs and outputs of the model, respectively. Among the five pharmaceutical

companies in Esfahan, three companies, namely, Amin, Farabi, and Goldaru, were determined as efficient pharmaceutical companies.

In addition, in order to provide more insight towards the organization for managers, two area charts were presented for SFs of TQM and KM, respectively, and results were analyzed and discussed. Accordingly, the factors which were located in area 1 need to be improved instantly. In this group, "organizational culture," "benchmarking," "supplier relationship management," and "human resource management" as the four SFs of TQM and "organizational culture" and "performance measurement" as the two SFs of KM need immediate improvement. Also, factors which were located in area 2 were in the favorable situation, and organizations need to continue applying the current procedure about these factors. In this regard, "leadership and top-management support," "organizational structure," and "reward and encouragement" as the three SFs of TQM and "leadership and top-management support" and "information technology" as the two SFs of KM should be retained in the current situations. Further, factors in area 3 including "TQM strategy," "customer relationship management," and "continuous improvement" as the three SFs for TQM and KM strategy and "reward and encouragement," "training and education," "human resource management," and "benchmarking" as the five SFs of KM are factors that are not a threat to organizations and do not need to be considered. However, "process management, control, and improvement," "data analysis," "resource innovation," and "training and education" as the four SFs of TQM and "organizational infrastructure management," "processes and activities," and "removing resource limitations" as the three SFs of KM were located in area 4. In consequence, these SFs should be reconsidered and their resources should be reassigned to the factors in area 1. Therefore, in selecting managers of different departments in organizations, it should be taken into account whether the department needs to improve or should it retain the current situation.

It is noteworthy that, among the SFs of TQM and KM in the five pharmaceutical companies in Esfahan, Jey Pharmed Spandana, and Farabi have had the highest performance in the factor of leadership and top-management support. Given that the success of any project in the organization requires the participation and support of senior managers of that organization, the style of leadership and top-management support in these two companies can be considered as a benchmark, so that if other companies implement this style within their organization, they will observe an improvement in their performance.

Apart from advantageous aspects of this research which have been highlighted within the text in detail, similar to other studies, there are limitations as well. Considering that this research has been conducted in the city of Esfahan, as one of the most developed regions of Iran, the results have been obtained according to the specific culture and social conditions of this city. However, since pharmaceutical companies are expanded across the country, caution should be exercised in generalizing the results of this research to pharmaceutical companies which are located in other cities and regions that are less developed. Also, the research findings are limited to the case study (pharmaceutical companies in Esfahan) and by changing the case, probably different results are obtained. Therefore, it is not possible to generalize the findings to other organizations. Moreover, it should be noted that, in addition to the SFs of TQM and KM which were used in this study, other factors may affect the success of TQM and KM that were not considered in this study. Another limitation of this study is that the validity of the research findings depends on the time of data collection. Therefore, if the data is collected at other intervals, it is possible that the results will change.

Because, in this study, only the questionnaire tool was used to collect data, researchers in future studies can apply different methods of interview to collect data. Performance evaluation based on the other SFs of TQM and KM can also be considered in future research. On the other hand, due to the fact that the K-means algorithm has been used in the present study, in future research, other clustering algorithms or other statistical methods can be employed to determine the relationship between variables and select input and output variables. Since the benefits of TQM and KM are not limited to the pharmaceutical companies, effects of these approaches on performance of other industries and companies could be addressed in future research. In fact, by using SFs of TQM and KM as the inputs and outputs of the DEA model, other firms and industries can enjoy our proposed framework for performance evaluation and analyze the effects of TQM and KM on their performances. Furthermore, because the data of this study were collected in a specific time period, data collection at different intervals and using other performance evaluation methods such as dynamic DEA or the Malmquist index method which consider different time intervals in data analysis could be addressed in future research.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

[1] Z. Danial, "Effect of total quality management in determining the educational needs of critical wards nurses," *Iranian Journal of Critical Care Nursing*, vol. 2, no. 3, pp. 117–120, 2009.

[2] D. I. Prajogo and A. S. Sohal, "The integration of TQM and technology/R&D management in determining quality and innovation performance," *Omega*, vol. 34, no. 3, pp. 296–312, 2006.

[3] K. B. Hendricks and V. R. Singhal, "Firm characteristics, total quality management, and financial performance," *Journal of Operations Management*, vol. 19, no. 3, pp. 269–285, 2001.

[4] A. Jantunen, "Knowledge-processing capabilities and innovative performance: an empirical study," *European Journal of Innovation Management*, vol. 8, no. 3, pp. 336–349, 2005.

[5] C. Garavelli, M. Gorgoglione, and B. Scozzi, "Knowledge management strategy and organization: a perspective of analysis," *Knowledge and Process Management*, vol. 11, no. 4, pp. 273–282, 2004.

[6] M. Asadpour, O. Boyer, and R. Tavakkoli-Moghaddam, "A blood supply chain network with backup facilities considering blood groups and expiration date: a real-world application," *International Journal of Engineering*, vol. 34, no. 2, pp. 470–479, 2021.

[7] I. Bahrami, R. M. Ahari, and M. Asadpour, "A maximal covering facility location model for emergency services within an $M(t)$/M/m/m queuing system," *Journal of Modelling in Management*, 2020, in Press.

[8] A. A. Rajaee, A. Amindoust, and M. Asadpour, "Applying simulated annealing algorithm for parallel machine tardiness problem subject to job splitting," in *Proceedings of the International Conference on Computers &Industrial Engineering (CIE48)*, Auckland, New Zealand, December 2018.

[9] R. Jafarpisheh, M. Karbasian, and M. Asadpour, "A hybrid reliability-centered maintenance approach for mining transportation machines: a real case in Esfahan," *International Journal of Quality & Reliability Management*, 2020, in Press.

[10] I. Ajripour, M. Asadpour, and L. Tabatabaie, "A model for organization performance management applying MCDM and BSC: a case study," *Journal of Applied Research on Industrial Engineering*, vol. 6, no. 1, pp. 52–70, 2019.

[11] S.-H. Hsu and H.-P. Shen, "Knowledge management and its relationship with TQM," *Total Quality Management & Business Excellence*, vol. 16, no. 3, pp. 351–361, 2005.

[12] L. M. Molina, J. Lloréns-Montes, and A. Ruiz-Moreno, "Relationship between quality management practices and knowledge transfer," *Journal of Operations Management*, vol. 25, no. 3, pp. 682–701, 2007.

[13] A. Y.-L. Chong, K.-B. Ooi, B. Lin, and P.-L. Teh, "TQM, knowledge management and collaborative commerce adoption: a literature review and research framework," *Total Quality Management & Business Excellence*, vol. 21, no. 5, pp. 457–473, 2010.

[14] S. Daud and W. F. W. Yusoff, "The influence of soft and hard TQM factors on knowledge management: perspective from Malaysia," in *Proceedings of the International Conference on Management and Service Science*, vol. 8, pp. 17–22, Bangkok, Thailand, May 2011.

[15] K.-B. Ooi, "TQM: a facilitator to enhance knowledge management? A structural analysis," *Expert Systems with Applications*, vol. 41, no. 11, pp. 5167–5179, 2014.

[16] H. A. M. Alshatnawi and M. B. Abd Ghani, *The Effect of Total Quality Management and Knowledge Management on Organizational Performance in Higher Education Institution in Jordan*, USIM NILAI, Nilai, Malaysia, 2018.

[17] D. Obeidat, B. Yousef, L. Hashem, and R. E. Masa'deh, "The influence of knowledge management uses on total quality management practices in commercial banks of Jordan," *Modern Applied Science*, vol. 12, no. 11, 2018.

[18] Z. Alkhazali, I. Aldabbagh, and A. Abu-Rumman, "TQM potential moderating role to the relationship between HRM practices, KM strategies and organizational performance: the case of Jordanian banks," *Academy of Strategic Management Journal*, vol. 18, no. 3, pp. 1–16, 2019.

[19] J. Abbas, "Impact of total quality management on corporate sustainability through the mediating effect of knowledge management," *Journal of Cleaner Production*, vol. 244, Article ID 118806, 2020.

[20] R. Y. Y. Hung, B. Yang, B. Y.-H. Lien, G. N. McLean, and Y.-M. Kuo, "Dynamic capability: impact of process alignment and organizational learning culture on performance," *Journal of World Business*, vol. 45, no. 3, pp. 285–294, 2010.

[21] A. Honarpour, A. Jusoh, and K. Md Nor, "Knowledge Management,Total quality management and innovation: a new look," *Journal of Technology Management & Innovation*, vol. 7, no. 3, pp. 22–31, 2012.

[22] A. Honarpour, A. Jusoh, and K. Md Nor, "Total quality management, knowledge management, and innovation: an empirical study in R&D units," *Total Quality Management & Business Excellence*, vol. 29, no. 7-8, pp. 798–816, 2018.

[23] C. T. Kuah and K. Y. Wong, "Data Envelopment Analysis modeling for measuring knowledge management performance in Malaysian higher educational institutions," *Information Development*, vol. 29, no. 3, pp. 200–216, 2013.

[24] H. Shirouyehzad, F. M. Rafiee, and N. Berjis, "Performance evaluation and prioritization of organizations based on knowledge management and safety management approaches using DEA," *Journal of Modelling in Management*, vol. 12, no. 1, pp. 77–95, 2017.

[25] L. Salhieh and J. Abu-Doleh, "The relationship between total quality management practices and their effects on bank's technical efficiency," *International Journal of Commerce and Management*, vol. 25, no. 2, 2015.

[26] A. Hadi-Vencheh and A. Yousefi, "Selecting six sigma project: a comparative study of DEA and LDA techniques," *International Journal of Lean Six Sigma*, vol. 10, no. 2, 2018.

[27] Y. Wen, Q. An, X. Xu, and Y. Chen, "Selection of Six Sigma project with interval data: common weight DEA model," *Kybernetes*, vol. 47, no. 7, pp. 1307–1324, 2018.

[28] A. Azadeh, B. Nasirian, V. Salehi, and H. Kouzehchi, "Integration of PCA and DEA for identifying and improving the impact of Six Sigma implementation on job characteristics in an automotive industry," *Quality Engineering*, vol. 29, no. 2, pp. 273–290, 2017.

[29] G. B. Kamath, "Intellectual capital and corporate performance in Indian pharmaceutical industry," *Journal of Intellectual Capital*, vol. 9, no. 4, pp. 684–704, 2008.

[30] A. A. A. Sharabati, S. N. Jawad, and N. Bontis, "Intellectual capital and business performance in the pharmaceutical sector of Jordan," *Management Decision*, vol. 48, no. 1, pp. 105–131, 2010.

[31] G. Mehralian, A. Rajabzadeh, M. R. Sadeh, and H. R. Rasekh, "Intellectual capital and corporate performance in Iranian pharmaceutical industry," *Journal of Intellectual Capital*, vol. 13, no. 1, pp. 138–158, 2012.

[32] C. I. Enekwe, C. I. Agu, and K. N. Eziedo, "The effect of financial leverage on financial performance: evidence of quoted pharmaceutical companies in Nigeria," *Journal of Economics and Finance*, vol. 5, no. 3, pp. 17–25, 2014.

[33] M. Tavana, K. Khalili-Damghani, and R. Rahmatian, "A hybrid fuzzy MCDM method for measuring the performance of publicly held pharmaceutical companies," *Annals of Operations Research*, vol. 226, no. 1, pp. 589–621, 2015.

[34] M. Varmaghani, A. H. Meshkini, F. Farzadfar et al., "Evaluation of productivity in Iranian pharmaceutical companies: a DEA-based Malmquist approach and panel data analysis," *Journal of Research in Pharmacy Practice*, vol. 4, no. 2, p. 51, 2015.

[35] F. Gascón, J. Lozano, B. Ponte, and D. de la Fuente, "Measuring the efficiency of large pharmaceutical companies: an

industry analysis," *The European Journal of Health Economics*, vol. 18, no. 5, pp. 587–608, 2017.

[36] A. Al-Refaie, C.-W. Wu, and M. Sawalheh, "DEA window analysis for assessing efficiency of blistering process in a pharmaceutical industry," *Neural Computing and Applications*, vol. 31, no. 8, pp. 3703–3717, 2019.

[37] T. Alam and R. Rastgi, "Trend analysis of cost efficiency for the pharmaceutical industry: a DEA approach," *Management Science Letters*, vol. 9, no. 5, pp. 749–754, 2019.

[38] Z. Liu and J. Lyu, "Measuring the innovation efficiency of the Chinese pharmaceutical industry based on a dynamic network DEA model," *Applied Economics Letters*, vol. 27, no. 1, pp. 35–40, 2020.

[39] J. W. Barrow, "Does total quality management equal organizational learning," *Quality Progress*, vol. 26, no. 7, pp. 39–43, 1993.

[40] R. Dubey and A. Gunasekaran, "Exploring soft TQM dimensions and their impact on firm performance: some exploratory empirical results," *International Journal of Production Research*, vol. 53, no. 2, pp. 371–382, 2015.

[41] G. K. Kanji, A. Malek, and B. A. Tambi, "Total quality management in UK higher education institutions," *Total Quality Management*, vol. 10, no. 1, pp. 129–153, 1999.

[42] R. Harnesk and L. Abrahamsson, "TQM: an act of balance between contradictions," *The TQM Magazine*, vol. 19, no. 6, 2007.

[43] F. Taddese and H. Osada, "Market development using TQM through system integration and customer service in Asia," *Asian Journal on Quality*, vol. 12, no. 3, 2011.

[44] Y. Kristianto, M. M. Ajmal, and M. Sandhu, "Adopting TQM approach to achieve customer satisfaction," *The TQM Journal*, vol. 24, no. 1, 2012.

[45] R. M. Grant, "Toward a knowledge-based theory of the firm," *Strategic Management Journal*, vol. 17, no. S2, pp. 109–122, 1996.

[46] Y. C. Hung, S. M. Huang, and Q. P. Lin, "Critical factors in adopting a knowledge management system for the pharmaceutical industry," *Industrial Management & Data Systems*, vol. 105, no. 2, pp. 164–183, 2005.

[47] T. H. Davenport, "Some principles of knowledge management," *Strategy & Business*, vol. 1, no. 2, pp. 34–40, 1996.

[48] C. S. O'Dell, C. O'dell, C. J. Grayson, and N. Essaides, *If Only We Knew what We Know: The Transfer of Internal Knowledge and Best Practice*, Simon and Schuster, New York, NY, USA, 1998.

[49] C. S. Choy and C. Y. Suk, "Critical factors in the successful implementation of knowledge management," *Journal of Knowledge Management Practice*, vol. 6, no. 1, pp. 234–258, 2005.

[50] D. Nevo and Y. E. Chan, "A Delphi study of knowledge management systems: scope and requirements," *Information & Management*, vol. 44, no. 6, pp. 583–597, 2007.

[51] S. Karkoulian, L. A. Halawi, and R. V. McCarthy, "Knowledge management formal and informal mentoring," *The Learning Organization*, vol. 15, no. 5, pp. 409–420, 2008.

[52] A. Charnes, W. W. Cooper, and E. Rhodes, "Measuring the efficiency of decision making units," *European Journal of Operational Research*, vol. 2, no. 6, pp. 429–444, 1978.

[53] M. J. Farrell, "The measurement of productive efficiency," *Journal of the Royal Statistical Society. Series A (General)*, vol. 120, no. 3, pp. 253–290, 1957.

[54] A. Charnes, W. W. Cooper, A. Y. Lewin, and L. M. Seiford, *Data Envelopment Analysis: Theory, Methodology, and Application*, Kluwer Academic Publishers, Norwell, MA, USA, 1994.

[55] W. W. Cooper, L. M. Seiford, and K. Tone, "Data envelopment analysis: a comprehensive text with models, applications, references and DEA-solver software," *Journal-operational Research Society*, vol. 52, no. 12, pp. 1408-1409, 2001.

[56] F. H. Lotfi, A. Ebrahimnejad, M. Vaez-Ghasemi, and Z. Moghaddas, *Data Envelopment Analysis with R*, Springer International Publishing, Berlin, Germany, 2020.

[57] A. Ebrahimnejad and N. Amani, "Fuzzy data envelopment analysis in the presence of undesirable outputs with ideal points," *Complex & Intelligent Systems*, vol. 7, pp. 379–400, 2021.

[58] A. Ebrahimnejad, M. Tavana, F. H. Lotfi, R. Shahverdi, and M. Yousefpour, "A three-stage data envelopment analysis model with application to banking industry," *Measurement*, vol. 49, pp. 308–319, 2014.

[59] A. Ebrahimnejad, M. Tavana, S. H. Nasseri, and O. Gholami, "A new method for solving dual DEA problems with fuzzy stochastic data," *International Journal of Information Technology & Decision Making*, vol. 18, no. 1, pp. 147–170, 2019.

[60] S. Maleki, A. Ebrahimnejad, and R. Kazemi Matin, "Pareto–Koopmans efficiency in two-stage network data envelopment analysis in the presence of undesirable intermediate products and nondiscretionary factors," *Expert Systems*, vol. 36, no. 3, Article ID e12393, 2019.

[61] N. Amani, H. B. Valami, and A. Ebrahimnejad, "Application of Malmquist productivity index with carry-overs in power industry," *Alexandria Engineering Journal*, vol. 57, no. 4, pp. 3151–3165, 2018.

[62] R. D. Banker, A. Charnes, and W. W. Cooper, "Some models for estimating technical and scale inefficiencies in data envelopment analysis," *Management Science*, vol. 30, no. 9, pp. 1078–1092, 1984.

[63] M. Asadpour and H. Shirouyehzad, "Performance evaluation and ranking of academy award winners for best original score applying data envelopment analysis: 1990–2016," *Operations Research Letters*, vol. 47, no. 5, pp. 371–376, 2019.

[64] K. Rahimpour, H. Shirouyehzad, M. Asadpour, and M. Karbasian, "A PCA-DEA method for organizational performance evaluation based on intellectual capital and employee loyalty," *Journal of Modelling in Management*, vol. 15, no. 4, pp. 1479–1513, 2020.

[65] F. Talib, Z. Rahman, and M. N. Qureshi, "The relationship between total quality management and quality performance in the service industry: a theoretical model," *International Journal of Business, Management and Social Sciences (IJBMSS)*, vol. 1, no. 1, pp. 113–128, 2010.

[66] C. Valmohammadi and M. Ahmadi, "The impact of knowledge management practices on organizational performance," *Journal of Enterprise Information Management*, vol. 28, no. 1, 2015.

Hindawi
Journal of Healthcare Engineering
Volume 2021, Article ID 5594649, 9 pages
https://doi.org/10.1155/2021/5594649

Research Article

DL-MRI: A Unified Framework of Deep Learning-Based MRI Super Resolution

Huanyu Liu,[1,2] Jiaqi Liu,[1,2] Junbao Li,[1,2] Jeng-Shyang Pan,[3] and Xiaqiong Yu[4]

[1]*School of Electronics and Information Engineering, Harbin Institute of Technology, Harbin 150001, China*
[2]*Center of AI Perception, AI Research Institute, Harbin Institute of Technology, Harbin 150001, China*
[3]*College of Computer Science and Engineering, Shandong University of Science and Technology, Qingdao 266590, China*
[4]*32021 Troops of the PLA, Beijing 100094, China*

Correspondence should be addressed to Junbao Li; lijunbao@hit.edu.cn

Received 25 February 2021; Revised 22 March 2021; Accepted 30 March 2021; Published 9 April 2021

Academic Editor: Hao Chun Lu

Magnetic resonance imaging (MRI) is widely used in the detection and diagnosis of diseases. High-resolution MR images will help doctors to locate lesions and diagnose diseases. However, the acquisition of high-resolution MR images requires high magnetic field intensity and long scanning time, which will bring discomfort to patients and easily introduce motion artifacts, resulting in image quality degradation. Therefore, the resolution of hardware imaging has reached its limit. Based on this situation, a unified framework based on deep learning super resolution is proposed to transfer state-of-the-art deep learning methods of natural images to MRI super resolution. Compared with the traditional image super-resolution method, the deep learning super-resolution method has stronger feature extraction and characterization ability, can learn prior knowledge from a large number of sample data, and has a more stable and excellent image reconstruction effect. We propose a unified framework of deep learning -based MRI super resolution, which has five current deep learning methods with the best super-resolution effect. In addition, a high-low resolution MR image dataset with the scales of ×2, ×3, and ×4 was constructed, covering 4 parts of the skull, knee, breast, and head and neck. Experimental results show that the proposed unified framework of deep learning super resolution has a better reconstruction effect on the data than traditional methods and provides a standard dataset and experimental benchmark for the application of deep learning super resolution in MR images.

1. Introduction

MRI [1] is a type of tomography, which uses magnetic resonance to obtain electromagnetic signals from the human organ tissue and reconstruct human information about the structure of human organs. MRI has been used in imaging diagnosis of various systems throughout the body, including craniocerebral [2], spinal cord [3], large heart blood vessels [4], joint bones [5], and soft tissues and pelvis [6]. High-resolution MR images have richer structural details, which is helpful for doctors to locate the lesions and diagnose the disease. The acquisition of MR high-resolution image needs to increase the magnetic field intensity and pulse radiation time, high intensity magnetic field, and ultra-long pulse radiation not only bring bad experience to patients but also produce image artifacts due to patients' movement, thus affecting the quality of imaging. Therefore, the software method for MR image super resolution has great significance, and it can improve the MR image resolution without causing harm to the human body.

Super resolution is to restore high-frequency detail from low-resolution image to improve image resolution. The simplest method is to use interpolation to improve the image resolution, such as bicubic interpolation [7] and nearest neighbor interpolation [8]. However, the interpolation method is not in essence to increase the image information, so it cannot recover the image high-frequency information. Subsequently, people put forward the super-resolution method based on spatial domain constraints because this algorithm has wide applicability and strong prior constraint capability. In this

respect, algorithms include reverse iterative projection method [9–11], statistical method based on maximum posterior probability [12–14], nonuniform sample interpolation method [15, 16], and convex set projection method [17]. These algorithms make full use of spatial correlation of data, introduce fuzzy, point diffusion, and other degradation parameters into image degradation model, and cover global motion and local motion. However, when dealing with the super resolution of multiframe images, the quality of reconstructed images declines rapidly as the super-resolution ratio increases or the number of available input images decreases. Under such circumstances, the reconstructed results will appear too smooth and lack important high-frequency details.

The learn-based super-resolution method is different from the image priori knowledge acquired under various assumptions. It directly learns the image priori knowledge from a large number of samples. The learning-based image super-resolution methods mainly include dictionary learning [18] and deep learning [19]. In dictionary learning, a high-resolution image block can utilize a high-resolution dictionary sparse representation, and the sparse coefficient can be obtained from the sparse representation of the low-resolution dictionary of the corresponding low-resolution image block. However, the super-resolution method based on sparse representation needs to solve the sparse representation of the super-complete dictionary. When the dictionary size or reconstructed image size is large, the computational time complexity is still very high, which still has a considerable gap with the real-time application. The super-resolution method based on deep learning has been validated in natural images. The network used by NISR mainly includes three categories: feedforward deep networks, feedback deep networks, and generative adversarial networks. DBPN [20] belongs to the feedforward deep network, which provides an error feedback and interdependent modules, which represent image degradation and high-resolution, making characteristics of the sampling phase are connected to improve the SR result. RDN [21] belongs to the feedback deep network. It proposes a residual dense block (RDB) and uses the densely connected convolutional layer to extract rich local features. SRFBN [22] belongs to the feedback deep network. This type of model implements this feedback method by using hidden states in constrained RNNs. EDSR [23] also belongs to the feedback deep network. Compared with SRResNet in structure, EDSR removes the batch normalization (BN) operation. SRGAN [24] is a generative adversarial network. It takes the residual network as the main network of feature extraction and adds the perceptive loss function. Alternate training generates network and discriminant network. The three major types of representative networks have made good contributions to super-resolution reconstruction of natural images.

At present, some deep learning methods are also applied to MR image super resolution. Chen et al. [25] apply DenseNet to brain MRI image super resolution, Chen et al. [26] propose a novel 3D CNN architecture, namely, mDCSRN, which provides appealing sharp SR images with rich texture details that are highly comparable with the referenced HR images.

We compared and analyzed bicubic, RDN, EDSR, SRGAN, DBPN, and SRFBN, algorithms on the constructed MR image dataset, covering traditional super-resolution learning and a batch of deep learning networks with excellent super-resolution performance. As can be seen from the experimental results, the super-resolution effect of the deep learning network on the MR image performs well, and the super-resolution effect of different parts of the same algorithm fluctuates greatly.

Our contributions include the following three points:

(1) We proposed a unified framework of deep learning-based MRI super resolution and carried out an experimental analysis of the traditional super-resolution methods and deep learning super-resolution methods on the MRI dataset.

(2) We build a dataset for super resolution of MR images, including 4300 high-low resolution pairs. It involves the skull, breasts, knees, and head and neck.

(3) We performed statistical analysis on the difficulty of MR images in different parts and provided suggestions for MR images in different parts.

2. A Unified Framework of Deep Learning and Dataset

2.1. Image Degradation. Because of tissue and organ movement in the process of imaging, the existence of the noise and artifacts makes image or part of the image blur, or resolution is not enough, so by the imaging model [27, 28], the low-resolution image is the high-resolution image by ideal after deformation, fuzzy, blood sampling, such as noise adding income after operation, and process of image degradation process. Figure 1 shows the degradation process of high-resolution images. X is an undergraded high-resolution image matrix with an ideal band limit not lower than Nyquist sampling frequency. Let x be the column vector of X. After the degradation process of x, such as deformation matrix M_k, fuzzy matrix B_k, and mass reduction sampling matrix D_k, noise matrix n_k is added to obtain the low-resolution image y_k. The mathematical model formula of image degradation is shown in Figure 1:

$$y_k = D_k B_k M_k x + n_k, \quad 1 \le k \le p. \tag{1}$$

Formula (1) can also be expressed as

$$Y = HX, \tag{2}$$

where H represents the image degradation matrix, Y represents low-resolution images, and X represents ground truth high-resolution images. The purpose of image super resolution is to find the inverse of the degenerate matrix H.

2.2. A Unified Framework of Deep Learning-Based MRI Super Resolution. We propose a unified framework of deep learning-based MRI super resolution. The aim of this framework is to apply the deep learning network for natural image super resolution to MR images. We have integrated five state-of-the-art deep learning networks, respectively,

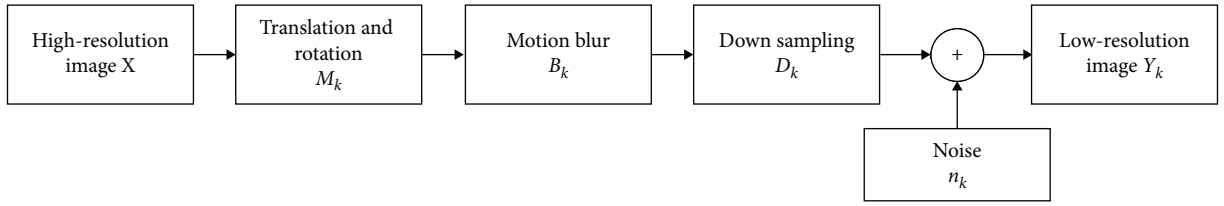


Figure 1: Image degradation model.

Deep Backprojection Networks For Super Resolution (DBPN), Residual Dense Network (RDN), Feedback Network for Image Super Resolution (SRFBN), Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network (SRGAN), and Enhanced Deep Residual Networks for Single Image Super Resolution (EDSR). The unified framework of deep learning-based MRI super resolution is shown in Figure 2.

DBPN: the structure of DBPN is shown in Figure 2(a). It consists of initial feature extraction, projection, and reconstruction.conv (f, n) represents the convolution layer, n represents the number of filters, and f represents the filter size.

Initial feature extraction: we use 3*3 convolution to extract features from low-resolution images. Then, we use 1*1 convolution to achieve feature reduction. n_0 is the number of filters used in the initial LR feature extraction phase. n_R is the number of filters used in each projection unit.

Backprojection stages: from Figure 2(A), the main component of the DBPN structure is the projection unit. As part of training the SR system, it maps LR features to HR features, or maps HR features to LR features in two opposite mapping relationships. The two mapping relationships are

$$\text{Upsampling}: H_0^t = \left(L^{t-1} * p_t\right)\uparrow s, \tag{3}$$

$$\text{Downsampling}: L_0^t = \left(H^{t-1} * g_t\right)\downarrow s, \tag{4}$$

$$\text{Residual}: e_t^l = L_0^t - L^{t-1}, \tag{5}$$

$$\text{Residualup sampling}: H_1^t = \left(e_t^l * q_t\right)\uparrow s, \tag{6}$$

$$\text{Output feature map}: H^t = H_0^t + H_1^t. \tag{7}$$

The above formula * represents a convolution operation, $\uparrow s$ stands for up sampling operation, $\downarrow s$ stands for down sampling operation, and p_t, g_t, q_t is the (de) convolutional layer of the stage t, which learns the mapping of low-resolution feature maps to high-resolution feature maps.

Reconstruction: the feature diagrams obtained by each reflection unit are connected to form $[H^1, H^2, \ldots, H^t]$. And, the combined features are transformed into reconstructed images by 3*3 convolution.

RDN: it can be seen from (B) Figure 2 that RDN network is composed of SFENet, RDBs, DFF, and UPNet. SFENet is composed of two convolutional layers for extracting shallow features. RDBs module is composed of the residual block and the dense block. This operation facilitates the training of RDB modules with larger growth rates. DFF module consists of global feature fusion and global residual learning, which realizes the global extraction of the characteristics of each layer. UPNet implements image upsampling.

SRFBN: we can see from (C) in Figure 2 that the network is mainly an improvement of a feedback mechanism based on the DRCN large framework, which is equivalent to turning the weight-sharing layer in DRCN into a weight-sharing module and add a skip connection. It consists of three parts: feature extraction, weight-sharing module, and learning strategy.

(1) Feature extraction: this network is mainly for shallow feature extraction, that is, shallow feature extraction is $F_{\text{in}}^t = f_{\text{LRFB}}$.

(2) Weight-sharing module: the output of the t weight-sharing module should be $F_{\text{out}}^t = f_{FB}(F_{\text{out}}^{t-1}, F_{\text{in}}^t)$, and the corresponding intermediate supervision output is $I_{\text{SR}}^t = I_{\text{Res}}^{t-1} + f_{\text{UP}}(\text{LR})$, among them, $I_{\text{Res}}^t = f_{\text{RB}}(F_{\text{out}}^t)$.

(3) Learning strategy: the learning supervision function used by this model is as follows:

$$L(\Theta) = \frac{1}{T}\sum_{t=1}^{T} W^t \left\| I_{HR}^t - I_{SR}^t \right\|_1. \tag{8}$$

The truth value of the intermediate supervision will be selected according to the difficulty of the task, such as a single bicubic downsampling degradation, and all truth values are the same; for BD (bicubic + blur) degradation, the first two intermediate supervised outputs use truth values with Gaussian blur, and the subsequent intermediate supervises use truth values without Gaussian blur.

SRGAN: it can be seen from Figure 2(D) that SRGAN is composed of the generator, discriminator, and loss function.

The generator is composed of multiple residual blocks in the generated network part (SRResNet). Each residual block is composed of two 3×3 convolution layers, the batch normalization layer (BN), PReLU as the activation functions are the latter items of the convolution layer, and two subpixel convolution layers.

The discriminator consists of 8 convolutional layers in the discriminating network part, and LeakyReLU is selected as the activation function. The role of the discriminator is to determine the difference between the high-resolution image output by the generator and the real high-resolution image.

The loss function in SRGAN is more special than other networks. It uses two loss functions: G loss and D loss. G loss and D loss can be expressed as

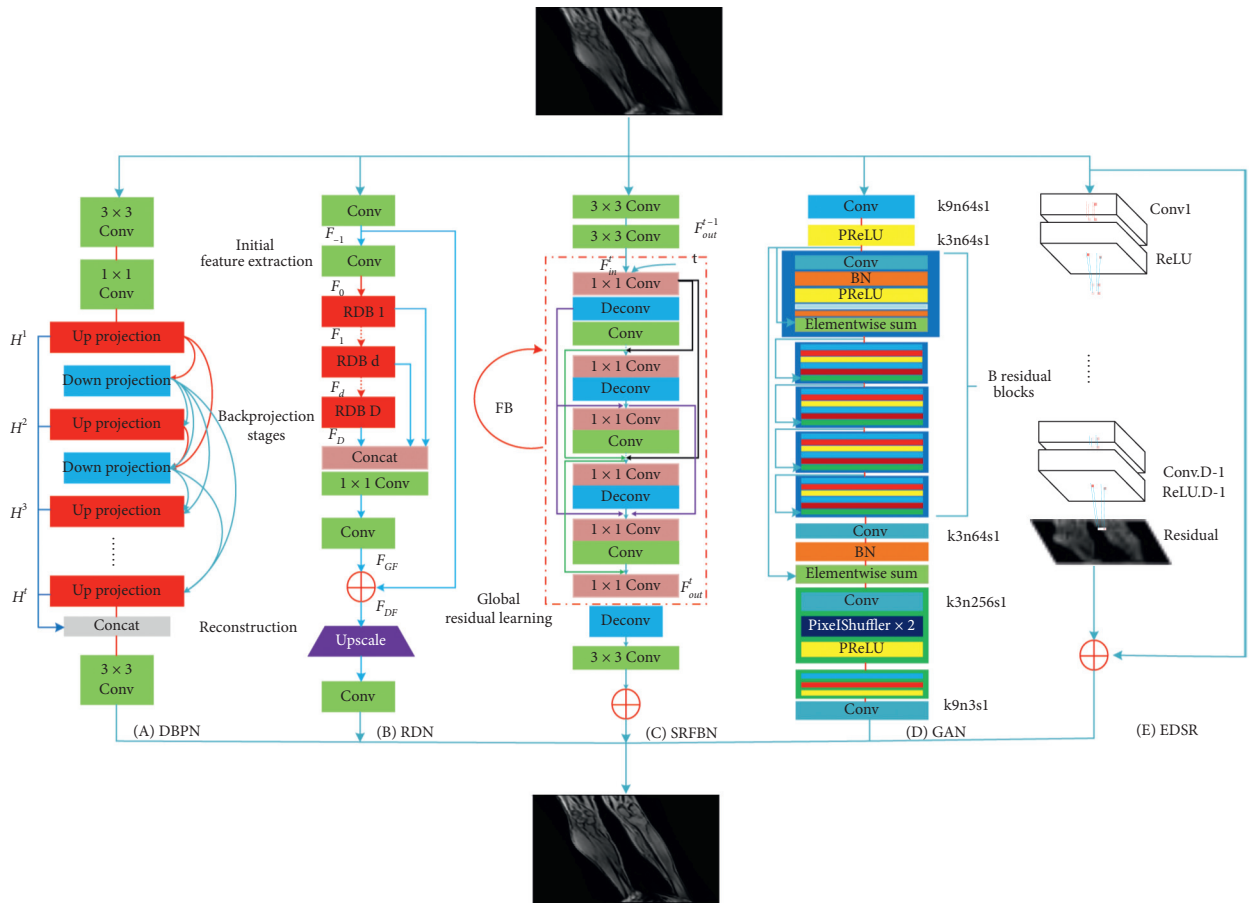


Figure 2: A unified framework of deep learning-based MRI super resolution.

$$l^{\mathrm{SR}} = l_X^{\mathrm{SR}} + 10^{-3} l_{\mathrm{Gen}}^{\mathrm{SR}}, \tag{9}$$

$$\min_{\theta_G} \max_{\theta_D} E_{I^{\mathrm{HR}} \sim p_{\mathrm{train}}(I^{\mathrm{HR}})}\left[\log D_{\theta_D}\left(I^{\mathrm{HR}}\right)\right] + E_{I^{\mathrm{LR}} \sim p_G(I^{\mathrm{LR}})}\left[\log\left(1 - G_{\theta_G}\left(I^{\mathrm{LR}}\right)\right)\right]. \tag{10}$$

It can be seen from the above formula that l_X^{SR} is content loss and $10^{-3} l_{Gen}^{SR}$ is against loss. The losses of the two are

$$l_{\mathrm{MSE}}^{\mathrm{SR}} = \frac{1}{r^2 WH} \sum_{x=1}^{rW} \sum_{y=1}^{rH} \left(I_{x,y}^{\mathrm{HR}} - G_{\theta_G}\left(I^{\mathrm{LR}}\right)_{x,y}\right)^2, \tag{11}$$

$$l_{\mathrm{VGG}/i,j}^{\mathrm{SR}} = \frac{1}{W_{i,j} H_{i,j}} \sum_{x=1}^{W_{i,j}} \sum_{y=1}^{H_{i,j}} \left(\phi_{i,j}\left(I^{\mathrm{HR}}\right)_{x,y} - \phi_{i,j} G_{\theta_G}\left(I^{\mathrm{LR}}\right)_{x,y}\right)^2. \tag{12}$$

EDSR: it can be seen from (E) in Figure 2 that it is similar to SRResnet, but the structure lacks a ReLU layer and a batchnorm layer, mainly, because the batch normalization layer normalizes the function. Therefore, by normalizing the functions, the scope flexibility of the network can be eliminated. Because the structure adopted by this method is too deep, the instability of the training process can easily cause numerical instability. To solve this problem, the model uses residual scaling to deal with it, by which the last convolutional layer output of the residual module is multiplied by 0.1.

2.3. Dataset for Training. We collected the MR image data of the four body parts of the head and neck, breast, bones, and skull on the open source website and used the bicubic downsampling method to construct high-low resolution MR image pairs of different scales, including ×2, ×3, and ×4, and divided the training set, validation set, and test set according to 7 : 2 : 1.

2.3.1. Collection and Quality Filtering of Raw Datasets. Our data comes from open source MR image data, including NYU fastMRI Dataset [29], IXI Dataset [30], TCIA MRI Dataset [31], and mridata.org [32]. Because different datasets include different parts, some data contain only one human body part. For example, NYU fastMRI dataset contains the skull and knee, IXI dataset contains the skull only, and TCIA MRI dataset contains the breast, skull, head and neck, and bladder. Therefore, this dataset is mainly based on the MRI image data downloaded from TCIA. The skull data also comes from NYU fastMRI dataset and IXI dataset; the knee parts are from mridata.org and NYU fastMRI dataset.

MR image data is stored in DICOM format, which is an international standard for medical images. We used *Python*'s third-party library pydicom to parse the obtained raw MR data. Thus, MR images of various organs are acquired. The MR image information of each part is shown in Table 1.

To reduce the impact caused by the signal-to-noise ratio, contrast, motion artifacts, and chemical artifacts of the image and in addition to considering the cost borne by the MRI equipment, we collect and use the mainstream magnetic field strength. Because we acquire data of different parts, according to different imaging standards of medical target organs, we obtain original images with different resolutions.

As shown in Table 1, we obtained a large amount of raw MR image data, but not all data are suitable for MR image super resolution. Unqualified data will reduce the effect of image super resolution. Therefore, we conducted a quality assessment of the data obtained for each organ and proposed data that did not meet the requirements. We screened MR images using both manual and machine methods.

Firstly, according to the advice of professional doctors, we manually removed the MR image data with obvious quality problems. Secondly, the performance of deep learning algorithm is positively correlated with the quality and quantity of data. However, more samples do not mean better performance. Poor quality data will not help the deep learning training, but will reduce the quality of reconstructed images. In this paper, we used a method based on gray consistency and gradient combined to evaluate MR image quality. The number of filtered images is shown in Table 1. The filtered data will be used as the training data of deep learning.

2.3.2. Training Set Generation. Various types of deep learning networks need to be based on prior knowledge which is a pair of high-low resolution MR images. We adopt the downsampling method based on the bicubic method. Downsampling based on cubic interpolation first requires the construction of a bicubic function. Its expression is

$$W(x) = \begin{cases} (a+2)|x|^3 - (a+3)|x|^2 + 1, & \text{for} |x| \le 1, \\ (a+2)|x|^3 - 5a|x|^2 + 8a|x| - 4a, & \text{for} 1 < |x| < 2, \\ 0, & \text{otherwise.} \end{cases} \tag{13}$$

Secondly, to treat the interpolated image points, take the nearby 4×4 area. Interpolate as follows.

$$f(x, y) = \sum_{i=0}^{3} \sum_{j=0}^{3} f\left(x_i, y_j\right) W(x - x_i) W\left(y - y_j\right). \tag{14}$$

Thirdly, the image obtained by upsampling is processed by downsampling, and the corresponding downsampling is

$$g(x, y) = f(x, y)\downarrow, \tag{15}$$

The HR is the original image. $g(x, y)$ is the LR image obtained by downsampling.

In order to simulate the real MR image acquisition process, we sampled HR images down and added Gaussian noise to obtain corresponding LR images. The downsampling method based on LR plus noise is to add the original image without noise and add 5% Gaussian noise to the LR image obtained after downsampling:

$$\text{HR}\downarrow s \longrightarrow \text{LR}_s + 5\%\text{GN}. \tag{16}$$

Among them, 5%GN represents 5% of Gaussian noise. $\text{LR}_s + 5\%\text{GN}$ represents low-resolution images of s size corresponding to this method.

We use the method of direct downsampling of the original image based on cubic interpolation to generate MR-SR high-low resolution pairs as input to the deep learning network. We mainly use the bicubic downsampling to implement the HR- > LR process. The downsampling scales of the dataset are ×2, ×3, and ×4. The generated multiscale MR image training set is shown in Figure 3.

3. Experimental Results and Analysis

3.1. Experimental Setup and Evaluation Index

3.1.1. Experimental Setup. During the training process, the hardware configuration we used is Core i9-9900K processor and dual-pass 1080ti graphics card, 32G memory. The software configuration is super-resolution network training using pytorch framework. We set the parameters of the 5 types of networks and iterate 200 rounds. Set the learning rate to 0.01, and its network parameter setting table is shown in Table 2. We found that the learning rates of the five models were fine-tuned within a certain range, and they all reached convergence before 200 rounds of training iterations. After 200 rounds of training, the losses of the five models hardly changed. For better comparative analysis, we set the same number of iterations.

3.1.2. Evaluation Index. In this paper, PSNR and SSIM are used to comprehensively measure the effect of MR image super-resolution reconstruction.

PSNR: this evaluation standard is the most commonly used. The higher the PSNR, the better the reconstructed image quality. The calculation formula is as follows:

$$\text{MSE} = \frac{1}{MN} \|X - Y\|^2, \tag{17}$$

where X stands for the high-resolution image, Y represents the reconstructed image, M and N, respectively, represent the height and width of the image, and L represents the largest gray value in the gray level, where $L = 255$.

SSIM: this evaluation indicates the degree of structural similarity between the reconstructed image and the original image. The larger the value, the more similar the reconstructed image and the original image and the better the reconstruction effect. The calculation formula is

$$\text{SSIM}(X, Y) = [l(X, Y)]^{\alpha} [c(X, Y)]^{\beta} [s(X, Y)]^{\gamma}, \tag{18}$$

Table 1: Sorting of raw MR data.

Data name	Magnetic field strength (T)	Resolution	Raw/filtered data
Knee	1.5	256*256	4517/1000
Skull	3.0	256*256	7917/2200
Breast	1.5	288*288	3816/600
Head and neck	3.0	512*512	2653/500

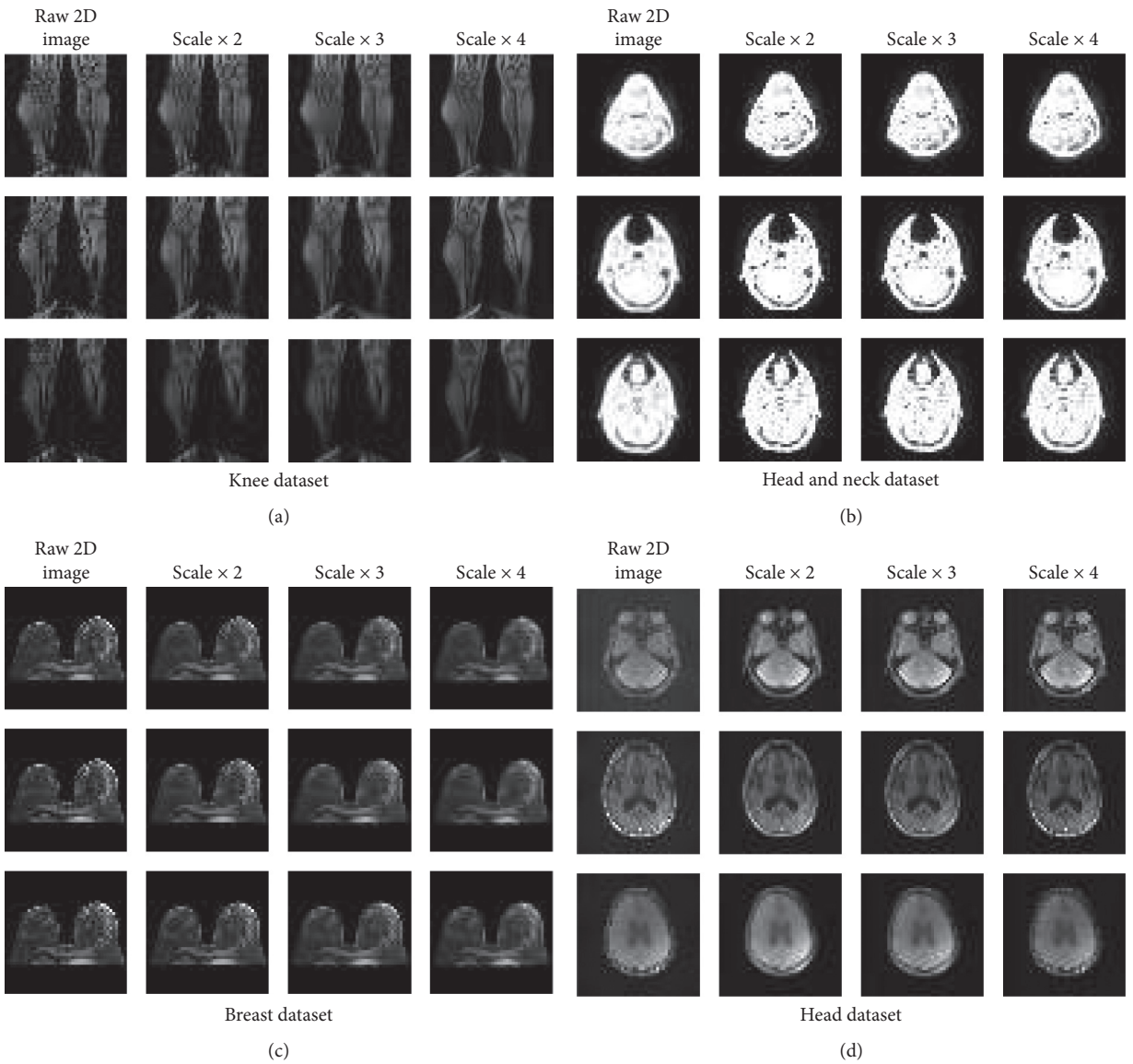


Figure 3: MR image super-resolution pairs, including three different high-low resolution pairs of MR images.

Table 2: Parameter settings of each network.

Parameter	RDN	EDSR	GAN	FBN	DBPN
Lr	0.0001	0.0001	0.0001	0.0001	0.0001
bath_size	64	64	128	64	128
Number of iterations	200	200	200	200	200
Activation	Relu	Relu	Leak Relu	Relu	Relu

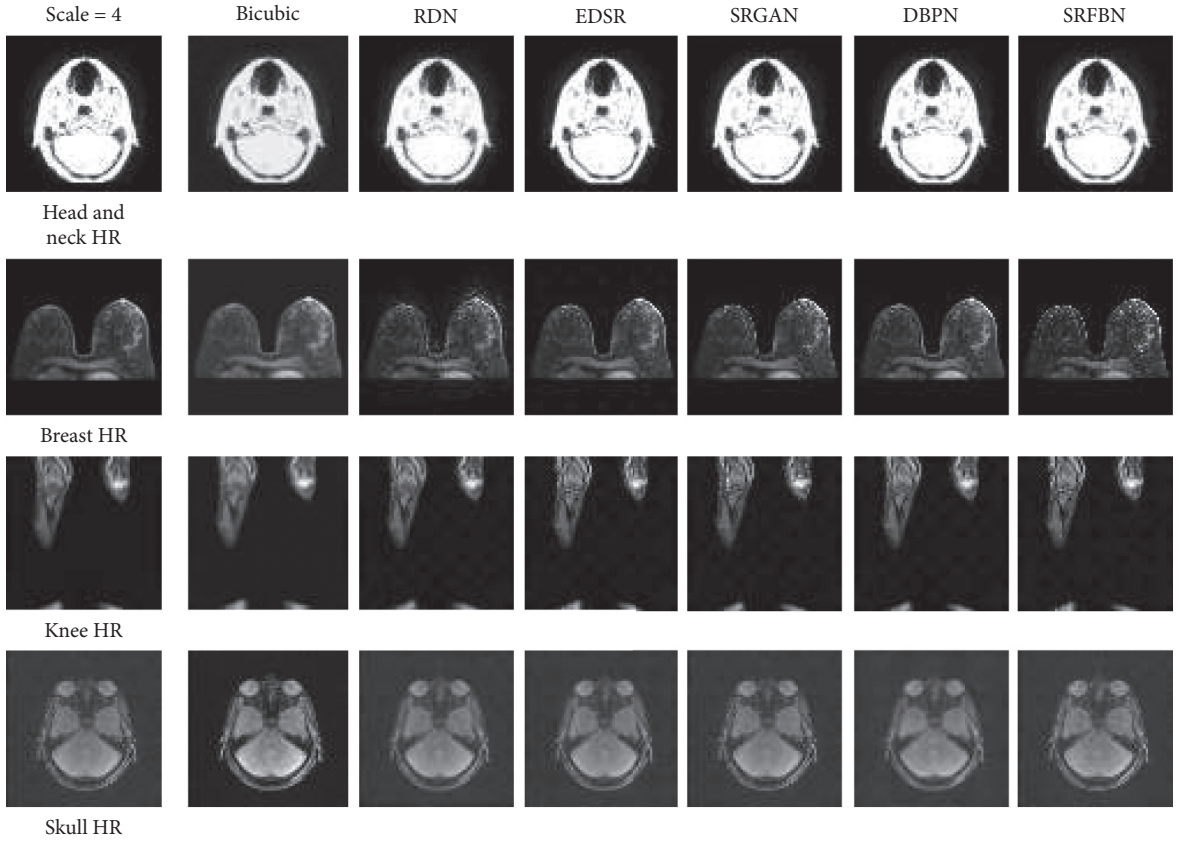


Figure 4: HR comparison of different super-resolution deep networks for each MRI data with scale = 4.

where $l(X,Y)$ stands for the brightness contrast operator, $c(X,Y)$ stands for the contrast operator, $s(X,Y)$stands for the structure contrast operator, and $\alpha,\beta,\gamma>0$ is used to adjust the weights of the three operators. The calculation formulas are as follows.

Brightness contrast operator:

$$l(X,Y)=\frac{2\mu_x\mu_y+C_1}{\mu_x^2+\mu_y^2+C_1},\qquad C_1=(K_1L)^2. \tag{19}$$

Contrast operator:

$$c(X,Y)=\frac{2\sigma_x\sigma_y+C_2}{\sigma_x^2+\sigma_y^2+C_2},\qquad C_2=(K_2L)^2. \tag{20}$$

Structure contrast operator:

$$s(X,Y)=\frac{\sigma_{xy}+C_3}{\sigma_x\sigma_y+C_3},\qquad C_3=(K_3L)^2, \tag{21}$$

where μ represents the average value of the image, that is, the average brightness of the image, σ represents the standard deviation of the signal and estimates the contrast of the signal, C represents the normalization factor that overcomes zero, and K is a constant. The mean μ and variance σ are calculated as

$$\mu_x=\frac{1}{MN}\sum_{i=1}^{MN}x_i, \tag{22}$$

$$\sigma_x=\sqrt{\frac{1}{MN}\sum_{i=1}^{MN}(x_i-\mu_x)^2}. \tag{23}$$

3.2. Experimental Result and Analysis. After training of various deep network models, the results obtained by each model in our MR dataset have obvious manifestations. As can be seen from Figure 4 and Table 3, the SRGAN algorithm is richer in visual texture. The SRFBN algorithm is better at the PSNR and SSIM indicators. We can draw this conclusion. First of all, for the same network, the SR effect of different organ parts is significantly different, which is mainly caused by the original resolution of each organ, the magnetic field intensity during the acquisition process, and the proportion of free water and bound water in the organ. Secondly, we can customize a network specifically adapted to this organ for different parts. For example, in Table 3, although the RDN network does not perform well in the head and neck, breast, knee, and other organs, this network

Table 3: PSNR/SSIM of MRI dataset on different networks.

Dataset	Scale	Bicubic	RDN	EDSR	SRGAN	D-DBPN	SRFBN
Head and neck	×2	25.44/0.6928	29.05/0.8824	37.34/0.9861	37.69/0.9868	35.74/0.9800	38.03/0.9879
	×3	23.04/0.6289	27.41/0.8326	31.14/0.8693	—/—	31.61/0.9474	32.79/0.9595
	×4	21.58/0.5442	24.53/0.7144	29.25/0.9164	28.78/0.9093	29.28/0.9182	30.23/0.9304
Breast	×2	28.56/0.8848	33.50/0.9373	34.57/0.9447	34.06/0.9332	31.94/0.9243	34.28/0.9428
	×3	21.85/0.6130	30.61/0.8876	31.25/0.8949	—/—	28.56.0.8568	32.14/0.9040
	×4	20.50/0.4830	23.68/0.6118	30.00/0.8631	27.99/0.8137	28.13/0.8351	30.68/0.8735
Knee	×2	30.60/0.9185	36.34/0.9641	37.49/0.9719	36.93/0.9670	37.67/0.9727	37.95/0.9744
	×3	24.36/0.6202	32.13/0.9103	32.85/0.9230	—/—	33.75/0.9360	33.82/0.9377
	×4	22.69/0.5649	30.02/0.8597	28.34/0.8968	29.63/0.8643	31.29/0.8940	31.35/0.8964
Skull	×2	20.320.6025	38.05/0.9565	38.03/0.9565	35.80/0.9234	33.82/0.9062	39.19/0.9600
	×3	19.23/0.5843	28.81/0.7790	34.16/0.9211	—/—	31.38/0.8620	35.65/0.9300
	×4	18.60/0.5623	29.70/0.7962	32.22/0.8886	30.95/0.8500	29.39/0.7906	33.60/0.9089

Table 4: Model parameter estimation and calculation capabilities.

Methods	RDN	EDSR	SRGAN	D-DBPN	SRFBN
Param (M)	22.317	43.103	214.376	10.426	3.546
FLOPs (T)	7.43	11.71	15.32	4.27	0.94

obviously has a better PSNR/SSIM of 38.05/0.9565 in the skull organs than other networks, indicating that different target organs can be customized to design super-resolution networks.

We can see from the table that the PSNR/SSIM index between the traditional bicubic algorithm and the deep network is particularly different, which is basically between 3.61–12.59/0.1896–0.2951, which also shows that the deep network has better image super-resolution reconstruction effect on MR image super resolution. It can be seen through experiments that directly applying deep learning networks can also achieve certain effects, indicating that deep networks have a certain role in MR image super resolution and are better than traditional methods. However, there is still a gap with the super-resolution effect of natural images. The main reason is that the natural image and MR image have different imaging mechanisms. Therefore, it is necessary to combine the characteristics of MR images of different human parts to design a more targeted deep super-resolution network of MR images.

We made statistics on the number of parameters and calculation amount of the five deep learning methods, as shown in Table 4.

It can be seen from Table 4 that the model parameters and calculation amount of different networks differ greatly. Combined with the super-resolution performance of the five methods in Table 3, the SRFBN method not only has better super resolution effect but also has the least model size and calculation amount, which is more suitable for practical application.

4. Conclusion

We propose a unified framework of deep learning-based MRI super resolution. We have integrated five state-of-the-art deep learning networks. Moreover, the deep learning method is experimentally verified on the self-built MR image dataset which covers the skull, knee, breast, and head and neck. Through data quality screening and analog imaging degradation, MR image dataset with certain scale and standard for image super resolution is formed. Compared with the traditional method, the deep learning method has better reconstruction performance on the data set. The reason of the difference of the super resolution of different organs is revealed from the structure level of each organ. We hope that our paper can provide data support for the application of deep learning networks in MR super resolution and inspire future research on MR image super resolution.

Data Availability

The two popular MRI datasets in this paper, fastMRI Dataset and IXI Dataset can be freely downloaded from https://fastmri.org/and http://www.brain-development.org/.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by National Science Foundation of China under Grant nos. 61671170 and 61872085, Science and Technology Foundation of National Defense Key Laboratory of Science and Technology on Parallel and Distributed Processing Laboratory(PDL) under Grant no. 6142110180406, Science and Technology Foundation of ATR National Defense Key Laboratory under Grant no. 6142503180402, China Academy of Space Technology (CAST) Innovation Fund under Grant no. 2018CAST33, and Joint Fund of China Electronics Technology Group Corporation and Equipment Pre-Research under Grant no. 6141B08231109.

References

[1] K. Jones, *Diffusion MRI: Theory, Methods, and Applications. Modern formal methods and applications*Springer, Berlin, Germany, 2010.

[2] A. Mallouhi, "Craniocerebral trauma: magnetic resonance imaging of diffuse axonal injury," *Der Radiologe*, vol. 54, no. 9, 2014.

[3] E. D. Schwartz, E. T. Cooper, Y. Fan et al., "MRI diffusion coefficients in spinal cord correlate with axon morphometry," *Neuroreport*, vol. 16, no. 1, pp. 73–76, 2005.

[4] P. Makowski, T. S. Sørensen, S. V. Therkildsen, A. Materka, H. Stødkilde-Jørgensen, and E. M. Pedersen, "Two-phase active contour method for semiautomatic segmentation of the heart and blood vessels from MRI images for 3D visualization," *Computerized Medical Imaging and Graphics*, vol. 26, no. 1, pp. 9–17, 2002.

[5] C. Peterfy, "SP0201 cartilage assessment by MRI in large and small joints in rheumatoid arthritis," *Annals of the Rheumatic Diseases*, vol. 72, no. 3, p. A46, 2013.

[6] Soft tissues of pelvis in patients with cervical cancer: MRI before, during, and after radiation therapy.

[7] F. P. Miller and A. F. Vandome, *John McBrewster. Bicubic interpolation*, Alphascript Publishing, Riga, Latvia, 2010.

[8] N. Jiang and L. Wang, "Quantum image scaling using nearest neighbor interpolation," *Quantum Information Processing*, vol. 14, no. 5, pp. 1559–1571, 2015.

[9] M. Irani and S. Peleg, "Super resolution from image sequences," in *Proceedings of the pattern recognition, 1990 10th International Conference on IEEE*, Atlantic City, NJ, USA, June 1990.

[10] I. Michal, "Improving resolution by image registration," *CVGIP: Graphical Models and Image Processing*, vol. 97, 1991.

[11] I. Michal, "Motion analysis for image enhancement: resolution, occlusion, and transparency," *Journal of Visual Communication and Image Representation*, vol. 4, 1993.

[12] R. R. Schultz and R. L. Stevenson, "A Bayesian approach to image expansion for improved definition," *IEEE Transactions on Image Processing*, vol. 3, no. 3, pp. 233–242, 1994.

[13] R. C. Hardie, K. J. Barnard, and E. E. Armstrong, "Joint MAP registration and high-resolution image estimation using a sequence of undersampled images," *IEEE Transactions on Image Processing*, vol. 6, no. 12, pp. 1621–1633, 1997.

[14] X. Zhang, K.-M. Lam, and L. Shen, "Image magnification based on a blockwise adaptive Markov random field model," *Image and Vision Computing*, vol. 26, no. 9, pp. 1277–1284, 2008.

[15] H. Ur and D. Gross, "Improved resolution from subpixel shifted pictures," *CVGIP: Graphical Models and Image Processing*, vol. 54, no. 2, pp. 181–186, 1992.

[16] M. S. Alam, J. G. Bognar, R. C. Hardie, and B. J. Yasuda, "Infrared image registration and high-resolution reconstruction using multiple translationally shifted aliased video frames," *IEEE Transactions on Instrumentation and Measurement*, vol. 49, no. 5, pp. 915–923, 2000.

[17] H. Stark and P. Oskoui, "High-resolution image recovery from image-plane arrays, using convex projections," *Journal of the Optical Society of America A*, vol. 6, no. 11, pp. 1715–1726, 1989.

[18] J. Yang, Z. Wang, Z. Lin et al., "Coupled dictionary training for image super-resolution," *Image Processing IEEE Transactions on*, vol. 21, no. 8, pp. 3467–3478, 2012.

[19] Z. Wang, J. Chen, and C. H. Hoi Steven, "Deep learning for image super-resolution: a survey," *IEEE transactions on pattern analysis and machine intelligence*, vol. 32, no. 99, 2020.

[20] M. Haris, G. Shakhnarovich, and N. Ukita, "Deep back-projection networks for super-resolution," 2018.

[21] Y. Zhang, Y. Tian, Y. Kong et al., "Residual dense network for image restoration," 2018.

[22] Z. Li, J. Yang, L. Zheng, X. Yang, and W. Wu, "Feedback network for image super-resolution," in *Proceedings of the 2019 IEEE/CVF conference on computer vision and pattern recognition (CVPR)*, Long Beach, CA, USA, June 2019.

[23] B. Lim, S. Son, H. Kim et al., "Enhanced deep residual networks for single image super-resolution," *Convolutional Neural Networks*, vol. 23, 2017.

[24] C. Ledig, L. Theis, F. Huszar et al., "Photo-realistic single image super-resolution using a generative adversarial network," 2016.

[25] Y. Chen, Y. Xie, Z. Zhou et al., "Brain MRI super resolution using 3D deep densely connected neural networks," *New Neural Network Architecture*, vol. 32, pp. 739–742, 2018.

[26] Y. Chen, A. G. Christodoulou, Z. Zhou et al., "MRI super-resolution with GAN and 3D multi-level densenet: smaller, faster, and better," 2020, https://arxiv.org/abs/2003.01217.

[27] L. K. Saul and S. T. Roweis, "An introduction to locally linear embedding," *Journal of Machine Learning Research*, vol. 7, 2001.

[28] S. Borman and L. Robert, "Super-resolution from image sequences-a review," *Proceeding of: Circuits and Systems*, vol. 47, pp. 1–3, 1998.

[29] J. Zbontar, F. Knoll, A. Sriram et al., "An open dataset and benchmarks for accelerated MRI," *Fastmri*, vol. 65, 2018.

[30] IXIDataset. http://www.brain-development.org/.

[31] K. Clark, B. Vendt, K. Smith et al., "The cancer imaging archive (TCIA): maintaining and operating a public information repository," *Journal of Digital Imaging*, vol. 26, no. 6, pp. 1045–1057, 2013.

[32] K. Hammernik, T. Klatzer, E. Kobler et al., "Learning a variational network for reconstruction of accelerated MRI data," *Magnetic Resonance in Medicine*, vol. 79, 2018.

Hindawi
Journal of Healthcare Engineering
Volume 2021, Article ID 5533208, 14 pages
https://doi.org/10.1155/2021/5533208

Research Article

Theory-Based Failure Modes and Effect Analysis for Medication Errors

Saeid Jafarzadeh Ghoushchi,[1] **Shadi Dorosti**,[1] **Mohd Nizam Ab Rahman**,[2] **Marzieh Khakifirooz**,[3] **and Mahdi Fathi**[4]

[1]*Faculty of Industrial Engineering, Urmia University of Technology, Urmia, Iran*
[2]*Department of Mechanical and Manufacturing Engineering, Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia, Bangi, Selangor, Malaysia*
[3]*Department of Industrial Engineering, Tecnologico de Monterrey, Monterrey,NL, Mexico*
[4]*Department of Information Technology & Decision Sciences, G. Brint Ryan College of Business, University of North Texas, Denton, TX, USA*

Correspondence should be addressed to Saeid Jafarzadeh Ghoushchi; s.jafarzadeh@uut.ac.ir

Received 7 January 2021; Revised 27 February 2021; Accepted 22 March 2021; Published 1 April 2021

Academic Editor: Chi-Hua Chen

Medication Errors (MEs) are still significant challenges, especially in nonautomated health systems. Qualitative studies are mostly used to identify the parameters involved in MEs. Failing to provide accurate information in expert-based decisions can provoke unrealistic results and inappropriate corrective actions eventually. However, mostly, some levels of uncertainty accompany the decisions in real practice. This study tries to present a hybrid decision-making approach to assigning different weights to risk factors and considering the uncertainty in the ranking process in the Failure Modes and Effect Analysis (FMEA) technique. Initially, significant MEs are identified by three groups of qualified experts (doctors, nurses, and pharmacists). Afterward, for assigning weights to the risk factors, *Z*-number couples with the Stepwise Weight Assessment Ratio Analysis (SWARA) method, named *Z*-SWARA, to add reliability concept in the decision-making process. Finally, the identified MEs are ranked through the developed Weighted Aggregated Sum Product Assessment (WASPAS) method, namely, *Z*-WASPAS. To demonstrate the applicability of the proposed approach, the ranking results compare with typical methods, such as fuzzy-WASPAS and FMEA. The findings of the present study highlight improper medication administration as the main failure mode, which can result in a fatality or patient injury. Moreover, the utilization of multiple-criteria decision-making methods in combination with *Z*-number can be a useful tool in the healthcare management field since it can address the problems by considering reliability and uncertainty simultaneously.

1. Introduction

Medication Errors (MEs) are the most common medical error that can disproportionately affect patients [1]. According to the available statistics, MEs affect approximately 1.5 million people each year [2]. MEs can occur throughout the medication-use system, such as prescribing medication, entering information into a system when medication is being prepared or dispensed, or when medication is given to or taken by a patient. Indeed, MEs are avoidable occurrences that may typically cause improper medication usage or patient damage. Hence, to prevent patients' injury or fatality, health centers endeavor to make practical measurements and develop various systems for MEs minimization [3].

One of the trends for reducing MEs is establishing advanced medication systems which are adopted in some countries, including the United States [4], Malaysia [5], New Zealand [6], and China [7]. However, improving systems depends on monitoring all steps of medical services and true recognition of failure modes to eliminate

the problem [8]. Failure Modes and Effect Analysis (FMEA) is a systematic tool to identify the potential failures, causes and effects of failures, and provide preventive actions [9, 10]. Therefore, it is an appropriate approach to use in the health-care field [11–13] and recognize possible failure modes in the MEs process [14, 15]. Based on the FMEA method, the risk priorities of the failure modes are specified through the Risk Priority Number (RPN) value. The RPN is the result of the occurrence, severity, and detection of potential failure modes multiplication [16, 17]. However, RPN scores have been criticized for some deficiencies, including uncertainty in team-based decision making, the shortage of full ranking for risks, and assumption of the equal importance of risk factors and not considering uncertainty concept [18–20]. Notably, to deal with the uncertainty in solving risk analysis problems, there are various useful theories including R-Number, G-number, and evidence theory [21–23]. Therefore, to address the shortcomings of RPN, it is necessary to develop a novel prioritization approach for prioritizing the identified MEs using Multiple-Criteria Decision-Making (MCDM) methods [24, 25]. In the meantime, the combination of the FMEA technique with the data envelopment analysis [12], best-worst method [19], MOORA [13], VIKOR [17], gray relational projection [21], and TOPSIS [24] can be mentioned.

Recently, researchers attend to apply MCDM methods in complicated decision-making processes, including healthcare sectors [26, 27]. For instance, Mardani et al. [26] concluded MCDM methods as effective techniques in different sections of hospitals and health centers to facilitate complex decisions making processes, assess various health centers and health services, and resolve uncertainty in different levels of decisions in healthcare centers. Hsieh et al. [27], for the first time, applied MCDM methods for reducing MEs. In their study, the Analytic Hierarchy Process (AHP) and technique for order of preference by similarity to ideal solution (TOPSIS) are applied to evaluate the crucial human error factors that are associated with MEs events. Besides, in the past years, researchers suggested various strategies to reduce MEs, for instance, utilizing lean six sigma method [28], simulation-based learning [29, 30], logistic regression [31], and carrying out qualitative studies [32].

The main contribution of this study is to provide an integrated approach to cover the deficiencies of traditional RPN score. Besides, the cost and time factors are added to RPN scoring because they are playing significant roles in MEs [33, 34]. In other words, MEs can be costly for both hospitals and patients and waste the time of proper treatment. In this paper, initially, ten main failure modes that induce the MEs are introduced by experts based on the FMEA method. Thereafter, with the aims of assigning weight to the quintuple factors of RPN, a combination of the Z-number and Stepwise Weight Assessment Ratio Analysis (SWARA) method, named Z-SWARA, is utilized. The Z-number theory is applied for considering the uncertainty and reliability simultaneously in expressing the values of RPN factors for each failure mode. In this way, the deficiencies of conventional FMEA are addressed by assigning different weights to RPN factors based on their importance. The SWARA method has the advantage of a more logical calculation of weights and relative importance of criteria. The ability to estimate experts' opinions about the importance ratio of the criteria in the process of their weight determination is the main element of this method [35]. Moreover, this method helps coordinate and gather data from experts. Furthermore, the SWARA method is uncomplicated, and experts can easily work together. The main advantage of this method in decision making is that in some problems priorities are defined based on policies of companies or countries and there are not any needs for evaluation to rank criteria [36]. In the third phase, the extended version of Weighted Aggregated Sum Product Assessment (WASPAS) using Z-number theory (Z-WASPAS) is applied for ranking ten failure modes, considering reliability in addition to the uncertainty using the Z-number theory, in comparison with fuzzy theory. The Z-numbers theory, like the D-numbers theory extended based on the Dempster–Shafer theory, represents an approach for intending uncertainty and imprecision in expert decisions [37]. In comparison with D-numbers theory, Z-numbers consider the fuzzy information and reliability of this information simultaneously. This theory is an effective tool to express the uncertainty in expert preferences, which is based on the reliability of choosing the appropriate criterion value. The main advantage of the WASPAS method is its high degree of reliability. Integration of rough numbers and the WASPAS method with advantages of both concepts presents very important support in decision-making in everyday conflicting situations [38].

The remainder of this paper is organized as follows: Section 2 introduces the fundamental theorem of Z-number followed by SWARA and WASPAS methods and transformation rules to Z-WASPAS and Z-SWARA. In Section 3, the research framework and ten main failure modes resulting in MEs are presented. In Section 4, the validation results of the proposed method in comparison with traditional methods are indicated. The suggested actions to remove or reduce failure modes and future research directions are compromised in Section 5.

2. Methodology

To explain the proposed approach, first, Z-number is introduced as a reliability increment method, along with the preliminary definitions and mathematical equations. Afterward, transformation rules are discussed in detail for both Z-SWARA and Z-WASPAS methods for ranking failure modes. In this study, the terms "failure mode" and "alternative" and "criteria" and "risk factors" are used interchangeably. Moreover, for simplicity, the fuzzy number term is used for the specific Triangular Fuzzy Number (TFN).

2.1. Terminologies and Notations. The mathematical terminologies and notations used in this study are as follows:

(i) Y: universal set
(ii) y: membership of Y
(iii) $\widetilde{G}$: a fuzzy set
(iv) $\mu_{\widetilde{G}}(y)$: membership function of dependency $y \in Y$ in $\widetilde{G}$ set
(v) (a, b, c): lower, medium, and upper bounds of TFN
(vi) $R(Y)$: fuzzy restriction in Y domain
(vii) $Z = (\widetilde{F}, \widetilde{L})$: a Z-number with the first component $\widetilde{F}$ and second component $\widetilde{L}$
(viii) $\widetilde{F}$: restriction component of Z-number with membership function of $\mu_{\widetilde{F}}(y)$
(ix) $\widetilde{L}$: reliability of restriction component in the Z-number with membership function of $\mu_{\widetilde{L}}(y)$
(x) i: failure mode/alternative index
(xi) j: risk factor/criteria index
(xii) m: number of alternatives ($m = 1, \ldots, i$)
(xiii) n: number of criteria ($n = 1, \ldots, j$)
(xiv) α: crisp value
(xv) $\widetilde{Z}'_j$: converted form of weighted Z-number to TFN for j-th criteria
(xvi) $\widetilde{q}_j$: fuzzy weight coefficient for j-th criteria
(xvii) $\widetilde{w}_j$: fuzzy relative weight for j-th criteria
(xviii) K_i: the utility function for i-th alternative
(xix) H: decision-making matrix with the Z-number elements
(xx) h_{ij}: elements of matrix H
(xxi) H: converted decision-making matrix with the TFN elements
(xxii) $\widetilde{\widetilde{h}}_{ij}$: elements of matrix H
(xxiii) $\widehat{h}_{ij}$: normalized form of the elements of H decision-making matrix
(xxiv) $\widetilde{Q}_i$: weighted sum for i-th alternative
(xxv) $\widetilde{P}_i$: weighted product for i-th alternative

2.2. Preliminary Definitions and Concepts. In this section, preliminary concepts required to develop the proposed approach of this study are presented.

Fuzzy Set. A fuzzy set is a membership function that shows a degree of membership in $[0, 1]$ interval. In (1), a fuzzy set like $\widetilde{G}$ defines y, the membership value, in Y reference set [39]:

$$\widetilde{G} = \left\{ y | \mu_{\widetilde{G}}: \; Y \longrightarrow [0, 1], \quad y \in Y \right\}, \tag{1}$$

where $\mu_{\widetilde{G}}(y)$ represents the degree of belonging of $y \in Y$ in $\widetilde{G}$ set.

TFN. The triplet (a, b, c) set is known as TFN. The membership function of a TFN like $\widetilde{G} = (a, b, c)$ is [40]

$$\mu_{\widetilde{G}}(y) = \begin{cases} \dfrac{y - a}{b - a}, & a \le y \le b, \\ 1, & y = b, \\ \dfrac{c - y}{c - b}, & b \le y \le c, \\ o & \text{otherwise.} \end{cases} \tag{2}$$

In this study, the TFN form of fuzzy numbers is considered.

Fuzzy Restriction. A fuzzy restriction is visualized as an elastic constraint on the values that are assigned to a variable. A restriction may be viewed as a generalized constraint. Suppose $\widetilde{F}$ is a fuzzy set; the restriction $R(Y)$ is as a probabilistic constraint such that [41]

$$R(Y): \; Y \text{ is } \widetilde{F}, \tag{3}$$

where $\widetilde{F}$ is playing the role of the possible distribution of Y.

The statement in equation (3) can be interpreted as follows:

$$R(Y): \; Y \text{ is } \widetilde{F} \longrightarrow \text{poss}(Y = y) = \mu_{\widetilde{F}}(y), \tag{4}$$

where $\mu_{\widetilde{F}}$ is a membership function of F and y is a generic value of Y. $\mu_{\widetilde{F}}$ can be constrained how $\widetilde{F}$ is associated with $R(Y)$.

Z-Number Theory. A Z-number is an ordered pair of $(\widetilde{F}, \widetilde{L})$, where $\widetilde{F}$ and $\widetilde{L}$ are assumed to be TFN. A Z-number associated with a real-valued uncertain variable $\mathbf{Y}$, where $\widetilde{F}$, as a first component, is a fuzzy subset from $\mathbf{Y}$ domain, and $\widetilde{L}$ is a fuzzy subset from a unite interval $[0, 1]$ [42]. A Z-number can provide information about an uncertain variable, where $\widetilde{F}$ is the restriction and $\widetilde{L}$ represents an idea of certainty or reliability. A collection of Z-valuations is referred to as Z-information. It should be noted that much of everyday reasoning and decision-making is based, in effect, on Z-information. According to the fuzzy restriction in equation (2), suppose that Y is a random variable; its probability distribution of Y illustrates the probabilistic restriction on Y. The probabilistic restriction is

$$R(Y): \; Y \text{ is } p, \tag{5}$$

and based on equation (5), the probability density function of Y is explained in the following equation:

$$R(Y): \; Y \text{ is } p \longrightarrow \text{prob}(u \le Y \le u + \mathrm{d}u) = p(u)\mathrm{d}u, \tag{6}$$

where p is the probability density function of Y and $\mathrm{d}u$ represents deferential of u.

To convert Z-number to TFN, assume first that $Z = [(a_1, b_1, c_1), (a_2, b_2, c_2)]$; the first part (a_1, b_1, c_1) plays the role of restriction and the second part (a_2, b_2, c_2) represents reliability [43]. Initially, the second part (reliability) converts into a crisp number, α, as follows:

$$\alpha = \frac{\int y \, \mu_{\widetilde{L}}(y)\mathrm{d}y}{\int \mu_{\widetilde{L}}(y)\mathrm{d}y}, \tag{7}$$

where $\mu_{\tilde{L}}(y)$ is as defined in equation (2).

Then, α as the weight of the second part (reliability) added to the first part (restriction), and the TFN form of weighted Z-number is obtainable through the following equation:

$$\tilde{Z}' = \left(\sqrt{\alpha} * a_1, \sqrt{\alpha} * b_1, \sqrt{\alpha} * c_1\right). \tag{8}$$

2.3. Z-SWARA Method. The fuzzy Stepwise Weight Assessment Ratio Analysis (fuzzy SWARA) is a multiple-attribute decision-making method [35] for calculating the weight of criteria and subcriteria [44] in a fuzzy environment. The fuzzy SWARA acts the same as the SWARA method [35], but the ambiguity in decision-making or incomplete information leads to the extension of the SWARA method to fuzzy SWARA. In the fuzzy SWARA method, experts play a significant role in assigning the weight of the criteria; therefore, the information accumulates based on experts' opinions [45–49]. In this study, the fuzzy SWARA method extended to the Z-SWARA method, and the reliability factor is added to enhance the certainty in the final results. The steps of the Z-SWARA method are as follows:

Step 1. Initially, the experts sort criteria, from the most important to less important in descending order, based on self-identification.

Step 2. Based on the initial opinion, experts need to assign linguistic variables to the relative importance of criteria j in relation to the previous $j-1$ criteria. Thereafter, experts set the value of the first component ($\tilde{F}_j$), according to Table 1 [40], and the reliability component ($\tilde{L}_j$) is assigned through Table 2 [43]. The result forms a Z-number for each criterion.

Step 3. To convert the Z-number as a result in Step 2 to a TFN, firstly, according to equation (6), the second part (reliability) converts into a crisp number, and the weight is added to the first part based on equation (7).

As a numerical example, suppose that for the j-th criteria, the relative importance in the form of linguistic variables is (VLI, M). By replacing the corresponding TFN values of VLI and M from Tables 3 and 4, respectively, the Z-number forms as [((2/7), (1/3), (2/5)), (0.35, 0.5, 0.75)]. The crisp value is $\alpha = 0.53$, and the TFN form of Z-number according to equation (7) is (0.21, 0.24, 0.29). Other transformations of Z-number to TFN are presented in Table 5.

Step 4. Based on the results of Step 3, coefficient $\tilde{q}_j$ as a fuzzy weight coefficient is defined as follows:

$$\tilde{q}_j = \frac{\tilde{q}_{j-1}}{\tilde{Z}'_j}, \tag{9}$$

where $\tilde{q}_j$ is TFN and $\tilde{q}_1 = (1, 1, 1)$.

Step 5. Finally, considering n evaluation criteria, the relative weights of the j-th evaluation criteria are determined as follows:

Table 1: Linguistic variables for weighting criteria [49].

Linguistic variables	TFNs
Equally important (EI)	(1, 1, 1)
Moderately less important (MOL)	(2/3, 1, 3/2)
Less important (LI)	(2/5, 1/2, 2/3)
Very less important (VLI)	(2/7, 1/3, 2/5)
Much less important (MUL)	(2/9, 1/4, 2/7)

$$\tilde{w}_j = \frac{\tilde{q}_j}{\sum_{j=1}^{n} \tilde{q}_j}, \tag{10}$$

where $\tilde{w}_j$ is a TFN.

2.4. Z-WASPAS Method. The fuzzy-WASPAS is a multivariable decision-making method [50], which, such as the WASPAS method [51, 52], is usually used in highly sensitive cases by considering the certainty of the system [51] (for instance, see the example of reservoir flood control management in [53] and see the example of reservoir flood control management in [54]). The fuzzy-WASPAS method is a unique combination of two well-known MCDM approaches, the Weighted Sum Model (WSM) and Weighted Product Model (WPM). In the fuzzy-WASPAS method, the beneficial (e.g., profit, efficiency) or non-beneficial (e.g., cost) aspect of each risk factor must be determined based on experts' opinions. For the beneficial aspects higher values are always desirable, and for non-beneficial smaller values are always preferable. In this study, only beneficial aspects of criteria are considered, and the final output of fuzzy-WASPAS presents as a utility function (K_i) that can help to rank alternatives. The new extended Z-number of fuzzy-WASPAS, namely, Z-WASPAS, is used for ranking failure modes. The steps of the Z-WASPAS method are as follows:

(i) *Step 1.* First, experts determine a linguistic variable for each element, and then the corresponding values of each linguistic variable are assigned to each element to make decision matrix H. Consider a Z-number like $Z = (\tilde{F}_{ij}, \tilde{L}_{ij})$, linguistic value for $\tilde{F}_{ij}$ is selected using Table 3 [50], and similar to Z-SAWARA method, $\tilde{L}_{ij}$ can get the linguistic values from Table 4. Therefore, the decision-making matrix H with Z-number elements is determined as follows:

$$H = \begin{bmatrix} h_{11} & \dots & h_{1n} \\ \dots & \dots & \dots \\ h_{m1} & \dots & h_{mn} \end{bmatrix}, \tag{11}$$

where $h_{ij} = [(a_{ij}^f, b_{ij}^f, c_{ij}^f), (a_{ij}^l, b_{ij}^l, c_{ij}^l)]$, $i = 1, \dots, m$, $j = 1, \dots, n$, m indicates the number of alternatives, and n shows the number of criteria.

Table 2: Linguistics variables for determining reliability [43].

Linguistic variables	Very weak (VW)	Weak (W)	Medium (M)	High (H)	Very high (VH)
TFNs	(0, 0, 0.25)	(0.2, 0.35, 0.5)	(0.35, 0.5, 0.75)	(0.5, 0.75, 0.9)	(0.75, 1, 1)

Table 3: Linguistic variables for rating failure modes [50].

Linguistic variables	Very poor (VP)	Poor (P)	Medium poor (MP)	Fair (F)	Medium good (MG)	Good (G)	Very good (G)
TFNs	(0, 1, 2)	(1, 2, 3)	(2, 3.5, 5)	(4, 5, 6)	(5, 6.5, 8)	(7, 8, 9)	(8, 9, 10)

Table 4: Transformation rules to convert Z-number to TFN ($\widetilde{Z}'$) based on linguistics variables for rating failure modes.

Linguistics variables	TFNs
(VP, VW)	(0, 0.29, 0.58)
(VP, W)	(0, 0.59, 1.18)
(VP, M)	(0, 0.73, 1.46)
(VP, H)	(0, 0.85, 1.69)
(VP, VH)	(0, 0.96, 1.91)
(P, VW)	(0.29, 0.58, 0.87)
(P, W)	(0.59, 1.18, 1.77)
(P, M)	(0.73, 1.46, 2.19)
(P, H)	(0.85, 1.69, 2.54)
(P, VH)	(0.96, 1.91, 2.87)
(MP, VW)	(0.58, 1.01, 1.44)
(MP, W)	(1.18, 2.07, 2.96)
(MP, M)	(1.46, 2.56, 3.65)
(MP, H)	(1.69, 2.96, 4.23)
(MP, VH)	(1.91, 3.35, 4.79)
(F, VW)	(1.15, 1.44, 1.73)
(F, W)	(2.37, 2.96, 3.55)
(F, M)	(2.92, 3.65, 4.38)
(F, H)	(3.39, 4.23, 5.08)
(F, VH)	(3.83, 4.79, 5.74)
(MG, VW)	(1.44, 1.88, 2.31)
(MG, W)	(2.96, 3.85, 4.73)
(MG, M)	(3.65, 4.75, 5.84)
(MG, H)	(4.23, 5.50, 6.77)
(MG, VH)	(4.79, 6.22, 7.66)
(G, VW)	(2.02, 2.31, 2.60)
(G, W)	(4.14, 4.73, 5.32)
(G, M)	(5.11, 5.84, 6.57)
(G, H)	(5.93, 6.77, 7.62)
(G, VH)	(6.70, 7.66, 8.62)
(VG, VW)	(2.31, 2.60, 2.89)
(VG, W)	(4.73, 5.32, 5.92)
(VG, M)	(5.84, 6.57, 7.30)
(VG, H)	(6.77, 7.62, 8.47)
(VG, VH)	(7.66, 8.62, 9.57)

Table 5: Transformation rules for Z-number to TFN based on linguistics variables for weighting criteria.

Linguistics variables	TFNs
(EI, VW)	(1, 1, 1)
(EI, W)	(1, 1, 1)
(EI, M)	(1, 1, 1)
(EI, H)	(1, 1, 1)
(EI, VH)	(1, 1, 1)
(MOL, VW)	(0.19, 0.29, 0.43)
(MOL, W)	(0.39, 0.59, 0.89)
(MOL, M)	(0.49, 0.73, 1.10)
(MOL, H)	(0.56, 0.85, 1.27)
(MOL, VH)	(0.64, 0.96, 1.44)
(LI, VW)	(0.12, 0.14, 0.19)
(LI, W)	(0.24, 0.30, 0.39)
(LI, M)	(0.29, 0.37, 0.49)
(LI, H)	(0.34, 0.42, 0.56)
(LI, VH)	(0.38, 0.48, 0.64)
(VLI, VW)	(0.08, 0.10, 0.12)
(VLI, W)	(0.17, 0.20, 0.24)
(VLI, M)	(0.21, 0.24, 0.29)
(VLI, H)	(0.24, 0.28, 0.34)
(VLI, VH)	(0.27, 0.32, 0.38)
(MUL, VW)	(0.06, 0.07, 0.08)
(MUL, W)	(0.13, 0.15, 0.17)
(MUL, M)	(0.16, 0.18, 0.21)
(MUL, H)	(0.19, 0.21, 0.24)
(MUL, VH)	(0.21, 0.24, 0.27)

(ii) *Step 2*. The decision matrix H with Z-number elements converts to TFN ($\widetilde{Z}'$). The transformed decision-making matrix H is as follows:

$$\widetilde{\widetilde{H}} = \begin{bmatrix} \widetilde{\widetilde{h}}_{11} & \dots & \widetilde{\widetilde{h}}_{1n} \\ \dots & \dots & \dots \\ \widetilde{\widetilde{h}}_{m1} & \dots & \widetilde{\widetilde{h}}_{mn} \end{bmatrix}, \tag{12}$$

where $\widetilde{\widetilde{h}}_{ij}$ is a TFN in the form of $\widetilde{Z}'$, $i = 1, \dots, m$, $j = 1, \dots, n$, m indicates the number of alternatives, and n shows the number of criteria.

As a numerical example, suppose that for the ith alternative and the jth criteria, the importance, $\widetilde{F}_{ij}$, is determined "medium poor" (MP) and the reliability, $\widetilde{L}_{ij}$, is selected "weak" (W) by the expert; consequently, the Z-number and the crisp value are $[(2, 3.5, 5), (0.20, 0.35, 0.50)]$, and $\alpha = 0.35$, respectively. Therefore, the transformed form of Z-number to a TFN ($\widetilde{Z}'$) is $(1.18, 2.08, 2.96)$. Other transformations of Z-number to TFN ($\widetilde{Z}'$) are presented in Table 4.

(iii) *Step 3*. Normalize $\widetilde{\widetilde{H}}$ matrix considering beneficial and nonbeneficial elements, using

$$\widehat{h}_{ij} = \begin{cases} \dfrac{\widetilde{\widetilde{h}}_{ij}}{\max\limits_{i} \widetilde{\widetilde{h}}_{ij}}, & \text{for } j\text{beneficial}, \\ \dfrac{\min\limits_{i} \widetilde{\widetilde{h}}_{ij}}{\widetilde{\widetilde{h}}_{ij}}, & \text{for } i \text{ nonbeneficial}. \end{cases} \tag{13}$$

(iv) *Step 4*. Determine the weighted normalized fuzzy decision-making matrix of $\widehat{h}_{ij}$ for the WSM ($\widetilde{Q}$ TFN) and WPM ($\widetilde{P}$ TFN) as follows:

$$\widetilde{Q}_i = \sum_{j=1}^{n} \widehat{h}_{ij}\widetilde{w}_j$$
$$\widetilde{P}_i = \prod_{j=1}^{n} \widehat{h}_{ij}^{\widetilde{w}_j}. \quad (14)$$

To facilitate defuzzification of the performance measurement, center-of-area of each TFN is considered for decision making as follows [49]:

$$\overline{Q}_i = \frac{1}{3}\left(a_i^Q + b_i^{Q_i} + c_i^Q\right),$$
$$\overline{P}_i = \frac{1}{3}\left(a_i^P + b_i^P + P_i^c\right). \quad (15)$$

(v) *Step 5*. Calculate the value of failure modes and rank them using utility function K_i for ith alternative as follows:

$$K_i = \lambda \sum_{j=1}^{m} \overline{Q}_i + (1-\lambda) \sum_{j=1}^{m} \overline{P}_i; \quad 0 \le \lambda \le 1,\ 0 \le K_i \le 1, \quad (16)$$

where $\lambda = (\sum_{i=1}^{m} \overline{P}_i / \sum_{i=1}^{m} \overline{Q}_i + \sum_{i=1}^{m} \overline{P})$.

Finally, alternatives can be ranked from the highest value of K_i to the lowest one.

3. Research Framework

The main purpose of the current study is to introduce a novel approach to identify and prioritize failure modes that result in MEs. In the first stage of this approach, the experts identify five factors that play an important role in MEs occurrence, including severity (S), occurrence (O), detection (D), cost (C), and time (T). The main reasons for the contribution of the aforementioned factors can be described as follows:

(i) S: identifying the severity of effects that patients will perceive in each failure mode

(ii) O: understanding the likelihood of failure modes occur during the process

(iii) D: detecting failure mode before the patient faces failure

(iv) C: determining the amount of cost that can be wasted in each failure mode

(v) T: considering the urgency of time in each failure mode

Within the risk assessment range, the values of SODCT factors are defined using Table 6. Then, ten primary failure modes have been determined by the FMEA team through brainstorming (see Table 7). Afterward, the value of each risk factor for each failure mode is specified by experts. In this step, experts are asked to express the values of uncertainty and reliability using linguistic variables in order to use them in the following stages. In the second stage, the weights of risk factors are determined using the Z-SWARA method, introduced in Section 2.3, in an uncertain environment. In the third stage, the Z-WASPAS method, introduced in Section 2.4, is used to prioritize the identified failure modes with the aim of providing a distinct prioritization. To put it precisely, using the developed methods in the second and third stages can tackle some of the main disadvantages of the conventional FMEA technique, such as not assigning different weights to risk factors and not considering uncertainty and reliability in the process of determining the weight of factors and prioritizing failures. Finally, the results of the proposed approach are evaluated in comparison with conventional FMEA and fuzzy-WASPAS. A summary of the research framework has been presented in Figure 1.

4. Analysis of the Results

4.1. Proposed Approach Implementation. In this section, the results of implementing the proposed approach for assessing the influential failure modes in the MEs are presented. Initially, the value of each risk factor is determined by the FMEA team, and the results are presented in Table 8.

Then, according to FMEA team judgments, the values of SODCT factors for each failure mode are determined in the form of linguistic variables. These linguistic values of SODCT factors for ten identified failures are presented in Table 9.

In the second step of the proposed approach, for assigning weights to SODCT factors, the Z-SWARA method is used. For this purpose, first, experts in each group considered the most important criteria and then ordered other criteria in a descending form, from higher to lower importance, in comparison with the higher-ranked criteria. Thereafter, the order of each criterion is explained using linguistic variables and transformed into the form of Z-number based on Table 1 and Table 2. The results in the form of linguistic variables are indicated in Table 10. According to Table 5, these values are transformed to TFN, and then following the steps of Z-SWARA in Section 2.2, the weights of SODCT factors are calculated and summarized in Table 11.

To prioritize failure modes based on the Z-WASPAS method, initially, the decision-making matrix H of the Z-WASPAS method in the form of Z-numbers' components (considering reliability) was formed, where rows represent failure modes, and columns represent SODCT factors. Then, the decision-making matrix transforms into a matrix of TFN H, and consequently, $\widetilde{Q}_i$ and $\widetilde{P}_i$ are obtained. The results are shown in Table 12.

In Table 13, the average value of $Q_i$and P_i, the utility value K_i, and the final rank of failure modes are presented. The value of λ is obtained equal to 0.48.

Based on Table 14, in the conventional FMEA method, F_5 (entering patients ID numbers manually instead of scanning from bracelet) with RPN = 8640 is at the top of the priority list to investigate in MEs occurrence. Furthermore,

Table 6: Traditional ratings for SODCT factors [55].

Rating	S	O	D	C	T
10	Hazardous with warning	Very high	Absolute	Repair cost close to the original price	Repair time extremely high
9	Hazardous without warning	Almost inevitable failure	Uncertainty	Repair cost extremely high	
8	Very high	High	High	Repair cost high	Repair time high
7	High	Repeated failures	Repeated failures		
6	Moderate	Moderate: occasional failures	Moderate: occasional failures	Repair cost moderately high	Repair time moderate
5	Low			Repair cost moderate	
4	Very low			Repair cost relatively low	
3	Minor	Low	Low	Repair cost low	Repair time low
2	Very minor	Relatively few failures	Relatively few failures	Repair cost very low	
1	None	Remote: failure is unlikely	Remote: failure is unlikely	Repair at nearly no cost	Repair cost very low

Table 7: Significant failure modes result in MEs.

Failure modes	Causes	Effects
	Medication orders' confirmation	
F_1: unable to verify medication orders	Absence of proper electronic health record (EHR).	Unable to know the dose that was given before.
	Print medication list from electronic record	
F_2: inaccessibility to EHR	Busy timetable, inappropriate communication.	Extra dose given.
	Getting medication from pharmacy	
F_3: mislabeling and incorrect medication dispensed in shelves in pharmacy	Misreading labels and incorrect stock of medications in pharmacy.	Giving incorrect medication, incorrect dose, or spending more time for giving the correct medication.
F_4: unable to verify the correctness of given dose	Absence of supporting documentation to prove if patient received the dose before.	Incorrect medication administration.
	Scan patient identification (ID) numbers	
F_5: assigning incorrect ID number to patient	ID band is not scanned.	Medications might be administered to the wrong patient.
F_6: inappropriate scanning	Busy timetable, lack of knowledge about medication administration.	Incorrect medication administration
	Scan medication barcode and administrate	
F_7: wrong medication or wrong time of administration process in pharmacy.	Mislabeling in pharmacy or physician changes the medication order.	Patient may not take the correct medication or receive medication at the right time.
F_8: medication may be administered incorrectly via wrong route or dosage.	Wrongly reading order. Unfamiliarity with medicine.	The patient is negatively affected through incorrect route of medication.
F_9: system overriding by manually entering medication barcode of medicines' containers.	Mislabeling of container. System malfunction.	Receiving the wrong medication, dose, or medication administration.
	Prepare medications	
F_{10}: medications may be prepared wrongly.	Incomprehensible medication label or physician order.	Patient may receive the incorrect dosage of medication or take the medication via the wrong route.
Incorrect dosage or incorrect route.	Not double-checking order previous to preparation.	

F_8 (medication may be administered incorrectly via wrong route or dosage) is the second priority with RPN = 8100, and F_3 (mislabeling and incorrect medication dispensed in shelves in pharmacy) is the third with RPN = 5670. The study of failures priority is indicated that based on conventional FMEA, ten failures have categorized into nine groups; the reason is the repetition of rank four in F_4 (inability to verify the correctness of the given dose) and F_6 (inappropriate ID scanning). This issue indicates that prioritizing, based on a traditional index, has not been perfectly done, and this issue results in the confusion of experts and decision-makers for planning corrective and preventive actions. Incomplete prioritizing of failure modes is the result of overlooking the weight of SODCT factors and uncertainty in ranking alternatives (see Table 14). In fuzzy-WASPAS and the proposed *Z*-WASPAS method, failure modes with greater K_i value are the main influential factors and rank in higher priorities. Utilizing the fuzzy-WASPAS method leads to improving the

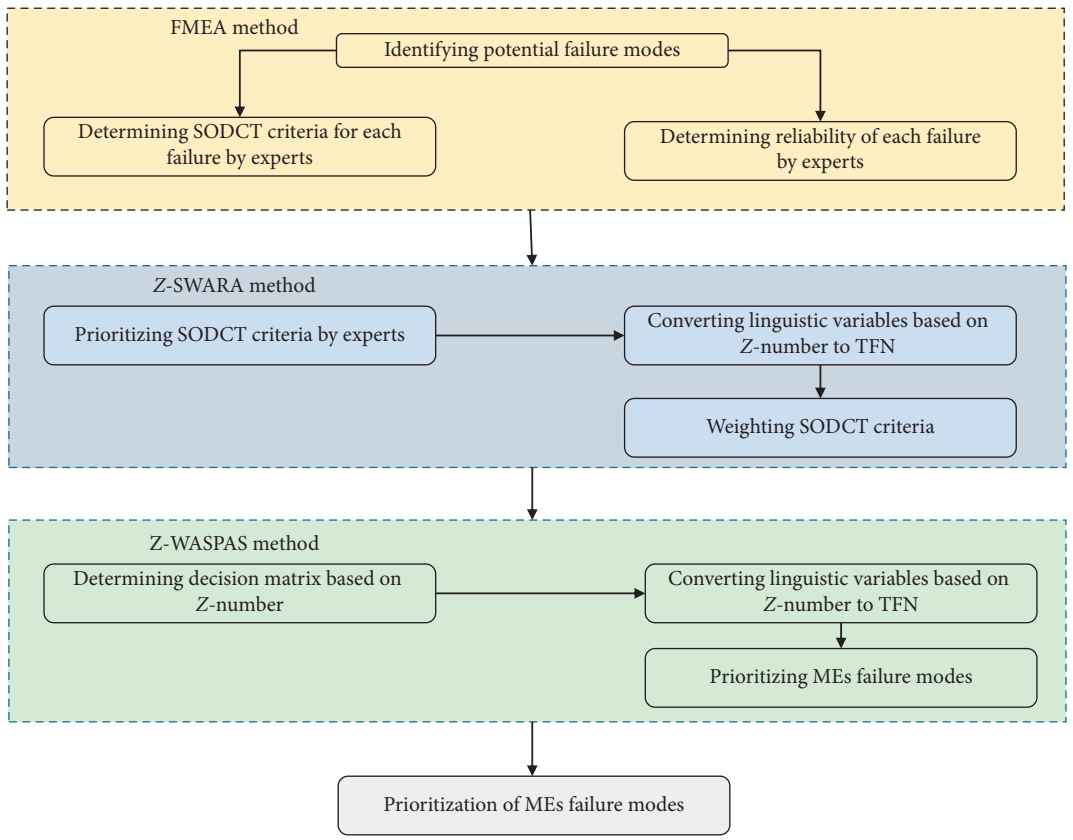


Figure 1: Proposed research approach for prioritizing MEs failure modes.

Table 8: Scoring risk factors based on FMEA team.

Failure mode	S			O			D			C			T		
	TM_1	TM_2	TM_3	TM_1	TM_2	TM_3	TM_1	TM_2	TM_3	TM_1	TM_2	TM_3	TM_1	TM_2	TM_3
F_1	3	4	5	4	5	5	1	1	2	5	3	4	4	3	5
F_2	2	2	1	5	3	5	3	3	2	4	4	5	5	4	6
F_3	8	7	5	3	3	4	9	7	6	9	8	9	4	5	6
F_4	4	4	5	7	5	6	4	3	4	6	6	7	3	4	5
F_5	8	8	9	5	6	5	8	9	10	6	4	6	4	3	5
F_6	5	4	4	6	4	7	4	5	4	4	5	6	6	5	3
F_7	2	2	3	4	2	3	10	8	10	3	2	4	5	3	4
F_8	9	7	9	4	4	3	7	10	8	5	4	5	5	5	4
F_9	8	7	8	4	3	4	5	8	7	4	4	5	3	4	2
F_{10}	7	7	7	2	3	2	4	3	5	6	4	5	4	5	4

incomplete prioritizing problem of conventional FMEA in which the failure modes have been categorized into ten groups. Based on the fuzzy-WASPAS method, F_8 (medication may be prepared wrongly via incorrect route or dosage) with $K_i = 1.070$, F_5 with $K_i = 1.046$, and F_3 with $K_i = 1.020$ are in first, second, and third priorities, respectively. Although the certainty is considered in weighting SODCT factors by the fuzzy-WASPAS method, the reliability has not contributed to the values of failure modes. Concerning the fact that reliability plays an important role in determining the most important failure modes in MEs, the proposed *Z*-WASPAS approach aims to synthesize the reliability of rating failure modes based on the expert's identification. In other words, including the reliability in the process of decision-making about significant failure modes can bring real results that are closer to reality and less injury to the patients. In the *Z*-WASPAS method, F_8 with $K_i = 1.01$, F_5 with $K_i = 1.00$, and F_3 with $K_i = 0.98$ are in the first, second, and third priority, respectively.

4.2. Results Comparison. In this section, the outputs of the proposed approach have been compared with other similar methods to demonstrate its applicability. According to Table 14, in the fuzzy-WASPAS and *Z*-WASPAS method, the most significant failure mode is F_8, whereas in the conventional FMEA, F_5 is the most crucial failure mode. The main reason for this change is the nature of F_8, in which there is considerable uncertainty about the wrong preparation of

Table 9: The linguistic variable for the SODCT factors for each failure mode.

Risk factor	Teams	Failure modes									
		F_1	**F_2**	**F_3**	**F_4**	**F_5**	**F_6**	**F_7**	**F_8**	**F_9**	**F_{10}**
S	**TM_1**	(ML, H)	(ML, M)	(H, VH)	(ML, M)	(EH, M)	(ML, VH)	(ML, VH)	(MH, H)	(H, M)	(MH, H)
	TM_2	(M, M)	(L, VH)	(MH, M)	(MH, VH)	(H, H)	(M, M)	(L, H)	(H, M)	(MH, H)	(M, VH)
	TM_3	(M, H)	(L, H)	(MH, M)	(M, M)	(H, M)	(M, VH)	(ML, M)	(VH, M)	(MH, M)	(MH, M)
O	**TM_1**	(M, M)	(M, M)	(L, H)	(M, H)	(M, VH)	(M, H)	(L, M)	(M, H)	(ML, VH)	(L, H)
	TM_2	(M, VH)	(MH, VH)	(ML, VH)	(M, VH)	(ML, H)	(MH, H)	(M, VH)	(ML, VH)	(ML, H)	(ML, H)
	TM_3	(MH, M)	(MH, H)	(M, M)	(MH, H)	(MH, M)	(M, M)	(M, H)	(M, M)	(M, M)	(L, VH)
D	**TM_1**	(L, H)	(M, M)	(MH, M)	(ML, H)	(MH, H)	(MH, M)	(H, H)	(MH, M)	(MH, VH)	(M, H)
	TM_2	(L, VH)	(ML, H)	(H, M)	(ML, M)	(ML, VH)	(M, VH)	(H, M)	(H, VH)	(M, H)	(MH, VH)
	TM_3	(ML, VH)	(L, H)	(H, VH)	(L, VH)	(MH, H)	(M, H)	(MH, VH)	(VH, H)	(MH, VH)	(M, H)
C	**TM_1**	(M, H)	(ML, VH)	(MH, H)	(MH, H)	(M, H)	(ML, H)	(M, H)	(ML, H)	(M, M)	(ML, VH)
	TM_2	(MH, M)	(ML, M)	(ML, M)	(MH, VH)	(MH, VH)	(ML, VH)	(M, H)	(M, H)	(MH, M)	(M, H)
	TM_3	(M, M)	(M, M)	(M, H)	(M, H)	(M, H)	(ML, VH)	(ML, M)	(MH, M)	(M, M)	(ML, M)
T	**TM_1**	(ML, H)	(M, H)	(MH, VH)	(ML, M)	(ML, M)	(M, H)	(ML, VH)	(M, H)	(L, H)	(ML, H)
	TM_2	(M, VH)	(M, VH)	(M, H)	(M, M)	(ML, M)	(MH, VH)	(M-M)	(MH, VH)	(ML, M)	(ML, VH)
	TM_3	(ML, H)	(M-H)	(MH-H)	(MH, VH)	(M-H)	(MH, VH)	(M-H)	(M-H)	(L-M)	(M-H)

Table 10: Prioritizing the SODCT factors based on their importance in TMs view.

TM_1		TM_2		TM_3	
S		*S*		*C*	
C	(MOL, H)	***C***	(MOL, VH)	***S***	(VLI, M)
O	(VLI, H)	***T***	(LI, M)	***T***	(MOL, VH)
T	(MUL, M)	***D***	(VLI, H)	***O***	(LI, H)
D	(LI, VH)	***O***	(MUL, VH)	***D***	(MUL, H)

Table 11: Final weights of SODCT factors with *Z*-SWARA method.

Risk factor	**TM_1**			**TM_2**			**TM_3**			Final weight		
	a	**B**	**c**	**a**	**b**	**C**	**a**	**B**	**c**	**a**	**b**	**c**
S	0.329	0.362	0.404	0.359	0.412	0.484	0.152	0.211	0.289	0.280	0.328	0.392
C	0.257	0.293	0.336	0.148	0.211	0.296	0.342	0.388	0.450	0.249	0.297	0.361
O	0.106	0.151	0.205	0.059	0.099	0.153	0.094	0.141	0.201	0.086	0.130	0.187
T	0.068	0.106	0.154	0.100	0.156	0.230	0.114	0.165	0.232	0.094	0.143	0.205
D	0.055	0.088	0.130	0.075	0.122	0.185	0.058	0.095	0.146	0.062	0.102	0.154

medication via incorrect route or dosage failure mode, among decision-makers (see Table 8). Assigning a weight to the SODCT factors and contributing uncertainty and unreliability in fuzzy-WASPAS and *Z*-WASPAS resulted in F_8 being known as the main failure mode. The importance of F_8 is the principal reason for this contradiction, which is determined by considering reliability. The importance of this failure mode was approved in several papers [56, 57]; any error occurrence in medication preparation can cause irreversible results for the patients or fatality.

F_3 in all the methods ranks third place, which means that this failure mode has not been affected by weight changes, and the uncertainty is minimum about this failure mode based on experts' decision because of any mistake in stock of medication by the pharmacy can result in incorrect medication. The other failure modes affected by involving weight, uncertainty, and reliability and their ranks changed between four to ten.

The weight that is assigned to the S factor can have a considerable effect on the ranking since the severity plays an important role in MEs prevention. For instance, the fourth rank attributed to the F_4 in conventional FMEA, but in fuzzy-WASPAS, the rank of this failure mode, one unit increased and stayed in the fifth rank. The main reason for this increment is the contributing weights of factors and certainty. However, when the reliability is involved in the ranking process, in the *Z*-WASPAS method, the ranking result for the F_4 is the fourth alternative. Figure 2 shows the comparisons of prioritizing failure modes according to three different methods.

One of the main problems that decision-makers face when using the conventional FMEA technique is assigning similar scores to different failure modes. In this case, decision-makers cannot properly identify critical failure modes and take corrective action to reduce their negative effects due to limited organizational resources. Therefore, this study tried to present the results with high separability compared to the FMEA by developing this conventional method based on the SWARA and WASPAS methods and using the *Z*-

Table 12: Aggregated weighted normalized decision matrix of WSM (WPN).

Failure modes	*S*			*O*			*D*			*C*			T		
	a	**b**	**c**	**A**	**b**	**c**	**A**	**b**	**c**	**a**	**b**	**C**	**a**	**b**	**C**
$\mathbf{F_1}$	0.08 (0.62)	0.17 (0.80)	0.32 (0.98)	0.05 (0.87)	0.11 (0.97)	0.22 (1.07)	0 (0.71)	0.02 (0.85)	0.06 (0.94)	0.11 (0.74)	0.23 (0.92	0.45 (1.13)	0.03 (0.79)	0.08 (0.92)	0.18 (1.03)
$\mathbf{F_2}$	0.01 (0.32)	0.06 (0.58)	0.19 (0.83)	0.07 (0.92)	0.13 (1.0)	0.25 (1.1)	0.01 (0.8)	0.03 (0.88)	0.07 (0.96)	0.05 (0.59)	0.15 (0.81)	0.35 (1.04)	0.05 (0.85)	0.11 (0.96)	0.22 (1.065)
$\mathbf{F_3}$	0.2 (0.85)	0.3 (0.97)	0.47 (1.12)	0.02 (0.77)	0.06 (0.9)	0.15 (1)	0.07 (0.95)	0.07 (0.99)	0.14 (1.03)	0.1 (0.72)	0.220.91	0.45 (1.13)	0.07 (0.9)	0.14 (0.99)	0.27 (1.09)
$\mathbf{F_4}$	0.11 (0.7)	0.2 (0.85)	0.36 (1.03)	0.06 (0.9)	0.12 (0.99)	0.24 (1.08)	0.01 (0.75)	0.03 (0.86)	0.06 (0.95)	0.16 (0.82)	0.3 (1.0)	0.58 (1.22)	0.05 (0.85)	0.1 (0.95)	0.21 (1.05)
$\mathbf{F_5}$	0.23 (0.89)	0.33 (1.0)	0.48 (1.13)	0.05 (0.87)	0.1 (0.96)	0.21 (1.6)	0.04 (0.9)	0.07 (0.96)	0.12 (1.01)	0.13 (0.78)	0.27 (0.97)	0.53 (1.19)	0.02 (0.77)	0.07 (0.90)	0.16 (1.01)
$\mathbf{F_6}$	0.09 (0.64)	0.18 (0.82)	0.34 (1.01)	0.06 (0.89)	0.11 (0.98)	0.22 (1.07)	0.04 (0.9)	0.07 (0.95)	0.11 (1.0)	0.04 (0.53)	0.15 (0.81)	0.36 (1.05)	0.08 (0.91)	0.15 (1.0)	0.28 (1.1)
$\mathbf{F_7}$	0.02 (0.42)	0.1 (0.66)	0.23 (0.88)	0.03 (0.83)	0.08 (0.93)	0.18 (1.04)	0.07 (0.95)	0.1 (0.99)	0.14 (1.03)	0.08 (0.67)	0.19 (0.87)	0.4 (1.09)	0.04 (0.81)	0.09 (0.93)	0.19 (1.04)
$\mathbf{F_8}$	0.23 (0.88)	0.32 (0.98)	0.46 (1.11)	0.04 (0.84)	0.09 (0.95	0.19 (1.05)	0.08 (0.97)	0.11 (1.0)	0.14 (1.03)	0.09 (0.70)	0.21 (0.90)	0.43 (1.11)	0.06 (0.88)	0.13 (0.98)	0.25 (1.08)
$\mathbf{F_9}$	0.19 (0.83)	0.28 (0.95)	0.44 (1.10)	0.02 (0.80)	0.07 (0.92	0.17 (1.03)	0.05 (0.93)	0.08 (0.98)	0.14 (1.03)	0.11 (0.73)	0.22 (0.90)	0.43 (1.11)	0 (0.6)	0.031 (0.79)	0.1 (0.94)
$\mathbf{F_{10}}$	0.16 (0.78)	0.26 (0.92)	0.43 (1.09)	0 (0.65)	0.04 (0.84)	0.12 (0.98)	0.04 (0.91)	0.07 (0.96)	0.12 (1.01)	0.06 (0.61)	0.17 (0.83)	0.37 (1.07)	0.03 (0.78)	0.08 (0.92)	0.18 (1.03)

Table 13: Final result of ranking failure modes for $\lambda = 0.48$

Failure modes	$\mathbf{Q_i}$	$\mathbf{P_i}$	$\mathbf{K_i}$	Rank
$\mathbf{F_1}$	0.709	0.648	0.677	8
$\mathbf{F_2}$	0.593	0.492	0.541	10
$\mathbf{F_3}$	0.918	0.874	0.895	3
$\mathbf{F_4}$	0.866	0.788	0.825	4
$\mathbf{F_5}$	0.942	0.893	0.917	2
$\mathbf{F_6}$	0.760	0.710	0.734	5
$\mathbf{F_7}$	0.647	0.583	0.614	9
$\mathbf{F_8}$	0.941	0.904	0.922	1
$\mathbf{F_9}$	0.781	0.703	0.740	6
$\mathbf{F_{10}}$	0.710	0.659	0.683	7

Table 14: Failure modes prioritization using the proposed approach compared to other methods.

Failure modes	Conventional FMEA		Fuzzy-WASPAS		*Z*-WASPAS	
	RPN	Rank	$\mathbf{K_i}$	Rank	$\mathbf{K_i}$	Rank
$\mathbf{F_1}$	800	8	0.791	6	0.74	8
$\mathbf{F_2}$	600	9	0.605	10	0.60	10
$\mathbf{F_3}$	5670	3	1.020	3	0.98	3
$\mathbf{F_4}$	3600	4	0.908	5	0.91	4
$\mathbf{F_5}$	8640	1	1.046	2	1.00	2
$\mathbf{F_6}$	3600	4	0.764	7	0.81	5
$\mathbf{F_7}$	1215	6	0.685	9	0.68	9
$\mathbf{F_8}$	8100	2	1.070	1	1.01	1
$\mathbf{F_9}$	3360	5	0.917	4	0.81	6
$\mathbf{F_{10}}$	1120	7	0.758	8	0.75	7

number theory. Besides, applying the concepts of uncertainty and reliability simultaneously in the *Z*-WASPAS method compared to the fuzzy-WASPAS can lead to more realistic results. In conclusion, based on FMEA methodology and the opinion of three experts (doctor, pharmacist, and nurse), F_5 or manually entering the patient's ID and not scanning the patient bracelet is the principal reason of the MEs inducement, which would result in identifying the wrong ID or misuse of the medicine by the patient. However, in the proposed method, by involving reliability factors in identifying failure modes, F_8 or medication administration error is known as a critical failure mode. Wu et al. [56] also confirmed the importance of this failure mode in MEs occurrence. The results also point to the importance of F_8 as the main failure

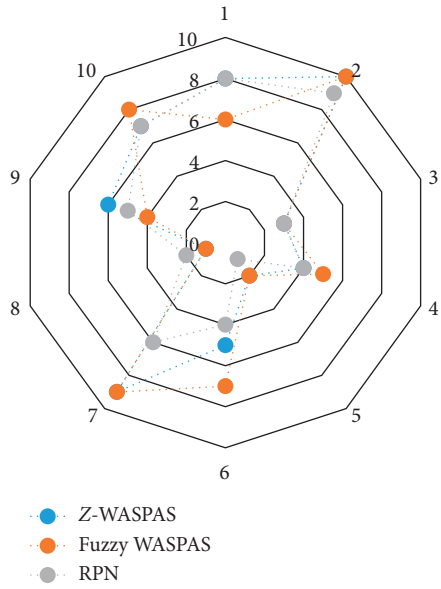


Figure 2: Failure modes comparative ranking for the three different methods.

Table 15: Weights and crisp value of weights of SODCT factors in different cases.

Risk factor	Case $_0$	Case $_1$	Case $_2$	Case $_3$	Case $_4$
$\mathbf{w_S}$	(0.28, 0.33, 0.39)	(0.26, 0.30, 0.35)	(0.22, 0.33, 0.29)	(0.53, 0.38, 0.48)	(0.02, 0.12, 0.09)
α_{w_S}	0.333	0.3	0.28	0.46	0.08
$\mathbf{w_O}$	(0.09, 0.13, 0.19)	(0.09, 0.06, 0.14)	(0.02, 0.05, 0.09)	(0.02, 0.01, 0.03)	(0.20.12, 0.26)
α_{w_O}	0.134	0.1	0.05	0.02	0.19
$\mathbf{w_D}$	(0.06, 0.10, 0.15)	(0.26, 0.15, 0.35)	(0.29, 0.23, 0.32)	(0.12, 0.08, 0.14)	(0.35, 0.25, 0.28)
α_{w_D}	0.106	0.25	0.28	0.11	0.29
$\mathbf{w_C}$	(0.25, 0.30, 0.36)	(0.29, 0.23, 0.39)	(0.15, 0.01, 0.17)	(0.39, 0.20, 0.29)	(0.25, 0.32, 0.29)
α_{w_C}	0.302	0.3	0.11	0.29	0.29
$\mathbf{w_T}$	(0.09, 0.14, 0.21)	(0.02, 0.08, 0.05)	(0.22, 0.29, 0.32)	(0.10, 0.09, 0.18)	(0.20, 0.09, 0.15)
α_{w_T}	0.147	0.05	0.28	0.12	0.15

mode. The proposed corrective action is to ensure the information of the medicine, including dosage, route, and other administration specifics, that are immediately available in barcode medication administration. Moreover, F_5 (entering patients identification numbers manually instead of scanning from bracelet) and F_3 (incorrect medication dispensed in the drawer, or refrigerator, and mislabeling in pharmacy) were introduced as second and third main failure modes in MEs. Avoiding entering ID number unless approved by the charge nurse manually is the corrective action of F_5 and having more than one pharmacy technician dispense medications at each section to double-check that medications are correctly dispensed is the corrective action for F_3.

4.3. Sensitivity Analysis. A sensitivity analysis is administered by transforming the weight values of criteria in five different cases (see Table 15). Case $_0$ represents the original crisp weights of the criteria that are obtained based on the *Z*-SWARA method in this research. For evaluating how the rank of alternatives changes in possible conditions, the crisp value of weights is assigned to the SODCT factors and made Case $_1$ to Case $_4$. The result of sensitivity analysis for ranking outcomes of ten failure modes and different cases is shown in Table 16. In this paper, the aggregated decision of three groups of decision-makers indicated that the order of importance in SODCT factors is *S*, *C*, *D*, *T*, and *O*, respectively. For instance, factor *S* has a significant impact on MEs' control in comparison with *C*. According to Table 16, F_8 is the failure mode with the highest risk priority in all cases. Considering the SODCT factors, in Case $_0$, Case $_1$, and Case $_3$, F_5 is the second important failure mode because of the high weight of *S*, while in Case $_2$ and Case $_4$, F_3 is the second significant failure mode because of the lower weight of *S*. This comparison is valid for other criteria and failure modes. The sensitivity analysis indicates that the weight of criteria can have a significant influence on the final ranking orders of failure modes. Therefore, determining the acceptable weight for criteria, according to the real situation, is of importance and advantage to the risk prioritization of failure modes and the subsequent corrective actions.

As stated, this study attempted to propose an extended approach using FMEA, SWARA, and WASPAS methods to prioritize MEs failure modes. In this approach, new factors of cost and time based on the case study in addition to traditional factors were added in comparison with the FMEA

Table 16: Ranking results of failure modes with respect to the different cases.

Failure modes	Case $_0$	Case $_1$	Case $_2$	Case $_3$	Case $_4$
$\mathbf{F_1}$	8	9	9	8	9
$\mathbf{F_2}$	10	10	10	10	10
$\mathbf{F_3}$	3	3	2	3	2
$\mathbf{F_4}$	4	5	8	5	6
$\mathbf{F_5}$	2	2	3	2	3
$\mathbf{F_6}$	5	7	4	7	4
$\mathbf{F_7}$	9	8	7	9	5
$\mathbf{F_8}$	1	1	1	1	1
$\mathbf{F_9}$	6	4	6	4	7
$\mathbf{F_{10}}$	7	6	5	6	8

technique. Besides, this study tried to assign different weights to risk factors using the developed *Z*-SWARA method. This method has fewer pairwise comparisons compared to other conventional methods, such as the AHP method, and this research has considered the concepts of uncertainty and reliability in the process of determining the weights of risk factors simultaneously using the *Z*-SWARA method. In other methods like AHP or ANP, the model is created based on criteria and experts' evaluations that will affect priorities and ranks. So, SWARA can be useful for some issues whose priorities are known in advance according to situations [57]. Also, by developing the WASPAS method based on the *Z*-number theory and using it in the proposed approach, a more distinct prioritization of failures compared to the traditional RPN score has been presented.

5. Conclusion

One of the major causes of injury to patients when providing medical services is related to MEs. Identifying the MEs and reducing the likelihood of their occurrence is very important in order to increase the patient's level of safety. In this regard, using a decision-making approach based on the FMEA, a popular technique in this field, can help decision-makers identify and prioritize these errors. However, since this technique has some major drawbacks, this study has developed its proposed decision-making approach based on the *Z*-number theory. The proposed decision-making approach incorporates *Z*-SWARA and *Z*-WASPAS methods with the FMEA technique to address some disadvantages of the RPN score. In this research, by implementing and comparing the proposed approach with the conventional FMEA and fuzzy-WASPAS methods, the results indicated that prioritizing failure modes with the proposed method is closer to reality because of reliability factor involvement. On the other hand, the decision-makers can provide a series of appropriate corrective/preventive measures for critical failures, implement the corrective actions with the relevant departments, and perform a reassessment to examine the new situation and the effectiveness of these measures. Notably, failing to observe the cause-and-effect relation of failure modes is the main limitation of this study. Besides, not considering importance-necessity and uncertainty in decision making and addressing the relative importance among experts are the other issues that can be considered in future investigation using R-number, G-number, and evidence theories, respectively.

Data Availability

The data are extracted by experts sorting criteria, from the most important to less important in descending order, based on self-identification.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this paper.

References

[1] W. B. Tshiamo, M. Kgositau, E. Ntsayagae, and M. B. Sabone, "The role of nursing education in preventing medication errors in Botswana," *International Journal of Africa Nursing Sciences*, vol. 3, pp. 18–23, 2015.

[2] B. A. Da Silva and M. Krishnamurthy, "The alarming reality of medication error: a patient case and review of Pennsylvania and national data," *Journal of Community Hospital Internal Medicine Perspectives*, vol. 6, no. 4, p. 31758, 2016.

[3] S. M. Handler, S. Perera, E. F. Olshansky et al., "Identifying modifiable barriers to medication error reporting in the nursing home setting," *Journal of the American Medical Directors Association*, vol. 8, no. 9, pp. 568–574, 2007.

[4] T. A. Edgar, D. S. Lee, and D. D. Cousins, "Experience with a national medication error reporting program," *American Journal of Health-System Pharmacy*, vol. 51, no. 10, pp. 1335–1338, 1994.

[5] A. Samsiah, N. Othman, S. Jamshed, M. A. Hassali, and W. M. Wan-Mohaina, "Medication errors reported to the national medication error reporting system in Malaysia: a 4-year retrospective review (2009 to 2012)," *European Journal of Clinical Pharmacology*, vol. 72, no. 12, pp. 1515–1524, 2016.

[6] P. Nishtala, H. Ndukwe, T. Y. Chyou, M. Salahudeen, and S. Narayan, "An overview of pharmacoepidemiology in New Zealand: medical databases, registries and research achievements," *The New Zealand Medical Journal*, vol. 130, no. 1449, pp. 52–61, 2017.

[7] Y. Chen, X. Wu, Z. Huang et al., "Evaluation of a medication error monitoring system to reduce the incidence of medication errors in a clinical setting," *Research in Social and Administrative Pharmacy*, vol. 15, no. 7, pp. 883–888, 2019.

[8] D. M. Benjamin, "Reducing medication errors and increasing patient safety: case studies in clinical pharmacology," *The Journal of Clinical Pharmacology*, vol. 43, no. 7, pp. 768–783, 2003.

[9] M. A. Onari, S. Yousefi, and M. J. Rezaee, "Risk assessment in discrete production processes considering uncertainty and reliability: *Z*-number multi-stage fuzzy cognitive map with fuzzy learning algorithm," *Artificial Intelligence Review*, vol. 54, no. 3, pp. 1–35, 2020.

[10] M. J. Rezaee, S. Yousefi, M. Valipour, and M. M. Dehdar, "Risk analysis of sequential processes in food industry integrating multi-stage fuzzy cognitive map and process failure mode and effects analysis," *Computers & Industrial Engineering*, vol. 123, pp. 325–337, 2018.

[11] H.-C. Liu, "FMEA for proactive healthcare risk analysis: a systematic literature review," *Improved FMEA Methods for Proactive Healthcare Risk Analysis*, vol. 1, pp. 15–45, 2019.

[12] M. J. Rezaee, S. Yousefi, M. Eshkevari, M. Valipour, and M. Saberi, "Risk analysis of health, safety and environment in chemical industry integrating linguistic FMEA, fuzzy inference system and fuzzy DEA," *Stochastic Environmental Research and Risk Assessment*, vol. 34, no. 1, pp. 201–218, 2020.

[13] R. Dabbagh and S. Yousefi, "A hybrid decision-making approach based on FCM and MOORA for occupational health and safety risk analysis," *Journal of Safety Research*, vol. 71, pp. 111–123, 2019.

[14] H.-C. Liu, L. Liu, and N. Liu, "Risk evaluation approaches in failure mode and effects analysis: a literature review," *Expert Systems with Applications*, vol. 40, no. 2, pp. 828–838, 2013.

[15] H.-C. Liu, X.-Q. Chen, C.-Y. Duan, and Y.-M. Wang, "Failure mode and effect analysis using multi-criteria decision making methods: a systematic literature review," *Computers & Industrial Engineering*, vol. 135, pp. 881–897, 2019.

[16] M. Baghery, S. Yousefi, and M. J. Rezaee, "Risk measurement and prioritization of auto parts manufacturing processes based on process failure analysis, interval data envelopment analysis and grey relational analysis," *Journal of Intelligent Manufacturing*, vol. 29, no. 8, pp. 1803–1825, 2018.

[17] Y. Ozdemir, M. Gul, and E. Celik, "Assessment of occupational hazards and associated risks in fuzzy environment: a case study of a university chemical laboratory," *Human and Ecological Risk Assessment: An International Journal*, vol. 23, no. 4, pp. 895–924, 2017.

[18] S. Dorosti, M. Fathi, S. J. Ghoushchi, M. Khakifirooz, and M. Khazaeili, "Patient waiting time management through fuzzy based failure mode and effect analysis," *Journal of Intelligent & Fuzzy Systems*, vol. 38, pp. 1–12, 2020.

[19] M. Gul, M. Yucesan, and E. Celik, "A manufacturing failure mode and effect analysis based on fuzzy and probabilistic risk analysis," *Applied Soft Computing*, vol. 96, Article ID 106689, 2020.

[20] S. Yousefi, M. Jahangoshai Rezaee, and A. Moradi, "Causal effect analysis of logistics processes risks in manufacturing industries using sequential multi-stage fuzzy cognitive map: a case study," *International Journal of Computer Integrated Manufacturing*, vol. 33, no. 10-11, pp. 1055–1075, 2020.

[21] D. Wu and Y. Tang, "An improved failure mode and effects analysis method based on uncertainty measure in the evidence theory," *Quality and Reliability Engineering International*, vol. 36, no. 5, pp. 1786–1807, 2020.

[22] H. Seiti, A. Hafezalkotob, and L. Martínez, "*R*-numbers, a new risk modeling associated with fuzzy numbers and its application to decision making," *Information Sciences*, vol. 483, pp. 206–231, 2019.

[23] S. J. Ghoushchi and M. Khazaeili, "*G*-numbers: importance-necessity concept in uncertain environment," *International Journal of Management and Fuzzy Systems*, vol. 5, no. 1, pp. 27–32, 2019.

[24] V. Başhan, H. Demirel, and M. Gul, "An FMEA-based TOPSIS approach under single valued neutrosophic sets for maritime risk evaluation: the case of ship navigation safety," *Soft Computing*, vol. 24, no. 24, pp. 18749–18764, 2020.

[25] S. J. Ghoushchi, S. Yousefi, and M. Khazaeili, "An extended FMEA approach based on the *Z*-MOORA and fuzzy BWM for prioritization of failures," *Applied Soft Computing*, vol. 81, Article ID 105505, 2019.

[26] A. Mardani, R. E. Hooker, S. Ozkul et al., "Application of decision making and fuzzy sets theory to evaluate the healthcare and medical problems: a review of three decades of research with recent developments," *Expert Systems with Applications*, vol. 137, pp. 202–231, 2019.

[27] M.-C. Hsieh, E. M.-Y. Wang, W.-C. Lee et al., "Application of HFACS, fuzzy TOPSIS, and AHP for identifying important human error factors in emergency departments in Taiwan," *International Journal of Industrial Ergonomics*, vol. 67, pp. 171–179, 2018.

[28] Y. Trakulsunti and J. Antony, "Can lean six sigma be used to reduce medication errors in the health-care sector?" *Leadership in Health Services*, vol. 31, no. 4, pp. 426–433, 2018.

[29] D. G. Ford, A. L. Seybert, P. L. Smithburger, L. R. Kobulinsky, J. T. Samosky, and S. L. Kane-Gill, "Impact of simulation-based learning on medication error rates in critically ill patients," *Intensive Care Medicine*, vol. 36, no. 9, pp. 1526–1531, 2010.

[30] C. de Leon, A. Silva, L. Ribeiro, L. Guarda, and L. Fonseca, "Development and validation of clinical cases to be used in maternal-child nursing education," *Revista de Enfermagem Referência*, vol. IV, no. 18, pp. 51–62, 2018.

[31] M. Härkänen, J. Ahonen, M. Kervinen, H. Turunen, and K. Vehviläinen-Julkunen, "The factors associated with medication errors in adult medical and surgical inpatients: a direct observation approach with medication record reviews," *Scandinavian Journal of Caring Sciences*, vol. 29, no. 2, pp. 297–306, 2015.

[32] J. Maiden, J. M. Georges, and C. D. Connelly, "Moral distress, compassion fatigue, and perceptions about medication errors in certified critical care nurses," *Dimensions of Critical Care Nursing*, vol. 30, no. 6, pp. 339–345, 2011.

[33] S. Yousefi, A. Alizadeh, J. Hayati, and M. Baghery, "HSE risk prioritization using robust DEA-FMEA approach with undesirable outputs: a study of automotive parts industry in Iran," *Safety Science*, vol. 102, pp. 144–158, 2018.

[34] M. J. Rezaee, A. Salimi, and S. Yousefi, "Identifying and managing failures in stone processing industry using cost-based FMEA," *The International Journal of Advanced Manufacturing Technology*, vol. 88, no. 9-12, pp. 3329–3342, 2017.

[35] V. Keršuliene, E. K. Zavadskas, and Z. Turskis, "Selection of rational dispute resolution method by applying new step-wise weight assessment ratio analysis (SWARA)," *Journal of Business Economics and Management*, vol. 11, no. 2, pp. 243–258, 2010.

[36] S. H. Zolfani, M. H. Aghdaie, A. Derakhti, E. K. Zavadskas, and M. H. M. Varzandeh, "Decision making on business issues with foresight perspective; an application of new hybrid MCDM model in shopping mall locating," *Expert Systems with Applications*, vol. 40, no. 17, pp. 7111–7121, 2013.

[37] Y. Deng, "*D*-numbers: theory and applications," *Journal of Information & Computational Science*, vol. 9, no. 9, pp. 2421–2428, 2012.

[38] R. Akbari, R. Dabbagh, and S. J. Ghoushchi, "HSE risk prioritization of molybdenum operation process using extended FMEA approach based on fuzzy BWM and *Z*-WASPAS," *Journal of Intelligent & Fuzzy Systems*, vol. 38, no. 4, pp. 5157–5173, 2020.

[39] D.-Y. Chang, "Applications of the extent analysis method on fuzzy AHP," *European Journal of Operational Research*, vol. 95, no. 3, pp. 649–655, 1996.

[40] S.-M. Chen, "Fuzzy system reliability analysis using fuzzy number arithmetic operations," *Fuzzy Sets and Systems*, vol. 64, no. 1, pp. 31–38, 1994.

[41] L. A. Zadeh, *Calculus of Fuzzy Restrictions. In Fuzzy Sets and Their Applications to Cognitive and Decision Processes*, Academic Press, Cambridge, MA, USA, 1975.

[42] L. A. Zadeh, "The concept of a *Z*-number-A new direction in uncertain computation," in *Proceedimgs of the 2011 IEEE International Conference on Information Reuse & Integration*, pp. xxii–xxiii, IEEE, Las Vegas, NV, USA, 2011 August.

[43] H. Aboutorab, M. Saberi, M. R. Asadabadi, O. Hussain, and E. Chang, "ZBWM: the *Z*-number extension of best eorst method and its application for supplier development," *Expert Systems with Applications*, vol. 107, pp. 115–125, 2018.

[44] Ž. Morkūnaitė, D. Kalibatas, and D. Kalibatienė, "A bibliometric data analysis of multi-criteria decision making methods in heritage buildings," *Journal of Civil Engineering and Management*, vol. 25, no. 2, pp. 76–99, 2019.

[45] M. KeshavarZ-Ghorabaee, M. Amiri, E. Zavadskas, Z. Turskis, and J. Antucheviciene, "An extended step-wise weight assessment ratio analysis with symmetric interval type-2 fuzzy sets for determining the subjective weights of criteria in multi-criteria decision-making problems," *Symmetry*, vol. 10, no. 4, p. 91, 2018.

[46] B. Juodagalvienė, Z. Turskis, J. Šaparauskas, and A. Endriukaitytė, "Integrated multi-criteria evaluation of house's plan shape based on the EDAS and SWARA methods," *Engineering Structures and Technologies*, vol. 9, no. 3, pp. 117–125, 2017.

[47] S. J. Ghoushchi, K. Gharibi, E. Osgooei, M. N. Ab Rahman, and M. Khazaeili, "Risk prioritization in failure mode and effects analysis with extended SWARA and MOORA Methods based on *Z*-numbers theory," *Informatica*, vol. 32, no. 1, pp. 1–27, 2020.

[48] S. Jafarzadeh Ghoushchi, M. N. Ab Rahman, D. Raeisi, E. Osgooei, and M. Jafarzadeh Ghoushji, "Integrated decision-making approach based on SWARA and GRA methods for the prioritization of failures in solar panel systems under *Z*-information," *Symmetry*, vol. 12, no. 2, p. 310, 2020.

[49] R. K. Mavi, M. Goh, and N. Zarbakhshnia, "Sustainable thir*D*-party reverse logistic provider selection with fuzzy SWARA and fuzzy MOORA in plastic industry," *The International Journal of Advanced Manufacturing Technology*, vol. 91, no. 5-8, pp. 2401–2418, 2017.

[50] Z. Turskis, E. K. Zavadskas, J. Antucheviciene, and N. Kosareva, "A hybrid model based on fuzzy AHP and fuzzy WASPAS for construction site selection," *International Journal of Computers Communications & Control*, vol. 10, no. 6, pp. 113–128, 2015.

[51] S. Chakraborty and E. K. Zavadskas, "Applications of WASPAS method in manufacturing decision making," *Informatica*, vol. 25, no. 1, pp. 1–20, 2014.

[52] E. K. Zavadskas, Z. Turskis, J. Antucheviciene, and A. Zakarevicius, "Optimization of weighted aggregated sum product assessment," *Elektronika Ir Elektrotechnika*, vol. 122, no. 6, pp. 3–6, 2012.

[53] M. K. Ghorabaee, E. K. Zavadskas, M. Amiri, and A. Esmaeili, "Multi-criteria evaluation of green suppliers using an extended WASPAS method with interval type-2 fuzzy sets," *Journal of Cleaner Production*, vol. 137, pp. 213–229, 2016.

[54] A. R. Mishra and P. Rani, "Interval-valued intuitionistic fuzzy WASPAS method: application in reservoir flood control management policy," *Group Decision and Negotiation*, vol. 27, no. 6, pp. 1047–1078, 2018.

[55] R. Lane, N. A. Stanton, and D. Harrison, "Applying hierarchical task analysis to medication administration errors," *Applied Ergonomics*, vol. 37, no. 5, pp. 669–679, 2006.

[56] J. Wu, J. Tian, and T. Zhao, "Failure mode prioritization by improved RPN calculation method," in *Proceedings of the 2014 Reliability and Maintainability Symposium*, pp. 1–6, IEEE, Colorado Springs, CO, USA, 2014 January.

[57] K. M. Davies, I. D. Coombes, S. Keogh, K. Hay, C. Hurst, and K. M. Whitfield, "Medication administration evaluation and feedback tool: simulation reliability testing," *Clinical Simulation in Nursing*, vol. 32, pp. 1–7, 2019.

Hindawi
Journal of Healthcare Engineering
Volume 2021, Article ID 5531327, 10 pages
https://doi.org/10.1155/2021/5531327

Research Article

Research on Medical Knowledge Graph for Stroke

Binjie Cheng,[1] **Jin Zhang**,[1] **Hong Liu**,[1] **Meiling Cai**,[1] **and Ying Wang**[2]

[1]*College of Information Science and Engineering, Hunan Normal University, Changsha 410081, China*
[2]*School of Humanities and Management, Hunan University of Chinese Medicine, Changsha 410208, China*

Correspondence should be addressed to Jin Zhang; jinzhang@hunnu.edu.cn

Received 24 February 2021; Revised 5 March 2021; Accepted 16 March 2021; Published 25 March 2021

Academic Editor: Hao-Chun Lu

Knowledge graph can effectively analyze and construct the essential characteristics of data. At present, scholars have proposed many knowledge graph models from different perspectives, especially in the medical field, but there are still relatively few studies on stroke diseases using medical knowledge graphs. Therefore, this paper will build a medical knowledge graph model for stroke. Firstly, a stroke disease dictionary and an ontology database are built through the international standard medical term sets and semiautomatic extraction-based crowdsourcing website data. Secondly, the external data are linked to the nodes of the existing knowledge graph via the entity similarity measures and the knowledge representation is performed by the knowledge graph embedded model. Thirdly, the structure of the established knowledge graph is modified continuously through iterative updating. Finally, in the experimental part, the proposed stroke medical knowledge graph is applied to the real stroke data and the performance of the proposed knowledge graph approach on the series of Trans $*$ models is compared.

1. Introduction

With the acceleration of urbanization and social aging, stroke has become one of the diseases with a high rate of death and disability in our country. Stroke is an acute cerebrovascular disease that causes brain tissue loss due to blockage or sudden rupture of blood vessels in the brain. It has the characteristics of high morbidity, mortality, disability, and recurrence. The human brain has an extremely complex nervous system and blood vessels, which brings great challenges to the treatment of brain diseases. So far, the current treatment methods for stroke are still limited and mainly prevent stroke to reduce the occurrence of the apoplexy. Considering that stroke involves many kinds of disease complications, the relationship between medical entities about stroke is complicated. It is impossible to directly use traditional technological means to effectively analyze the medical entity relationship of stroke, which brings a lot of inconvenience to the further prevention and treatment of stroke.

Thanks to the continuous development of Internet technology, the storage and sharing of knowledge has become more convenient. The knowledge of medical field can be integrated and developed through the Internet. More and more science and technology are used in the medical field to contribute to the smart medical field. Knowledge graph is a structured method to express knowledge through visual graphs, which can effectively show some directed lines with mark to depict the association between knowledge. Knowledge graph is essentially a semantic network that reveals the relationship between entities. The nodes in the network represent entities, and the edges between nodes represent the relationship between entities [1]. Knowledge graphs currently have more research and applications in related fields such as in-depth reading, finance, data analysis, medical fields, and other relevant fields. Especially in the medical field, it has very broad application prospects. We can apply the knowledge graph to intelligent question answering, disease-assisted diagnosis, risk assessment, and decision-making systems in the medical field and perform information screening and representation on the information of medical entities thereby establishing a database of medical knowledge relations. Through intuitive relationship expression, the problems in medical diseases can be analyzed more visually, and technical support can be provided for better treatment of diseases. The medical knowledge graph

contains a large number of disease knowledge and symptom characteristics, with a wider coverage of entities and concepts and more diverse semantic relationships, which can be used as the basis of machine language cognition. At present, the general medical knowledge graph has more in-depth research and practical application, but there is little attention paid to the knowledge graph for stroke.

According to the China Stroke Prevention and Treatment Report 2019 [2], the current form of stroke disease has become more and more serious, and it is urgent for scientific researchers to brainstorm and contribute to the study of stroke. Therefore, it is very necessary to design a reasonable medical knowledge graph model for stroke, to dig the physical associations related to stroke, and provide a powerful strategy for effective prevention and treatment of stroke further. This paper uses the public information database of the vertical medical website to extract the entities, select the ones with higher confidence to join the knowledge base, and build the connection between the entities. Then, a preliminary medical knowledge graph of stroke is designed through the bottom-up and top-down construction methods. Finally, the structure of the established knowledge graph is modified continuously through iterative updating. A stroke medical knowledge graph with comprehensive coverage, complete structure, and accurate knowledge description is constructed, which lays the foundation for the follow-up stroke intelligent question answering system and auxiliary decision-making system.

The rest of this paper is structured as follows. Some related works are reviewed in Section 2. In Section 3, we design a medical knowledge graph for stroke via knowledge fusion and knowledge graph embedding. In Section 4, we use a real stroke data to verify the performance of the proposed model. Finally, we make a conclusion and give the prospect of future research in Section 5.

2. Related Work

From the Semantic Web in the 1990s to the official introduction of the Knowledge Graph in 2012 by Google, the current public knowledge base and a large number of general knowledge graphs have emerged, proving that the research and application of knowledge graphs in various industries and fields have attracted great attention. Knowledge graphs can be simply divided into general knowledge graphs and domain knowledge graphs. There are many general knowledge graphs, large public knowledge bases in the research field, such as DBpedia [3] based on the structured knowledge extracted from Wikipedia entries, connection database YAGO [4] that integrates some Chinese information, and contributions from community partners large-scale cooperative knowledge base called Freebase [5], and Chinese open knowledge graph library OpenKG. Knowledge graph [6] is used by Google in the engineering field on the Google search engine, knowledge graph is used by Baidu for Baidu Search "*Zhixin*," and *Sogou's* "Knowledge Cube" is used in *Sougo Search*. The medical knowledge graphs studied in this article belong to a category of domain knowledge graphs. Many such knowledge graphs have also been constructed, such as Watson Health of IBM, which mainly uses knowledge graph reasoning and AI analysis of tumor medical images to assist decision-making [3]. And the Chinese medical knowledge graph is also promoted by many domestic institutions such as the knowledge graph constructed by Shanghai Shuguang Hospital mainly used for TCM knowledge QA system and medication recommendations [7] and the Chinese symptom database constructed by East China University of Science and Technology, which has been published on OpenKG [8]. In terms of the construction method of the knowledge graph, a knowledge-based syndrome reasoning method in computer-assisted diagnosis uses the reinforcement learning algorithm to mine the hidden relationship among the entities and obtain the reasoning path. [9] And an improved method for web text affective cognition computing based on knowledge graph constitutes a binary relationship knowledge base. [10] Although the current domestic research on knowledge graphs has gradually improved, various tasks still require a lot of manual intervention to improve, and the research on knowledge graphs in the medical field in our country is still in its infancy.

As mentioned in the China Stroke Prevention and Treatment Report 2019, stroke has become the first year of life lost in China. In 2018, more than 20% of Chinese residents died of cerebrovascular diseases, and the mortality rate has been increasing year by year. The current research on stroke is mainly focused on prevention and treatment [11], so this research uses stroke and related knowledge as the entry point to study the knowledge graph of stroke and construct a medical knowledge graph for stroke.

3. Construction of the Medical Knowledge Graph of Stroke

3.1. Stroke Medical Knowledge Modeling. The construction of the medical knowledge graph can be divided into two stages. The first is the artificial participation in the construction of the stroke disease dictionary. This part belongs to the description system design. We have investigated the authoritative standard medical term sets at home and abroad, such as the ICD-10 code [12]. According to the medical terminology, we labeled and analyzed the cases and designed the relationship classification system. After comparative evaluation, a stroke disease dictionary was initially constructed. The second stage is the construction of the knowledge graph. Under the guidance of the stroke disease dictionary, the stroke disease-related symptoms, treatment methods, drugs and other attributes, and relationship data are obtained from vertical medical websites and public information databases such as Baidu Encyclopedia. The extracted structured data are manually annotated and automatically extracted to construct the stroke medical knowledge graph ontology. After the knowledge graph modeling and knowledge processing, the semiautomatic construction of the stroke medical knowledge graph is realized.

After determining the construction method, the overall construction framework is formed as shown in Figure 1. The semiautomated construction of the design pattern adopts the

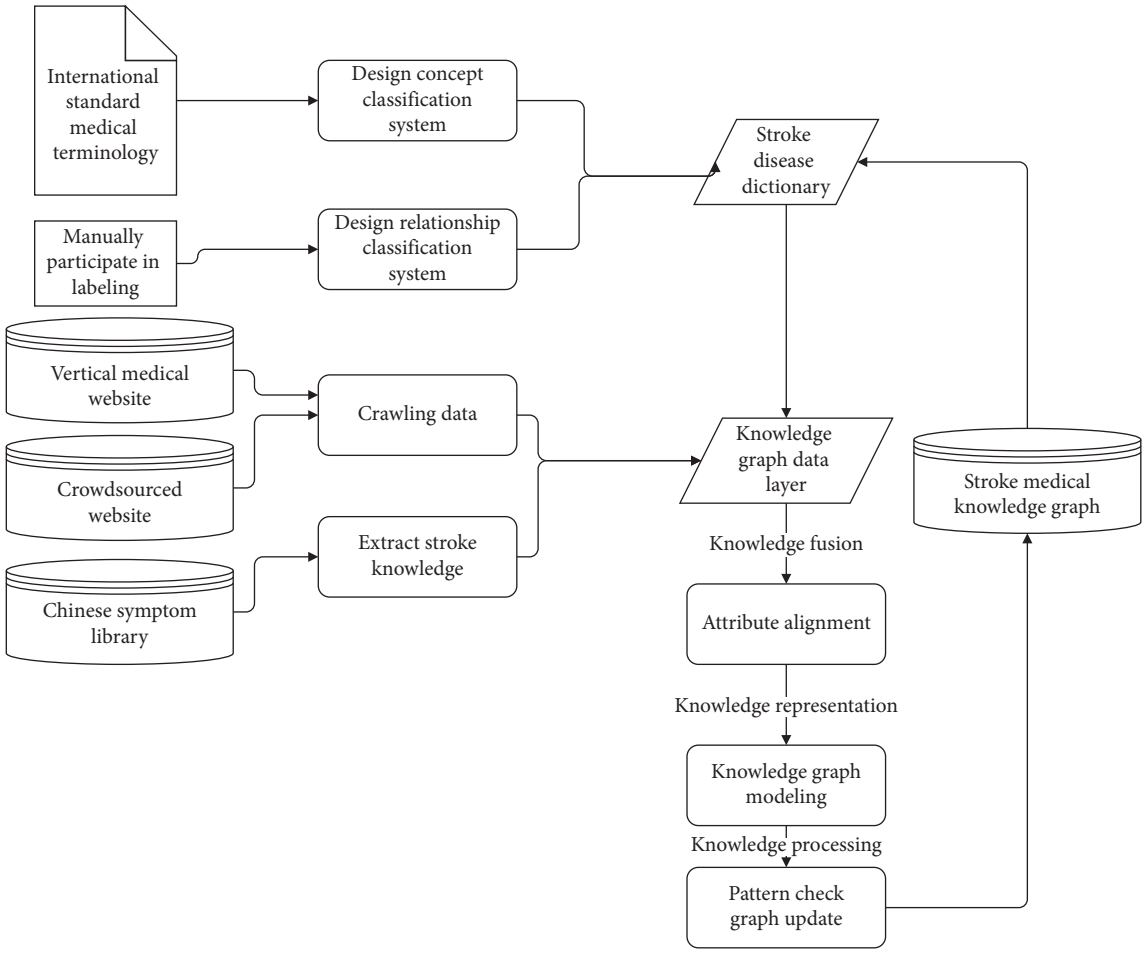


Figure 1: Stroke medical knowledge graph construction framework diagram.

cyclic iteration style. Each iteration includes the steps of graph pattern design, knowledge extraction, knowledge graph modeling, and knowledge processing. Such a design can realize the sustainable update of the knowledge graph, so each iteration will produce a new version of the knowledge graph.

3.2. Knowledge Graph Pattern Design. To construct a stroke-oriented medical knowledge graph, it is necessary to consider its actual application needs. The stroke medical knowledge graph is mainly used to provide patients with stroke disease self-examination and also provide doctors' medical knowledge base as a reference. The data in the graph contain data on stroke diseases, as well as related treatment methods, symptoms, inspection methods, and drugs in accordance with the requirements of the designated ontology library for manual participation.

Based on medical thesaurus, ICD-10 coding, and other medical terms as entities in the entity resource database, we manually participate in marking stroke diseases and classify them according to different types, respectively, from etiology, diagnostic methods, differentiation, epidemiology, and complications. Past medical history, prevention and recurrence, and other dimensions describe stroke. According to the above methods, a stroke disease dictionary was designed, combined with vertical medical websites, Baidu Encyclopedia, and medical literature to develop a stroke ontology database. In the stroke knowledge graph, the rules in the ontology database are used to constrain the model layer data. The stroke ontology describes all the concepts in the field of stroke and the relationships between them, such as proximity relationships, attribute relationships, and constraints.

There are three methods for constructing the ontology database of the knowledge graph [13]. The first is artificial construction, which is to invite medical experts to participate in the annotation process of the knowledge graph ontology and play a leading role in the description system of the knowledge graph. The second method is automatic construction, which relies on data-driven; the third is semiautomatic construction, which combines manual annotation and data-driven. Taking into account the small scale of the ontology library of stroke medicine implementation graphs, the third method can obtain the ontology with higher quality and save human resources. This article uses the third method to construct the ontology library. First, use manual intervention to build a simple stroke ontology database for the stroke disease dictionary, then extract pattern data with guaranteed accuracy from the

knowledge processing process in the subsequent steps, and add it to the stroke ontology database after manual verification is correct.

According to this construction process, after extracting high-confidence data from the vertical medical website and Baidu Encyclopedia, the entities, relationships, and attributes covered by the ontology library are refined, and the value range of each attribute is finally clarified. Figure 2 shows the information of the disease entity category.

3.3. Data Source. In the process of constructing a stroke medical knowledge graph, a key step is knowledge extraction, and data confidence is the basis for whether knowledge extraction is efficient and accurate. Figure 3 shows the knowledge extraction process of the stroke medical knowledge graph.

Currently, there is a lack of a stroke medical knowledge graph in the public knowledge base, so the data sources for constructing a stroke knowledge graph in this article mainly include vertical medical websites, crowdsourced encyclopedia websites, and general knowledge graphs in the public knowledge base.

Two medical websites were selected in this study: Xunyiwenyao.com, which is a doctor-patient communication platform. Doctors and hospitals participate voluntarily. The content contained in the website is highly professional and accurate. The degree is high, mainly including basic medical information data, health information, and medical questions and answers; *Dingxiangyuan* is a data service platform that includes disease encyclopedias, medication specifications, and medical science popularization. This study combined the open source data of these two vertical medical websites as the main source of the data layer of the stroke knowledge graph.

Baidu Encyclopedia is the largest crowdsourcing website in China. It is basically crowdsourced by experts in various fields. The structured data contained in it are the best supplement to the knowledge graph data layer. Baidu Encyclopedia is aimed at medical entity entries. The explanatory text contains a lot of information such as disease nicknames and context, and this part of the data has high value.

The public knowledge base, the Chinese symptom database of East China University of Science and Technology from *OpenKG* used in this article, extracts stroke nodes and relationships from it. This part can also complement the ontology database and data layer of the knowledge graph.

Different data sources require different processing methods. Since the data on vertical medical websites and crowdsourcing websites are structured, this article uses distributed crawlers to automatically crawl medical data. The knowledge graph from the public knowledge base can be directly integrated. After obtaining the data from the three sources, it is saved as a lightweight JSON file. For this kind of data, the first task is to clean up the data and clean up the default, garbled, and illegal characters.

3.4. Knowledge Fusion. The knowledge fusion in the knowledge graph can be realized in many ways. This study uses the two perspectives of entity attribute alignment and entity linking with the help of the similarity calculation method to carry out the knowledge fusion of the stroke knowledge graph. Entity alignment refers to merging the entities in the knowledge base of heterogeneous data sources into an entity with a globally unique identifier in the real world [14] and then linking the aligned entities to the stroke knowledge graph. Since this paper proposes to formulate a stroke disease dictionary, the entity alignment step has been completed in the data extraction stage, so the entity attribute alignment stage is mainly to align the attributes.

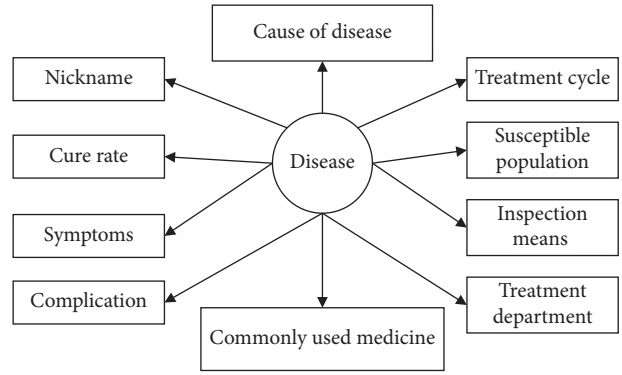


Figure 2: Partial structure of disease ontology.

3.4.1. Attribute Alignment. The effect of this step of attribute alignment is to improve the accuracy of entity links. Since there are few attributes in the stroke field, this paper constructs an attribute mapping table based on the constraints of the stroke ontology library and aligns the different expressions of the same attributes of the same entities. Table 1 shows some attribute alignment mappings of heterogeneous data sources.

After the attributes of heterogeneous data sources are aligned, the attribute values are standardized according to the constraint specifications in the model layer. This research divides the attributes as follows: numeric type, numeric interval type, entity object list type, string type, and Boolean type. These attribute values are standardized and structured according to the following constraint specifications:

(i) The unit of measurement for numeric attribute values is unified

(ii) Space characters and line breaks in the attribute value of string type must be deleted

(iii) For interval attribute values, keep the upper and lower limits and store them in the list

(iv) The attribute values of the entity object type are all stored in the list, and no attribute alignment operation is performed

3.4.2. Entity Linking. After completing the work of attribute alignment and normalizing attribute values, choose to calculate the semantic similarity of the result after attribute alignment to determine the relationship with the entity nodes in the knowledge graph and then decide whether to

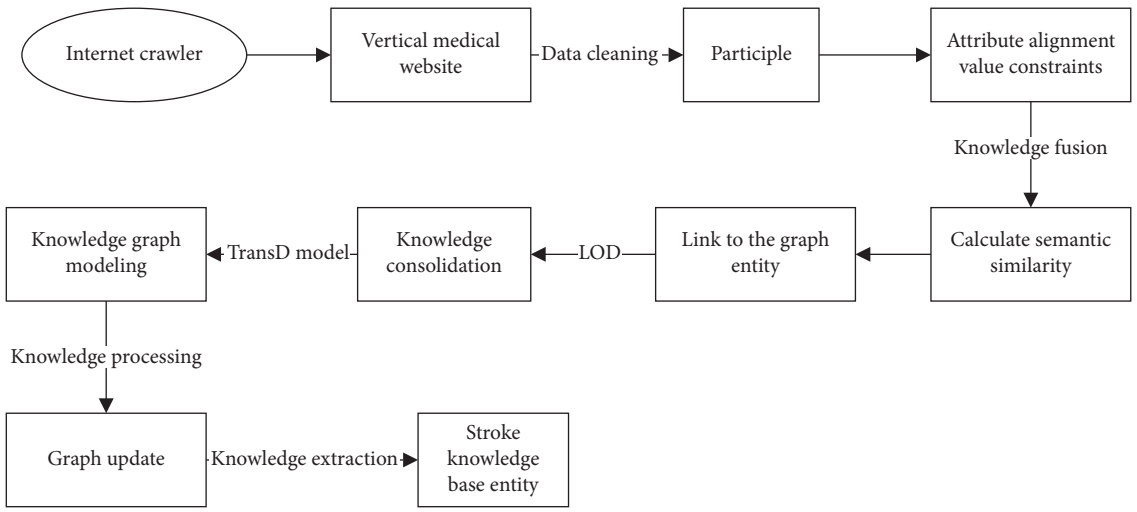


Figure 3: Knowledge extraction flowchart of stroke medical knowledge graph.

Table 1: Partial attribute alignment mapping table of heterogeneous data sources.

Sample	Data source	Original attribute	Aligned attributes
1	https://http://www.xywy.com/	Treatment department	Treatment department
	https://www.dxy.cn/	Registration department	
2	https://http://www.xywy.com/	Alias	Nickname
	https://www.dxy.cn/	Nickname	

link to the knowledge graph. In stroke medical entities, this study calculates the similarity of the abovementioned different types of attribute values to determine whether to link to the entities in the knowledge graph. For two entities *W1* and *W2*, the corresponding aliases and names merged into the name set are *S1* and *S2*, respectively. Calculate the similarity according to the following equation and links with high similarity to entities in the knowledge graph:

$$\mathrm{Sim}\left(W_1, W_2\right) = \frac{k}{1+N}\max\left(\sum_{i\in W1}^{N}\sum_{j\in W2}^{N}\frac{2\times lcs\left(i,j\right)}{L_i + L_j}\right). \tag{1}$$

Among them, $lcs\,(i, j)$ represents the length of the largest common subsequence of the names i and j, N is an adjustable parameter, L_i and L_j represent the length of the word in W_1 and W_2 in the entity, and k is a weight parameter to avoid the influence caused by the high similarity of the name set.

3.4.3. Knowledge Merger. Knowledge merging in the medical knowledge graph is to integrate structured knowledge and knowledge in the public knowledge base into the existing knowledge graph. The aforementioned Chinese symptom database is such a source of knowledge that can be incorporated into the stroke knowledge graph. The knowledge it possesses conforms to the knowledge norms and has high practicality and high knowledge quality.

This article will refer to the method of knowledge merging for LOD by Mendes et al. [15] and conclude that the process of merging the public knowledge base into the stroke knowledge graph is knowledge extraction, concept matching, entity alignment, and knowledge evaluation. The two steps of concept matching and entity alignment are to normalize the knowledge extracted from the Chinese symptom database with the artificial stroke disease dictionary, and the knowledge evaluation is used to detect the consistency and accuracy of the extracted knowledge.

3.5. Knowledge Graph Embedding. In the medical knowledge graph, each piece of medical knowledge can be represented by a triple such as ⟨head, relation, tail⟩, where head represents the head entity node, tail represents the tail entity node, and relation represents the relationship between nodes. For example, the results of the triple representation of ischemic stroke are shown in Table 2. At the same time, all the triples in Table 2 are combined to obtain a partial knowledge graph of "ischemic stroke," as shown in Figure 4. The knowledge graph modeling is to graph the relationship between the head and tail entity nodes and the entities to a continuous low-dimensional vector space and represent them as entity vectors, so that the semantic structure information between the triples can be saved and used to calculate the similarity between entities. This article uses TransE [16] and TransD [17] models that are currently widely used in knowledge representation. The TransE model is the first model proposed by the Trans series, which mainly associates entity

Table 2: Ischemic stroke description triad.

Head	Relation	Tail
Ischemic stroke	Nickname	Cerebral infarction
Ischemic stroke	Susceptible population	People with high blood pressure
Ischemic stroke	Susceptible population	Middle-aged and elderly
Ischemic stroke	Clinical manifestations	Hemiplegia
Ischemic stroke	Clinical manifestations	Hard to swallow
Ischemic stroke	Clinical manifestations	Numbness of the contralateral limb
Ischemic stroke	Clinical manifestations	Sensory disturbance
Ischemic stroke	Clinical manifestations	Aphasia

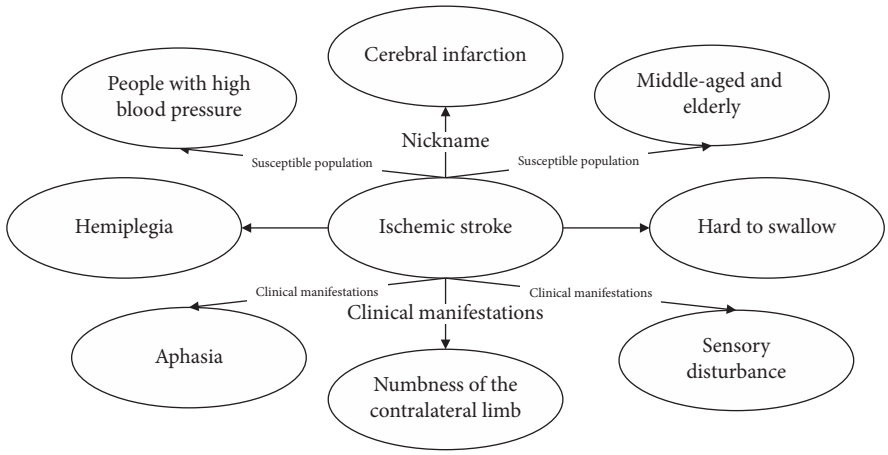


Figure 4: Example of ischemic stroke knowledge graph.

nodes by mapping the word vector to a low-dimensional space to calculate the similarity. The TransD model is a more advanced generalization of the TransE model. The latter is a special case of the former. Therefore, this paper will verify which model is better for the proposed knowledge graph construction.

The TransE model is the most representative translation model. For the given triples, the relation is interpreted as the translation vector from head to tail; when the triple of ⟨head, relation, tail⟩ is established, the relationship between nodes is shown like head + tailrelation. Otherwise, the tail node should not be linked to the other end of the head entity node and the relationship. The specific function formula is shown in the following equation:

$$f_r(h,t) = \|h + r - t\|_{L_1/L_2}. \tag{2}$$

Among them, L_1/L_2 refers to the distance of L_1 or L_2 when calculating. L_1 represents the translation distance from the r vector to the h vector. L_2 represents the translation distance from the r vector to the t vector. Due to the few model parameters and low computational complexity, the TransE model can handle one-to-one relationships well, but its performance is insufficient when dealing with one-to-many or even many-to-many relationships. Therefore, this article will also combine the TransD model and the TransR model that has been extended to TransE to perform knowledge representation processing on the triples of the subdivided medical knowledge graph.

The TransD model believes that the head and tail entity nodes in the triplet represent different semantics after the connection relationship. Therefore, the model uses the mapping matrices M_h and M_t to graph the head and tail entities to the relationship space and decomposes the mapping matrix of each relationship into the product of two vectors, and the function formula is shown in the following equation:

$$f_r(h,t) = \left\|M_h h = r - M_t t\right\|_{L_1/L_2}, \tag{3}$$

where $M_h = r_p h_p + I$, $M_t = r_p t_p + I$, and I is the identity matrix. Obviously, we can see that M_h and M_t are related to entities and relationships. The calculation speed of the model can be improved through vector operation conversion.

3.6. Knowledge Processing

3.6.1. Medical Ontology Model Layer Inspection. In the case that the new version of the ontology model layer formed by iteration and the existing model layer exists at the same time, the type and value constraints of the data that have been fused into the graph are applied to make the data layer meet the ontology model specifications defined in the layer. This article generates a rule database based on the model layer constructed by the stroke disease dictionary to test the data layer. The rule base includes entity type detection and attribute value interval detection.

3.6.2. Graph Update. In Section 3.2, this article developed a stroke disease dictionary based on actual needs and defined the model layer of the stroke knowledge graph based on the dictionary. However, it is difficult to ensure that the model layer can cover all data patterns in practical applications. Among the data obtained from heterogeneous data sources, there are some data patterns that are not clearly defined but have research value. We will generalize new data patterns from these new data and improve and supplement the stroke ontology database previously constructed.

Since the stroke medical knowledge graph constructed in this paper is different from the general medical knowledge graph, we have manually intervened the boundaries of the knowledge graph to ensure that the knowledge graph will not extend infinitely in breadth. For the relationship types that already exist in the current model layer and the head and tail node types are also known, they are linked to the existing entity relationships in the stroke knowledge graph through the semantic similarity calculation in Section 3.4. If there is no corresponding entity relationship, add a new entity relationship directly in the pattern layer; another case is that only one of the connected head and tail entities can find the corresponding relationship type in the pattern layer. This unknown type will appear in this article, the entity is extracted, the number of occurrences is calculated, and then its entropy value is normalized. The relationship type with larger entropy value and the entity type is added to the model layer as a candidate model. The stroke knowledge graph constructed by this loop iteration method will be relatively stable and complete, and the subsequent update direction depends on the application requirements. The follow-up knowledge graph update work can use deep learning models to learn the cascaded R–CNN and a correlation filter learning model of real data [18,19].

3.6.3. Quality Evaluation. The quality evaluation of the knowledge graph mainly quantifies the confidence of the knowledge in the graph and discards the knowledge with low confidence to ensure the quality of the entire knowledge graph [20]. In this paper, a semiautomated hierarchical ontology metric based on semiotics combined with manual calculation is used to evaluate the constructed stroke medical knowledge graph.

4. Experiment

4.1. Dataset. The experimental data in this article consist of two parts. One part of the data comes from the stroke-related description text data crawled from Xunyiwenyao.com and Baidu Encyclopedia by writing a crawler program; the other part is public Chinese obtained from OpenKG a triad of partial knowledge about stroke which is extracted from the symptom database and obtained a total of 4113 related entities related to stroke, with 8 attribute types (name of disease, introduction, susceptible population, etiology, treatment, treatment cycle, preventive measures, and cure probability) and 10 types of entity relationship. Specific information is as follows. Table 3 shows that there are 7 types of entities, and the specific type information is shown in Table 4.

4.2. Experiment Analysis. In this paper, the construction of the stroke knowledge graph is firstly through the semi-automatic labeling method plus human participation and the development of a stroke disease dictionary, combined with international general medical terminology and other professional information to construct a preliminary model layer. Then, a crawler is designed, combined with data cleaning to screen and crawl the open text data of *Xunyiwenyao.com*, *Dingxiangyuan.com*, and *Baidu Encyclopedia*, and the knowledge is extracted from the Chinese symptom database constructed by East China University of Science and Technology as stroke medicine. To supplement the knowledge graph, store these knowledge triples in the Neo4 J graph database. By means of knowledge fusion, knowledge graph embedding, and knowledge processing, iteratively update the stroke knowledge graph so that it is continuously improved while retaining the boundaries, forming the domain closure of the stroke knowledge graph.

Figure 5 is a partial knowledge graph of stroke disease information. The constructed stroke medical knowledge graph belongs to the Chinese knowledge graph, so the content shown in Figure 5 is Chinese. For example, hypertensive patients belong to the high-risk population of stroke diseases. In the database, you can query the information on the antihypertensive drugs of the predisposing factors of stroke and high blood pressure in the database. The database will return as follows. Figure 6 shows the subgraph. Since the *Xunyiwenyao* website and the *Dingxiangyuan* website belong to a doctor-patient communication platform, there will be consultations from most patients to medical experts. This part of the data is also crawled down in this study, and an intelligent question-and-answer system can be initially constructed.

4.3. Comparative Analysis of Trans Series Models. This paper selects the TransE model and the TransD model to express the knowledge of the knowledge graph. The experiment will select the TransD model as the main research model to compare with the TransE model and use the two models, respectively, to randomly extract the 200-dimensional head and tail entities and relationships from the knowledge graph. The triple vector is used for training. The experimental verification method is cross-validation, using 70% of the randomly selected vectors as the training set and 30% of the data as the test set. The three values of precision, recall, and F1 are used as evaluation indexes; these three indicators are mainly used to evaluate the effects of these two models in this article, so as to choose which model to use. The experiment was repeated 10 times, and the average value was taken as the final model evaluation result.

The experimental results are shown in Table 5. It is not difficult to see that the performance of the TransD model is better than that of the TransE model. In terms of accuracy, recall, and F1 value, the TransE model has different degrees

Table 3: Knowledge graph entity relationship type.

Entity relationship type	Example	Number of relationships
Department	*<Stroke, belongs to, Neurology>*	551
Commonly used drugs	*<Ischemia stroke, commonly used drugs, mannitol>*	2035
Suitable to eat	*<Ischemic stroke, suitable to eat, rich in plant protein>*	386
Drugs on sale	*<Ischemic stroke, medicines on sale, Tongmai granules>*	527
Examination	*<Ischemia stroke, examination, brain MRI examination>*	1029
Avoid eating	*<Ischemia stroke, avoid eating, high greasy fat>*	793
Recommended drugs	*<Ischemic stroke, recommended drugs, Tongmai granules>*	1348
Recommended recipe	*<Ischemic stroke, recommended diet, easy to digest>*	410
Symptom	*<Ischemia stroke, symptoms, deflection of tongue extension>*	949
Complication	*<Ischemic stroke, complications, pulmonary embolism>*	1202

Table 4: Stroke knowledge graph entity type.

Entity type	Example	Number of entities
Diagnostic inspection items	Head CT or brain MRI	301
Treatment department	Internal medicine, neurology	10
Disease	Ischemic stroke, hemorrhagic stroke	798
Medicine	Urokinase	2035
Food	Light and easy to digest food	507
Disease symptoms	Deviated tongue, sticking out	462

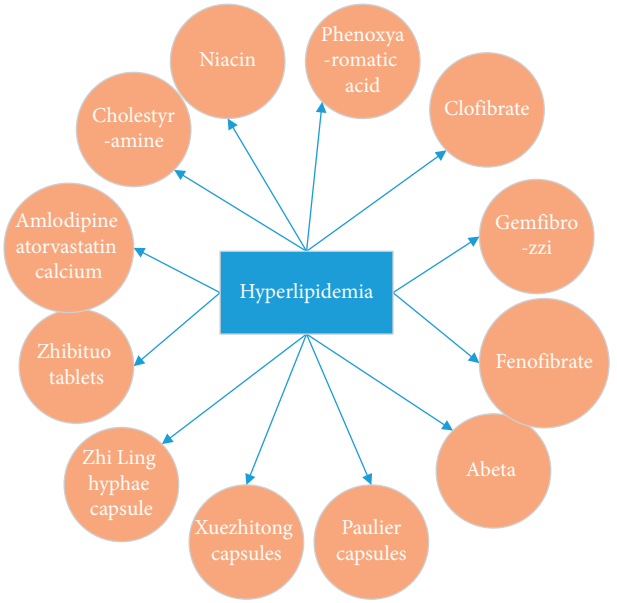


Figure 5: Part of ischemic stroke knowledge graph.

of improvement over the TransD model. Therefore, in the medical knowledge graph with complex semantic relationships, the TransD model graphs the head and tail entities to the low-dimensional vector space of the relationship through the mapping matrix to represent the semantic structure and better captures the nonlinearity between structured knowledge. Relationships reduce the loss of the vectorization process of physical nodes. This method is more reasonable and can play a greater role in the stroke ontology database.

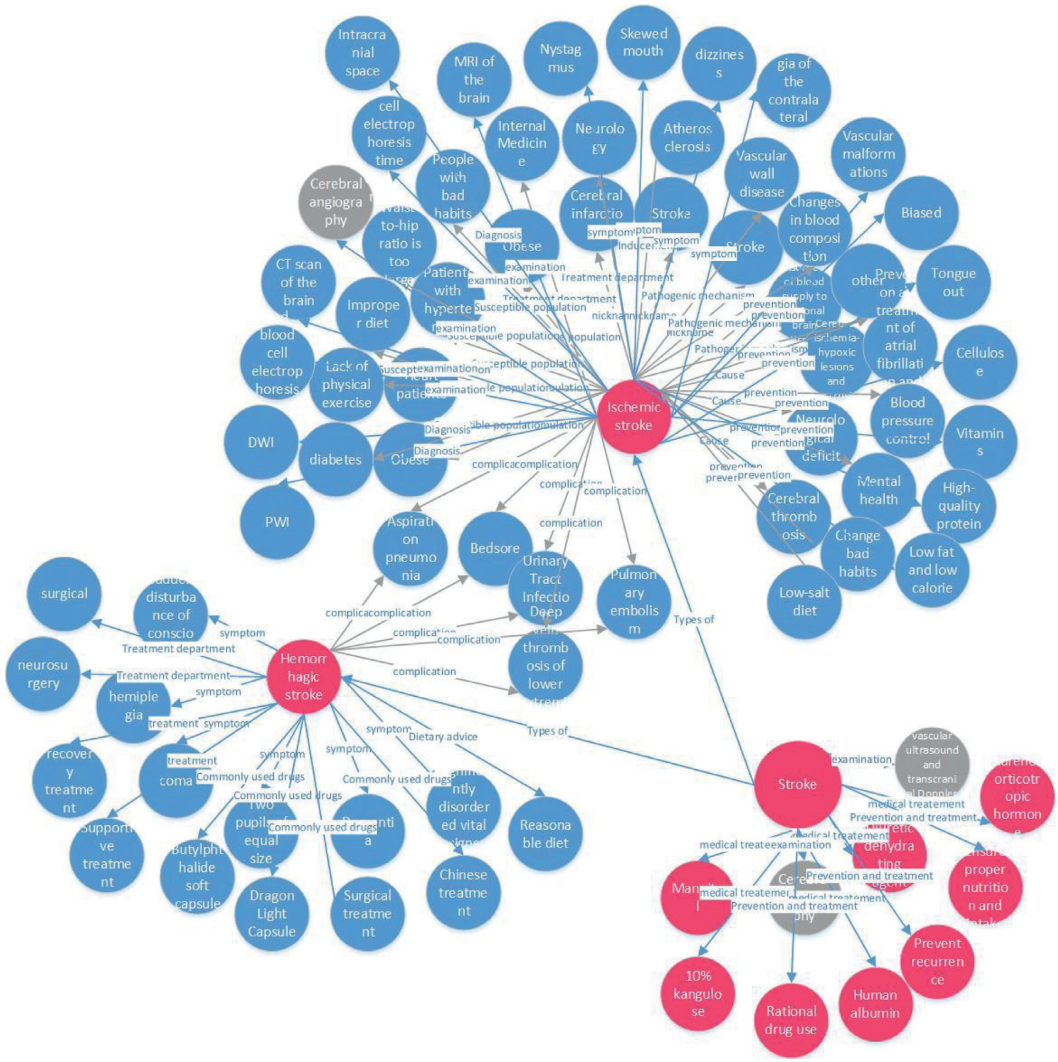


Figure 6: Subfigure of high blood lipid entity structure.

Table 5: Trans series model parameter table %.

Model	Precision	Recall	F1
TransE	71.7	62.9	67.0
TransD	85.8	85.45	85.6

5. Conclusion

Stroke is a disease that urgently needs to reduce the risk of treatment. The proposed stroke-oriented medical knowledge graph can effectively discover the associations between medical entities and establish a certain foundation for subsequent intelligent question-and-answer and medical assistance decision-making systems. Firstly, according to the actual application requirements, we manually participate in the development of a stroke dictionary using semiautomatic annotation and build a model layer of the knowledge graph combined with international standard medical terminology such as ICD-10. Secondly, the improved entity similarity measure is used to perform knowledge fusion on the processed stroke information and link the relationships between entities to the nodes of knowledge graph. Then, we perform the knowledge representation by the knowledge graph embedded model, and the constructed knowledge representation is updated and iterated at the same time to more accurately express the association between entities. In addition, the constructed knowledge graph of stroke already can be used in general medical question answering systems.

At present, there are a few medical knowledge graphs that have been applied to the actual medical scenarios, and their confidence has always been controversial. It is also a problem that needs to be solved as to how to systematically construct the medical knowledge graph for stroke. In future research work, we will focus on the application-level development of the stroke knowledge graph, such as exploring how to further intelligentize the question answering system and combine the deep learning model to extend the intelligent diagnosis.

Data Availability

The data used to support the findings of this study are included within the article (https://www.xywy.com).

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the education and research projects of Hunan Provincial Education Department (JG2018A012, XiangJiaoTong, 2019: nos. 291–410, 248–27, and 370; 2020: nos. 9, 90, and 233 HNKCSZ-2020-0122), the projects of the Ministry of Education of the People's Republic of China (201901051021), the Science and Technology Progress and Innovation Project of Hunan Provincial Department of Transportation (201927), and economic burden and policy of traditional Chinese and western medicine for stroke patients in Hunan Province: 2019 scientific research project of Hunan University of Chinese Medicine (2019XJJJ030).

References

[1] H. Q. Huang, J. Yu, X. Liao et al., "Review on knowledge graphs," *Computer Systems & Applications*, vol. 28, no. 6, pp. 1–12, 2019.

[2] Z. Li, "Summary of China stroke prevention and treatment Report 2019," *Chinese Journal of Cerebrovascular Diseases*, vol. 17, no. 05, pp. 272–281, 2019, in Chinese.

[3] C. Bizer, J. Lehmann, G. Kobilarov et al., "DBpedia - a crystallization point for the Web of Data," *Journal of Web Semantics*, vol. 7, no. 3, pp. 154–165, 2009.

[4] F. M. Suchanek, G. Kasneci, and G. Weikum, "Yago: a core of semantic knowledge," in *Proceedings of the of the 16th Int'l Conf. on World Wide Web*, pp. 697–706, ACM Press, Lyon, France, April 2007.

[5] K. Bollacker, C. Evans, P. Paritosh et al., "Freebase: a collaboratively created graph database for structuring human knowledge," in *Proceedings of the of the 2008 ACM SIGMOD Int'l Conf. on Management of Data*, pp. 1247–1250, ACM Press, Vancouver, BC, Canada, June 2008.

[6] K. J. Vang, "Ethics of google's knowledge graph: some considerations," *Journal of Information Communication & Ethics in Society*, vol. 11, no. 4, pp. 245–260, 2014.

[7] L. JIA, J. LIU, Y. U. Tong et al., "Construction of traditional Chinese medicine knowledge graph," *Journal of Medical Informatics*, vol. 36, no. 8, pp. 51–53, 2015.

[8] F. Hu, *Research on the Construction Method of Chinese Knowledge Graph Based on Multiple Data sources*, East China University of Science and Technology, Shanghai, China, 2015, in Chinese.

[9] Y. Xie, L. Hu, X. Chen, J. Feng, and D. Zhang, "Auxiliary diagnosis based on the knowledge graph of tcm syndrome," *Computers, Materials & Continua*, vol. 65, no. 1, pp. 481–494, 2020.

[10] B. Niu and Y. Huang, "An improved method for web text affective cognition computing based on knowledge graph," *Computers, Materials & Continua*, vol. 59, no. 1, pp. 1–14, 2019.

[11] G. Tsivgoulis, A. H. Katsanos, P. Mandava et al., "Blood pressure excursions in acute ischemic stroke patients treated with intravenous thrombolysis," *Journal of Hypertension*, vol. 39, no. 2, pp. 266–272, 2021.

[12] H. Zan, H. Dou, Y. Jia et al., "Construction of Chinese medical knowledge graph based on multi-source texts," *Journal of Zhengzhou University (Science Edition)*, vol. 52, no. 02, pp. 45–51, 2020, in Chinese.

[13] Y. Liu and H. Li, "A survey of domain knowledge graph research," *Computer System Application*, vol. 29, no. 6, pp. 1–12, 2020, in Chinese.

[14] L. Qiao, L. Yang, H. Duan et al., "Overview of knowledge graph construction technology," *Computer Research and Development*, vol. 53, no. 3, pp. 582–600, 2016, in Chinese.

[15] P. N. Mendes, H. Hleisen, and C. Bizer, "Sieve: linked data quality assessment and fusion," in *Proceedings of the 2012 Joint EDBT/ICDT Workshops*, pp. 116–123, Berlin, Germany, May 2012.

[16] A. Bordes et al., "Translating embeddings for modeling multi-relational data," in *Proceedings of the International Conference on Neural Information Processing Systems*, Lake Tahoe, NV, USA, December 2013.

[17] G. Ji, S. He, L. Xu, K. Liu, and J. Zhao, "Knowledge graph embedding via dynamic mapping matrix," in *Proceedings of the Meeting of the Association for Computational Linguistics & the International Joint Conference on Natural Language Processing*, Beijing, China, July 2015.

[18] J. Zhang, Z. Xie, J. Sun, X. Zou, and J. Wang, "A cascaded R-CNN with multiscale Attention and imbalanced samples for traffic sign detection," *Institute of Electrical and Electronics Engineers Access*, vol. 8, no. 8, pp. 29742–29754, 2020.

[19] J. Zhang, X. Jin, J. Sun, J. Wang, and K. Li, "Dual model learning combined with multiple feature selection for accurate visual tracking," *Institute of Electrical and Electronics Engineers Access*, vol. 7, no. 7, pp. 43956–43969, 2019.

[20] M. McDaniel, V. C. Storey, and V. Sugumaran, "Assessing the quality of domain ontologies: metrics and an automated ranking system," *Data & Knowledge Engineering*, vol. 115, pp. 32–47, 2018.

Hindawi
Journal of Healthcare Engineering
Volume 2021, Article ID 6692952, 8 pages
https://doi.org/10.1155/2021/6692952

Research Article

Role of an e-Health Intervention in Holistic Healthcare: A Quasiexperiment in Patients Undergoing Cardiac Catheterization in Taiwan

Jian-Rong Peng,[1] **Hung-Chi Su**,[2] **Chia-Pin Lin**,[2] **Chun-Chi Chen**,[2] **Chi-Jen Chang**,[2] **Siou-Ling Gong**,[2] **and Pao-Hsien Chu**[2,3]

[1]*Department of Cardiology, New Taipei Municipal TuCheng Hospital, Chang Gung Memorial Hospital, New Taipei, Taiwan*
[2]*Department of Cardiology, Chang Gung Memorial Hospital, Linkou Medical Center, Taoyuan, Taiwan*
[3]*College of Medicine, Chang Gung University, Taoyuan, Taiwan*

Correspondence should be addressed to Pao-Hsien Chu; taipei.chu@gmail.com

Received 7 December 2020; Revised 18 February 2021; Accepted 2 March 2021; Published 17 March 2021

Academic Editor: Hsiao-Ting Tseng

Background. The use of electronic health (e-health) resources is emerging as an alternative method to improve the secondary prevention of coronary artery disease (CAD). The aim of this study was to describe the influence of an e-health application in holistic healthcare for patients with CAD. *Methods*. A quasiexperiment with nonequivalent groups design recruited outpatients with a high risk of CAD admitted for cardiac catheterization. They were divided into two groups. Before the procedure, the control group received traditional patient education, and the intervention group watched videos on Internet-based social media. EQ-5D and FACIT–Sp-12 questionnaires were used as outcome measures of interest, and they were administered before and after the procedure and at the first return visit to the outpatient clinic after discharge. The effect of each intervention was tested using a linear mixed effects model. In addition, the 90-day readmission rate was also studied. *Results*. A total of 300 patients were divided into intervention and control groups (150 patients in each group). The interaction effect of EQ-5D was not statistically significant; however, improvements in FACIT–Sp-12 were greater in the intervention group from baseline to before discharge (regression coefficient (B) = 1.70, $p < 0.001$) and from baseline to postdischarge first outpatient visit ($B = 1.81$, $p < 0.001$). Moreover, the 90-day readmission rate was significantly lower in the intervention group (14% vs. 18.7%; $p = 0.016$, log-rank test). *Conclusions*. e-health intervention with easily accessible Internet-based social media is a promising model to meet the holistic needs of patients with CAD in the modern era.

1. Introduction

Electronic health (e-health) refers to health services and information delivered or enhanced through the Internet and related technologies [1]. Mobile health (m-health), a subset of e-health, means the use of mobile computing and communication technologies in healthcare and public health [2]. As most mobile devices can access the Internet, e-health is widely applied to the general public, including electronic medical records, personal health records, electronic communication between patients and providers, and education programs [3]. The European Society of Cardiology also recommends the use of e-health resources to support remote clinical care and improve psychosocial health in patients with cardiovascular diseases (CVDs) [4–6].

CVDs are the leading causes of deaths worldwide, accounting for 30% of all deaths. Of these deaths, 50% is due to heart attack, and ischemic heart disease causes the most deaths worldwide [7, 8]. For patients with suspected coronary artery disease (CAD), accurate diagnostic assessment with invasive coronary angiography and cardiac catheterization (CC) is necessary to assess the prognosis and guide

the choice of revascularization procedure such as percutaneous coronary intervention (PCI) or coronary artery bypass graft surgery (CABG) [9, 10].

With advances in material science and medical technology, CC has become a very common diagnostic procedure for patients with suspected CAD. As an invasive procedure, CC can cause significant patient anxiety [11], which may have negative implications on healthcare perception, expectations, decision-making, quality of life [12–15], and clinical outcomes [16]. To reduce patient anxiety and improve patient satisfaction, holistic healthcare can be used that considers not only physical health but also spiritual well-being [17].

In Taiwan, the average length of a scheduled hospital stay for CC and/or PCI is three days from admission to discharge. Within such a short period, it is not easy to provide adequate information to the patients. Therefore, a more effective informational education model is needed to provide holistic healthcare. The purpose of this study was to describe the influence of an e-health application in patients with suspected or established CAD undergoing CC.

2. Patients and Methods

2.1. Study Population. This study was a quasiexperimental research with a two-nonequivalent group prospective comparison design [18]. Outpatients from our cardiovascular outpatient department (CV OPD) with suspected or established CAD who were scheduled to undergo elective CC were assigned to two groups: the control group and intervention group (Figure 1). The patients in the control group received traditional ward-based patient education, including informed consent for CC by physicians and standard education and handouts on CC by clinical nurses before the procedure. In addition to the traditional patient education, those in the intervention group watched an instructional video produced by cardiologists and nurses on Internet-based social media. The length of the video was 11 minutes and included four major parts: (1) introduction to CAD, (2) indications for CC and PCI, (3) precautions after CC or PCI, and (4) essentials of primary or secondary prevention of CAD. The patients were given the link to the social media website, and after searching for the video using keywords, they could watch it online as many times as they wanted on their own mobile devices such as laptops, tablets, or smartphones. To avoid intergroup interference, e.g., the patients in the control group also watched the instructional video online, and we recruited the control group before the intervention group in different periods (Figure 1).

2.2. Instruments. The two instruments used as the indices of holistic healthcare in this study were the EQ-5D and the Functional Assessment of Chronic Illness Therapy–Spiritual Well-Being Scale (FACIT–Sp-12) questionnaires, which measure the respondents' self-reported health status and spiritual well-being, respectively.

2.2.1. EQ-5D-5L. The five-level of ED-5Q version(EQ-5D-5L) is a standardized instrument used to measure generic health status, also involving cardiovascular disease [19, 20], and is composed of two parts: EQ-5D-5L descriptive system and EQ visual analogue scale (EQ-VAS), both measures self-perceived heath status [21, 22]. The EQ-5D-5L is a five-dimension-specific rating scale including a 5-point scale ranging from 1 to 5 according to the severity in each case (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression). A lower total score indicates better subjective health status or quality of life. In contrast, the EQ-VAS is rated on a scale from 0 to 100, indicating the worst to the best imaginable health status.

2.2.2. FACIT–Sp-12. The FACIT–Sp-12 is a self-administered questionnaire supporting three or two factors (peace/meaning and faith). It contains 12 items (eight for peace/meaning, four for faith) with a 4-point Likert scale that measures spiritual well-being in people with chronic illnesses [23, 24]. Every subscale measures different aspects of spiritual well-being: peace for a sense of harmony and peace deriving from a connection deriving from something larger than one's self, mean for a sense of purpose and significance from a connection to something larger than one's self, and faith for a sense of strength and comfort from one's faith and spiritual beliefs [25]. A higher score indicates better spiritual well-being.

2.3. Study Design. The questionnaires were given three times in both groups. First, after the ward-based education and/or watching the online video and before CC. Second, after the procedure but prior to discharge. Third, at the first return visit to the CV OPD after discharge. Apart from the questionnaires, a 90-day follow-up visit after discharge was arranged. By analyzing the questionnaire results and outcomes of postdischarge follow-up between the two groups, the influence of watching the online video before the procedure was investigated.

2.4. Statistics. Baseline characteristics of the intervention and control groups were compared using the independent sample test for continuous variable and the chi-square test for categorical variable. EQ-5D and FACIT–Sp-12 scores were compared between the two groups at each of the three measurements (before catheterization, before discharge, and postdischarge first outpatient visit) using the independent sample test. The readmission rate during 90 days of follow-up was compared between the two groups using the log-rank test. The effect of the intervention was assessed using a linear mixed effects model which included main effects of intercept, covariates, the study group (1 = intervention; 0 = control) and measurement (three time points), and two-way interactions of "group × measurement." There were three random effects: the intercept, the slope of time, and the residual. The selected covariates were age, sex, smoking, coronary artery bypass graft, length of hospitalization, and use of nitrates. The effect of the intervention was confirmed if the two-way interaction effect was significant. All tests were two-tailed, and p value < 0.05 was considered to be statistically significant. No adjustments for multiple testing

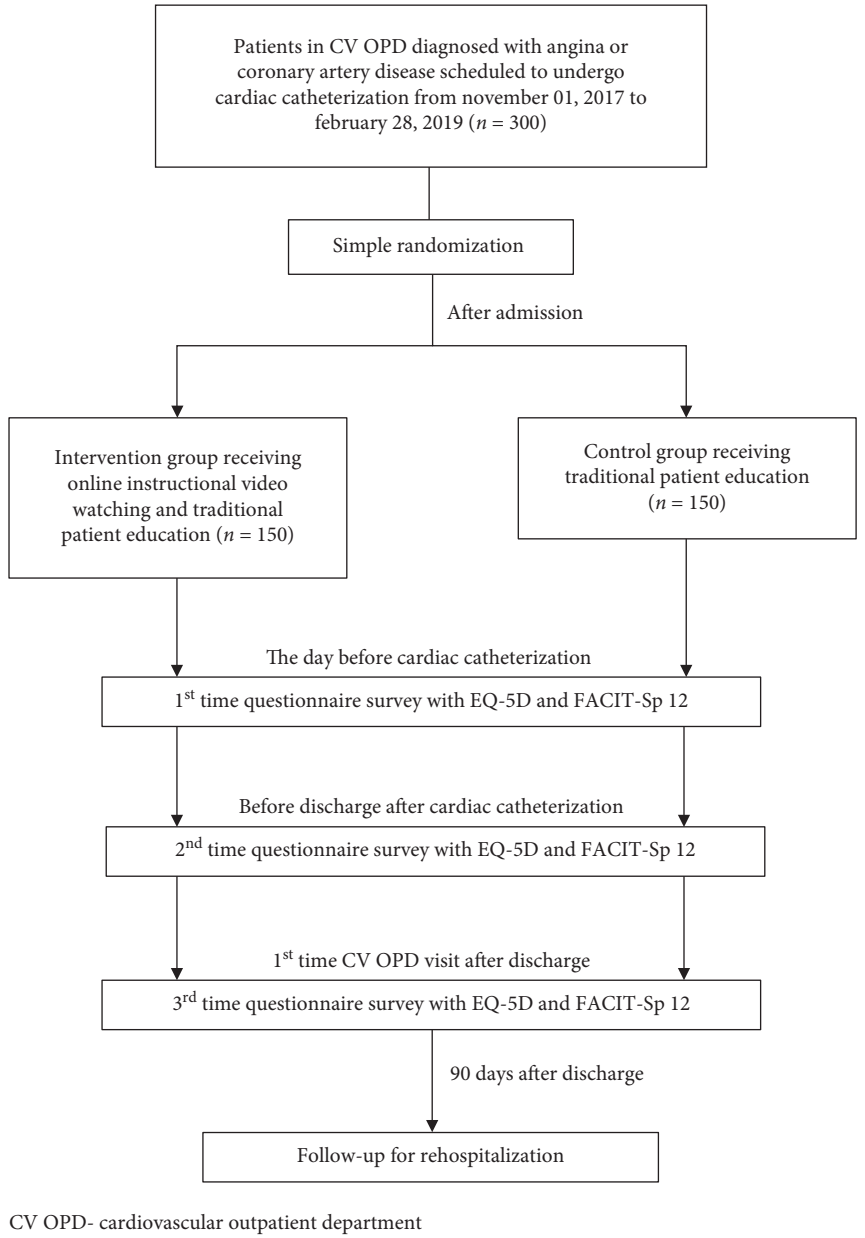


Figure 1: Flow chart of the study.

(multiplicity) were made in this study. Data analyses were conducted using SPSS version 25 (IBM SPSS Inc., Chicago, Illinois). The analytical results are summarized in Table S1.

3. Results

3.1. Baseline Characteristics. From November 1, 2017 through February 28, 2019, a total of 300 patients scheduled to undergo CC were randomly assigned to the intervention and control groups (150 patients in each group). There were no significant differences in baseline characteristics between the two groups except for age ($p = 0.019$) and length of hospital stay ($p = 0.004$). The patients in the intervention group were younger (62.85 years) than those in the control group (65.56 years), and the length of hospitalization was shorter in the intervention group (3.54 days) than in the control group (4.89 days) (Table 1).

3.2. Outcome Measures at Each Measurement. Patients in the intervention group reported significantly lower mean EQ-5D-5L scores than the control group at all three time points ($p = 0.014$, <0.001, and <0.001, respectively). There was no obvious difference in EQ-VAS between the two groups. In terms of spiritual well-being, the FACIT–Sp-12 score in the intervention group was significantly higher than that in the control group at the first return visit to the CV OPD ($p = 0.016$). In addition, the intervention group had better spiritual well-being with regards to faith before discharge ($p < 0.001$) and at first return to outpatient clinic than the control group ($p < 0.001$) (Table 2).

Table 1: Characteristics of the patients at baseline.

Characteristics	Intervention ($n = 150$)	Control ($n = 150$)	All cases ($n = 300$)	p value
Age (years)	62.85 ± 9.84	65.56 ± 10.15	64.2 ± 10.07	0.019
BMI (kg/m^2)	27.22 ± 4.36	27.13 ± 4.77	27.18 ± 4.57	0.873
Male sex (n, %)	126 (84)	114 (76)	240 (80)	0.083
Smoking (n, %)	50 (33.3)	30 (20)	80 (26.7)	0.08
Medical history (n, %)				
Diabetes mellitus	72 (48)	73 (48.7)	145 (48.3)	0.908
Hyperlipidemia	74 (49.3)	66 (44)	140 (46.7)	0.355
Heart failure	16 (10.7)	24 (16)	40 (13.3)	0.174
Stroke	7 (4.7)	10 (6.7)	17 (5.7)	0.454
Peripheral artery disease	1 (0.7)	4 (2.7)	5 (1.7)	0.176
COPD	2 (1.3)	5 (3.3)	7 (2.3)	0.251
Revascularization therapy (n, %)				
Percutaneous coronary intervention	68 (45.3)	70 (46.7)	138 (46)	0.817
Coronary artery bypass grafting	6 (4)	2 (1.3)	8 (2.7)	0.152
Length of hospital stay (days)	3.54 ± 2.55	4.89 ± 5.07	4.22 ± 4.07	0.004
Coronary artery disease medication (n, %)				
Antiplatelet	139 (92.7)	139 (92.7)	278 (92.7)	0.833
ACEI/ARB	90 (60.4)	83 (55.3)	173 (57.7)	0.375
Beta-blocker	107 (73.2)	114 (76)	223 (74.6)	0.572
Mineralocorticoid receptor antagonist	9 (6)	10 (6.7)	19 (6.3)	0.824
Statin	118 (79.2)	117 (78)	235 (78.3)	0.801
Nitrate	46 (30.9)	60 (40)	106 (35.3)	0.099

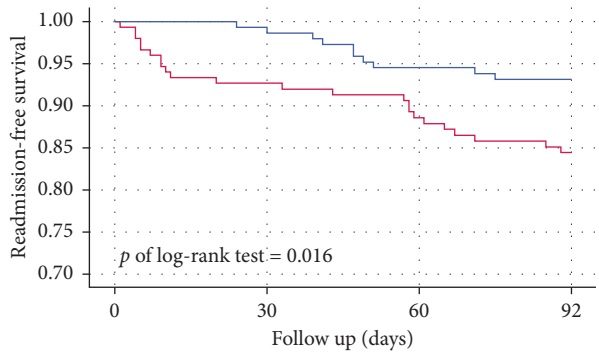
BMI, body mass index; COPD, chronic obstructive pulmonary disease; ACEI, angiotensin-converting enzyme inhibitors; ARB, angiotensin-receptor blockers. Data are presented as mean ± standard deviation or frequency and percentage in parenthesis.

Table 2: Results of the EQ-5D and FACIT–Sp12 for the quality of holistic healthcare.

	Intervention ($n = 150$)	Control ($n = 150$)	All cases ($n = 300$)	p value
EQ-5D				
EQ-5D-5L (5–25)				
Before catheterization	7.19 ± 1.39	7.71 ± 2.15	7.45 ± 1.83	0.014
Before discharge	5.46 ± 0.7	6.16 ± 1.84	5.81 ± 1.43	<0.001
Postdischarge, 1st OPD visit	5.23 ± 0.55	6.0 ± 1.76	5.61 ± 1.35	<0.001
EQ-VAS (0–100)				
Before catheterization	74.79 ± 9.53	73.20 ± 11.86	74 ± 10.77	0.201
Before discharge	78.76 ± 6.78	78.33 ± 8.86	77.55 ± 7.87	0.644
Postdischarge, 1st OPD visit	80.75 ± 5.74	79.67 ± 8.37	79.97 ± 8.28	0.201
FACIT–Sp12 (0–48)				
Before catheterization	31.83 ± 4.17	32.79 ± 4.77	32.31 ± 4.49	0.066
Before discharge	36.63 ± 3.59	35.88 ± 4.3	36.26 ± 3.97	0.106
Postdischarge, 1st OPD visit	37.61 ± 3.38	36.65 ± 4.54	37.13 ± 4.02	0.042
Mean (0–32)				
Before catheterization	22.48 ± 2.80	22.84 ± 3.37	22.66 ± 3.10	0.315
Before discharge	24.10 ± 2.47	24.63 ± 2.89	24.36 ± 2.70	0.090
Postdischarge, 1st OPD visit	24.14 ± 2.40	24.92 ± 3.01	24.53 ± 2.74	0.016
Faith (0–16)				
Before catheterization	9.35 ± 1.98	9.95 ± 2.55	9.65 ± 2.30	0.025
Before discharge	12.53 ± 1.86	11.25 ± 2.15	11.90 ± 2.10	<0.001
Postdischarge, 1st OPD visit	13.46 ± 1.71	11.73 ± 2.08	12.60 ± 2.09	<0.001

OPD, outpatient department; VAS, visual analogue scale; FACIT–SP, the Functional Assessment of Chronic Illness Therapy–Spiritual Well-Being Scale. Data are presented as mean ± standard deviation.

3.3. Readmission Rate. With respect to the postdischarge follow-up at 90 days, the overall readmission rate of patients in the control group was significantly higher than that in the intervention group (18.7% vs. 14%; $p = 0.016$, log-rank test; Figure 2). By analyzing the etiologies of readmission, patients in the control group were more likely to be readmitted within the first 3 months after the index hospitalization due to heart failure decompensation or other comorbidities such



Number of patient at risk:

Control	150	138	129	122
Intervention	150	146	137	134

Figure 2: Kaplan–Meier survival curves of 90-day readmission of patients in the control and intervention groups. The readmission rate within 90 days after discharge was significantly lower in the intervention group (red line) compared with the control (blue line).

as diabetes mellitus, chronic kidney disease, or peripheral artery disease (Table 3).

3.4. Intervention Effect on Outcome Measures. After controlling the selected covariates, including age, sex, smoking, CABG, length of hospitalization, and use of nitrates, the linear mixed effects model demonstrated that the interaction effects of EQ-5D-5L and EQ-VAS between the two groups were not statistically significant, suggesting that the improvement in outcome measures was not superior in the intervention group. However, the improvement in FACIT–Sp-12 was greater in the intervention group than in the control group from baseline to before discharge (regression coefficient (B) = 1.70, $p < 0.001$) and from baseline to postdischarge first CV OPD visit ($B = 1.81$, $p < 0.001$) (Figure 3).

4. Discussion

4.1. Effects of the e-Health Intervention. In this quasiexperimental prospective study involving cardiac outpatients scheduled to undergo CC, the preprocedural intervention of watching an online instructional video could more effectively improve the quality of holistic healthcare before discharge (after CC) and in postdischarge follow-up compared to traditional ward-based patient education.

Despite the improvement of subjective health status in EQ-5D not superior to the control, the patients in the intervention group indeed had significantly lower score in EQ-5D-5L before and after CC and in the first postdischarge visit to OPD. The EQ-5D-5L is widely used to assess general health status and evaluate four physical dimensions and one mental or spiritual dimension about anxiety/depression [21]. Because all the participants in this study were recruited from OPD, their physical condition should have been relatively stable. Accordingly, differences in EQ-5D-5L would result from the mental or spiritual dimension, suggesting less anxiety/depression the intervention group had at all the three timepoints.

The improvement of FACIT–Sp-12 from baseline to predischarge and first postdischarge OPD visit also corresponded to higher spiritual well-being in the intervention group. Interestingly, the intervention group had significantly lower faith subscale scores than the control before CC. It implicated that the online instructional video watching made the patients worried, which might come from a better insight of coronary artery disease obtained from the video content. For the same reason, the intervention group had significantly higher faith subscale scores before discharge and at the first OPD return visit.

The results of the 90-day postdischarge follow-up showed a higher total readmission rate in the control group, especially admissions for heart failure decompensation and other comorbidities. This may be due to the instructional online video containing shared information about CAD prevention with regards to heart failure and other vascular comorbidities such as diabetes and peripheral artery disease. On the other hand, the older average age in the control group may also have contributed to the admission rate for heart failure and comorbidities.

4.2. Barriers in the e-Health Implementation. The barriers in implementing e-health included technology disconnect and lack of the holistic approach [26]. Previous studies have shown that group patient education or education using multimedia such as a videotape or DVD before CC can improve spiritual well-being and satisfaction in patients scheduled to receive CC [27–29]. The biggest difference in this study is the application of Internet-based social media as the main tool for patient education, which overcomes the limits on time and space in addition to provision of standardized and comprehensive contents. The advantages of this tool include good accessibility, cost, and time efficiencies, and that the content can be accessed using a variety of devices. The patients can then select to watch the

Table 3: Readmission during 90 days of follow-up.

Postdischarge, 90-day follow-up	Intervention ($n = 150$)	Control ($n = 150$)	All cases ($n = 300$)	p value
Readmission, (n, %)	14 (9.3)	28 (18.7)	42 (14)	0.016
Etiology of readmission				—
Expected readmission	4 (2.7)	5 (3.3)	9 (3)	
Acute coronary syndrome	3 (2)	1 (0.7)	4 (1.3)	
Heart failure	0 (0)	5 (3.3)	5 (1.7)	
Infection	2 (1.4)	5 (3.3)	7 (2.3)	
Comorbidities (PAD, CKD, and DM)	0 (0)	5 (3.3)	5 (1.7)	
Others	5 (3.4)	7 (4.7)	12 (4)	

PAD, peripheral artery disease; CKD, chronic kidney disease; DM, diabetes mellitus.

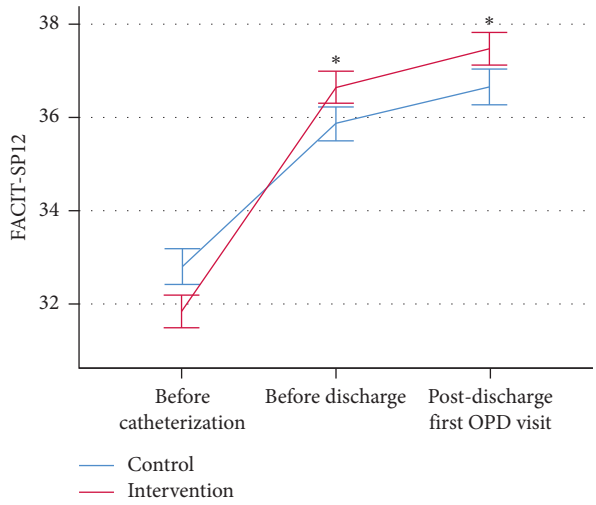


Figure 3: The means and standard errors of FACIT–SP12 of the intervention and control groups at each measurement. The improvement of the total FACIT–Sp-12 score from baseline was significantly more in the intervention group before discharge and in the first postdischarge visit to OPD. FACIT–SP, the Functional Assessment of Chronic Illness Therapy–Spiritual Well-Being Scale; *indicates that the improvement from baseline to follow-up was significantly different between the two groups ($p < 0.05$).

corresponding video scenes at anytime and anywhere to remind themselves about the periprocedural precautions or knowledge of CAD prevention in which they are interested in. Repeatedly watching the video will reinforce its contents and improve the level of self-care. This may be why there was stronger faith showed in FACIT–Sp-12 in the intervention group.

4.3. Prospects of e-Health in Holistic Healthcare. Holistic healthcare describes patient-centered approaches and interventions that are meant to satisfy a patient's physical, mental, emotional, and spiritual needs [30]. The primary goal for worldwide healthcare intervention is patient engagement [31], which is characterized by three dimensions: a behavior dimension (what the patient dose), a cognitive dimension (what the patient thinks and knows), and an emotional dimension (what the patient feels) [32]. The holistic approach with e-health interventions should be devoted to foster patient engagement in the future. Although the effects of the social media intervention in this study almost meet the above three dimensions for the patients with CAD, it does have room for improvement. In view of the rise of artificial intelligence (AI) in healthcare applications, AI is changing the way how healthcare is delivered, especially in personalized medicine and access to recommendations and automated treatments [33]. With the incorporation of AI and e-health, more novel patient-centered interventions other than social media application may be developed to provide more comprehensive and cost- and time-efficient accesses to satisfy the patient's need.

5. Limitations

The present study had several limitations. Based on a prospective quasiexperimental nonequivalent group design, selection bias of participants assigned to the control group and intervention group in different recruitment periods was inevitable in this study and might reduce the internal validity of this study [34]. In addition, many devices were able to connect to the Internet. The patients in the control group still could obtain the information about CC, PCI, or CABG by Internet search via their own mobile devices after traditional patient education in ward, which may add interference to the result of study and weaken the effect of the online institutional video watching as the measure of intervention in this study.

6. Conclusions

Access to the Internet is ubiquitous in most countries, and it continues to have an ever increasing impact on our daily life. The utilization of e-health/m-health efficiently provided information about CC to our patients. Our findings indicate that the more patients know, the better their spiritual well-being and the higher their satisfaction. To improve holistic healthcare in the modern era, the results of this study may provide some basis for the application of e-health to educate patients undergoing CC and potentially for other invasive procedures in other specialties.

Abbreviations

e-health: Electronic health
CVD: Cardiovascular disease
CAD: Coronary artery disease
CC: Cardiac catheterization
PCI: Percutaneous coronary intervention.

Data Availability

All the datas are available in the text.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Jian-Rong Peng, Hung-Chi Su, and Pao-Hsien Chu conceptualized the study and performed formal analysis. Jian-Rong Peng, Hung-Chi Su, Siou-Ling Gong, and Pao-Hsien Chu collected data. Jian-Rong Peng, Hung-Chi Su, Chia-Pin Lin, Chun-Chi Chen, Chi-Jen Chang, Siou-Ling Gong, and Pao-Hsien Chu investigated and administered the study. Siou-Ling Gong and Pao-Hsien Chu collected the resources. Pao-Hsien Chu supervised, reviewed, and edited the article. Peng and Hung-Chi Su wrote the original draft of the study and equally contributed to this study.

Acknowledgments

The authors would like to thank Alfred Hsing-Fen Lin for the statistical assistance during the completion of this article. This work has been supported by Chang Gung Memorial Hospital grant, #CORPG3G108 to Pao-Hsien Chu.

Supplementary Materials

Table S1. Multivariate linear mixed model to evaluate the intervention effect on the outcome measures. (*Supplementary Materials*)

References

[1] C. Pagliari, D. Sloan, P. Gregor et al., "What is eHealth (4): a scoping exercise to map the field," *Journal of Medical Internet Research*, vol. 7, no. 1, p. e9, 2005.

[2] C. Free, G. Phillips, L. Felix, L. Galli, V. Patel, and P. Edwards, "The effectiveness of M-health technologies for improving health and health services: a systematic review protocol," *BMC Research Notes*, vol. 3, p. 250, 2010.

[3] E. P. Ambinder, "Electronic health records," *Journal of Oncology Practice*, vol. 1, no. 2, pp. 57–63, 2005.

[4] M. R. Cowie, J. Bax, N. Bruining et al., "e-Health: a position statement of the European Society of Cardiology," *European Heart Journal*, vol. 37, no. 1, pp. 63–66, 2016.

[5] I. Frederix, E. G. Caiani, P. Dendale et al., "ESC e-cardiology working group position paper: overcoming challenges in digital health implementation in cardiovascular medicine," *European Journal of Preventive Cardiology*, vol. 26, no. 11, pp. 1166–1177, 2019.

[6] M. F. Piepoli, A. W. Hoes, S. Agewall et al., "2016 European guidelines on cardiovascular disease prevention in clinical practice: the sixth joint task force of the European society of Cardiology and other societies on cardiovascular disease prevention in clinical practice (constituted by representatives of 10 societies and by invited experts) developed with the special contribution of the European association for cardiovascular prevention & rehabilitation (EACPR)," *Atherosclerosis*, vol. 252, pp. 207–274, 2016.

[7] GBD 2015 Mortality and Causes of Death Collaborators, "Global, regional, and national life expectancy, allause mortality, and cause-pecific mortality for 249 causes of death, 1980–2015: a systematic analysis for the Global burden of disease study 2015," *Lancet*, vol. 388, no. 10053, pp. 1459–1544, 2016.

[8] G. A. Roth, C. Johnson, A Abajobir et al., "Global, regional, and national burden of cardiovascular diseases for 10 causes, 1990 to 2015," *Journal of the American College of Cardiology*, vol. 70, no. 1, pp. 1–25, 2017.

[9] J. Knuuti, W. Wijns, A. Saraste et al., "2019 ESC guidelines for the diagnosis and management of chronic coronary syndromes," *European Heart Journal*, vol. 41, no. 3, pp. 407–477, 2020.

[10] F. J. Neumann, M. Sousa-Uva, A. Ahlsson et al., "2018 ESC/EACTS guidelines on myocardial revascularization. The task force on myocardial revascularization of the European society of Cardiology (ESC) and European association for cardiothoracic surgery (EACTS)]," *Giornale Italiano di Cardiologia*, vol. 20, no. 1, pp. 1s–61s, 2019.

[11] Z. K. Nekouei, A. Yousefy, G. Manshaee, and S. Nikneshan, "Comparing anxiety in cardiac patients candidate for angiography with normal population," *ARYA Atherosclerosis*, vol. 7, no. 3, pp. 93–96, 2011.

[12] M. Ahmad and S. M. Ayasrah, "Periprocedural anxiety associated with cardiac catheterisation and the monitored physiological measures," *Journal of Perioperative Practice*, vol. 30, no. 5, pp. 130–134, 2019.

[13] H. M. Arthur, K. M. Smith, and M. K. Natarajan, "Quality of life at referral predicts outcome of elective coronary artery angiogram," *International Journal of Cardiology*, vol. 126, no. 1, pp. 32–36, 2008.

[14] M. S. Eastin and N. M. Guinsler, "Worried and wired: effects of health anxiety on information-seeking and health care utilization behaviors," *CyberPsychology & Behavior*, vol. 9, no. 4, pp. 494–498, 2006.

[15] R. Gallagher, R. Trotter, and J. Donoghue, "Preprocedural concerns and anxiety assessment in patients undergoing coronary angiography and percutaneous coronary interventions," *European Journal of Cardiovascular Nursing*, vol. 9, no. 1, pp. 38–44, 2010.

[16] S. Ercan, A. Unal, G. Altunbas et al., "Anxiety score as a risk factor for radial artery vasospasm during radial interventions," *Angiology*, vol. 65, no. 1, pp. 67–70, 2014.

[17] R. Kruizinga, M. Scherer-Rath, H. J. B. A. M. Schilderman, C. M. Puchalski, and H. H. W. M. van Laarhoven, "Toward a fully fledged integration of spiritual care and medical care," *Journal of Pain and Symptom Management*, vol. 55, no. 3, pp. 1035–1040, 2018.

[18] T. D. Cook and D. T. Campbell, *Quasi-Experimentation: Design and Analysis Issues for Field Settings*, Houghton Mifflin, Boston, MA, USA, 1979.

[19] J. J. Ellis, K. A. Eagle, E. M. Kline-Rogers, and S. R. Erickson, "Validation of the EQ-5D in patients with a history of acute coronary syndrome," *Current Medical Research and Opinion*, vol. 21, no. 8, pp. 1209–1216, 2005.

[20] D. Nowels, J. McGloin, J. M. Westfall, and S. Holcomb, "Validation of the EQ-5D quality of life instrument in patients after myocardial infarction," *Quality of Life Research*, vol. 14, no. 1, pp. 95–105, 2005.

[21] G. Balestroni and G. Bertolotti, "EuroQol-5D (EQ-5D): an instrument for measuring quality of life," *Monaldi Archives for Chest Disease*, vol. 78, no. 3, pp. 155–159, 2012.

[22] M. F. Janssen, A. S. Pickard, D. Golicki et al., "Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study," *Quality of Life Research*, vol. 22, no. 7, pp. 1717–1727, 2013.

[23] S. Monod, M. Brennan, E. Rochat, E. Martin, S. Rochat, and C. J. Büla, "Instruments measuring spirituality in clinical research: a systematic review," *Journal of General Internal Medicine*, vol. 26, no. 11, pp. 1345–1357, 2011.

[24] A. H. Peterman, G. Fitchett, M. J. Brady, L. Hernandez, and D. Cella, "Measuring spiritual well-being in people with cancer: the functional assessment of chronic illness therapy-spiritual well-being scale (FACIT-Sp)," *Annals of Behavioral Medicine*, vol. 24, no. 1, pp. 49–58, 2002.

[25] A. R. Munoz, J. M. Salsman, K. D. Stein, and D. Cella, "Reference values of the functional assessment of chronic illness therapy-spiritual well-being: a report from the American cancer society's studies of cancer survivors," *Cancer*, vol. 121, no. 11, pp. 1838–1844, 2015.

[26] B. Alkhaldi, T. Sahama, C. Huxley, and R. Gajanayake, "Barriers to implementing eHealth: a multi-dimensional perspective," *Studies in Health Technology and Informatics*, vol. 205, pp. 875–879, 2014.

[27] D. S. Chan and H. W. Cheung, "The effects of education on anxiety among Chinese patients with heart disease undergoing cardiac catheterization in Hong Kong," *Contemporary Nurse*, vol. 15, no. 3, pp. 310–320, 2003.

[28] T. M. Davis, T. O. Maguire, M. Haraphongse, and M. R. Schaumberger, "Undergoing cardiac catheterization: the effects of informational preparation and coping style on patient anxiety during the procedure," *Heart Lung*, vol. 23, no. 2, pp. 140–150, 1994.

[29] K. L. Wu, S. R. Chen, W. C. Ko et al., "The effectiveness of an accessibility-enhanced multimedia informational educational programme in reducing anxiety and increasing satisfaction of patients undergoing cardiac catheterisation," *Journal of Clinical Nursing*, vol. 23, no. 13-14, pp. 2063–2073, 2014.

[30] M. Jasemi, L. Valizadeh, V. Zamanzadeh, and B. Keogh, "A concept analysis of holistic care by hybrid model," *Indian Journal of Palliative Care*, vol. 23, no. 1, pp. 71–80, 2017.

[31] W. Hardyman, K. L. Daunt, and M. Kitchener, "Value co-creation through patient engagement in health care: a micro-level approach and research agenda," *Public Management Review*, vol. 17, no. 1, pp. 90–107, 2015.

[32] S. Barello, S. Triberti, G. Graffigna et al., "eHealth for patient engagement: a systematic review," *Frontiers in Psychology*, vol. 6, p. 2013, 2016.

[33] J. Cabestany, D. Rodriguez-Martín, C. Pérez, and A. Sama, "Artificial intelligence contribution to eHealth application," in *Proceedings of the 2018 25th International Conference on Mixed Design of Integrated Circuits and System (MIXDES)*, pp. 15–21, Gdynia, Poland, 2018.

[34] K. J. Flannelly, L. T. Flannelly, and K. R. B. Jankowski, "Threats to the internal validity of experimental and quasi-experimental research in healthcare," *Journal of Health Care Chaplaincy*, vol. 24, no. 3, pp. 107–130, 2018.

Hindawi
Journal of Healthcare Engineering
Volume 2021, Article ID 6658538, 15 pages
https://doi.org/10.1155/2021/6658538

Research Article

A Hybrid MADM Model for Newly Graduated Nurse's Competence Evaluation and Improvement

Fengmin Cheng,[1] Yanjun Jin,[1] Ching-Wen Chien,[2] Lei Xiong ,[3] and Yen-Ching Chuang [4]

[1]*Taizhou Hospital of Zhejiang Province Affiliated to Wenzhou Medical University, Taizhou, China*
[2]*Institute for Hospital Management, Tsinghua University, Shenzhen Campus, Beijing, China*
[3]*School of Architecture and Applied Art, Guangzhou Academy of Fine Arts, Guangzhou, China*
[4]*Institute of Public Health & Emergency Management, Taizhou University, Taizhou, China*

Correspondence should be addressed to Lei Xiong; kmt20005@gmail.com and Yen-Ching Chuang; yenching.chuang@gmail.com

Received 15 December 2020; Revised 31 January 2021; Accepted 22 February 2021; Published 9 March 2021

Academic Editor: Chi-Hua Chen

Nursing departments in hospitals must evaluate the practical competency of newly graduated nurses and assist them to increase their competence. Competency assessments often consider multiple qualitative attributes and use expert knowledge as the basis for decision-making. This study proposes a hybrid multiple attribute decision-making (MADM) model that determines practical competency of the newly graduated nurse as an evaluation framework. A causal influence-network diagram (CIND) and influential weights are obtained from nursing experts' clinical experience using the decision-making trial and evaluation laboratory (DEMATEL)-based analytical network process analysis (DANP). The MOORA-AS method is then used to evaluate the ability expectation ratio-gap for newly graduated nurses in practice. The CIND is used to allow each newly graduated nurse to reduce the performance ratio-gaps between the current level and the aspirational level from a systematic perspective. The empirical data applies to a third-class and a first-class hospital in China. The results show that the proposed hybrid MADM model has reliable results and allows nursing department decision-makers/managers to easily evaluate and systematically improve competencies for newly graduated nurses.

1. Introduction

The 2019 Coronavirus (COVID-19) pandemic has had a devastating impact on a global scale on the public health, the economy, the labor market, and other aspects of personal and social life [1]. During global health emergencies, nurses constitute the largest professional group for the provision of first-line healthcare [2, 3]. The clinical nursing and management abilities of nurses are key to identifying, isolating, and managing COVID-19 patients and to providing support for non-COVID-19 patients [3, 4]. 2020 has been named the International Year of the Nurse and Midwife by World Health Organization (WHO) [5].

Resignations among nurses are increasingly common [6]. This fact is worrying because high turnover rates lead to a shortage of nurses and reduce the quality of care, which can lead to adverse patient outcomes [7, 8]. However, more than 20% of all newly graduated nurses leave their jobs within one year, which exacerbates the shortage of nurses [7]. 27.6% of all newly graduated nurses in the United States left their jobs in 2019 [9]. Some studies list the factors that influence the intention of newly graduated nurses to leave their jobs, including working environment, clinical ability and support level, effective leadership, team cohesion, job satisfaction, job stress and coping with self-efficacy, rewards and recognition, and professional development opportunities [7, 10, 11]. Poor ability is one of the main reasons for intended and actual resignations among newly graduated nurses [11, 12]. From the outbreak of Severe Acute Respiratory Syndrome (SARS) in 2003 until the global epidemic involving COVID-19 2019, the increase in the risk of nosocomial infection, the change in nursing requirements, and the increased difficulty of the

job increase the willingness of new graduates to leave their profession.

In many cases, newly graduated nurses have insufficient knowledge and experience to supplement their expected job responsibilities, and their knowledge and experience are insufficient to follow practice [13, 14]. This occurs because the profession cannot fully identify and integrate nursing competencies and instill them in the early practice for newly graduated nurses [7]. Practical projects and processes that allow newly graduated nurses to progress have been extensively studied [15–17]. However, the level of competency for newly graduated nurses after graduation that meets the needs in practice is unclear [18], so nursing management faces new challenges.

One study systematically reviewed the assessment of clinical nursing competence and found that the definition of competence is vague and lacks the reliability and validity of systematic measurement tools or strategies [19]. In terms of evaluating competency, studies use multiple attribute decision-making (MADM) methods to evaluate personnel, for selection and for improvement. Chen et al. [20] used the analytical hierarchy process (AHP) to evaluate human resources performance in the logistics market. De Moura and Sobral [21] used the elimination and choice translating reality (ELECTRE) method to classify the competency of employees in a call center. Some MADM methods use interdependence relation perspective, such as the decision-making trial and evaluation laboratory (DEMATEL) [22], the analytical network process analysis (ANP) [23, 24], and the DEMATEL-based analytical network process analysis (DANP) [25]. MADM methods are used for decision-making processes to evaluate personnel, but no studies evaluate the ability of new nurses.

This study develops a hybrid MADM model. The DMATEL method is used to confirm the interaction between attributes and to determine the causal relationship. The results are visualized using a plot, called a causal influence-network diagram (CIND). The influence weights of interdependence between attributes are then derived using the DANP method. The MOORA-AS method is used to evaluate the practical competency of three newly graduated nurses and to determine the attribute of the largest aspirated ratio-gap. The CIND is then used to decrease the maximum aspirated ratio-gap for each newly graduated nurse, and the improvement strategy is based on the system perspective. The hybrid MADM model makes the following contributions to this field of study:

(i) The CIND visualizes the influential relationship structure within attributes and identifies the key factors among abilities from a systematic perspective in the real world.

(ii) The CIND is used as an improvement tool with a systematic perspective to allow nursing department decision-makers/managers to propose causal improvement suggestions that correspond to each newly graduated nurse's competency. This improvement strategy also avoids the problem of "treating the head when the head aches and treating the foot when the foot hurts" [26].

(iii) The concept of an aspiration level is used to prevent "choosing good apples from rotten apples," which is a feature of the original MOORA method [27].

The MADM methodology for this study allows nursing researchers to apply various MADM methods to future nursing-related management decision-making problems. The nursing departments of a third-class and first-class hospital in China are used as an empirical case. The results show that the model allows nursing department decision-makers/managers to propose appropriate strategies and improvements to prepare newly graduated nurses for practice.

The remainder of this paper is divided into five sections. Section 2 details the competency framework for a newly graduated nurse and the use of MADM models to evaluate the competence of personnel and the background of the MADM methods for the proposed model. Section 3 presents the hybrid MADM model and its corresponding approaches. Section 4 describes the computational process for the empirical case using the proposed model. Section 5 discusses the results of Section 4. Finally, the conclusion is presented in Section 6.

2. Literature Review

Evaluating the competence of a newly graduated nurse is a personnel assessment problem that involves MADM. It often involves intangible factors and is based on the practical experience of domain-experts. The competency framework for a newly graduated nurse is firstly described. The current use of MADM models for personnel assessment is also discussed. The MADM methods for the proposed hybrid MADM model are then detailed.

2.1. The Competency Framework for a Newly Graduated Nurse. Practice readiness refers to the comfort level of new graduates in assuming the professional duties of nurses (professional abilities) [28]. It also refers to thinking like a nurse (professional ability) [29]. The ability index system of a newly graduated nurse is based on the results of the conceptual analysis methods of Mirza et al. [30]. This study uses several databases (CINAHL, PubMed, EBM, and ProQuest) and adopts Rodgers' evolutionary method of conceptual analysis to combine the results of these databases into a conceptual model for the practical preparation of newly graduated nurses. A newly graduated nurse who is preparing to become a qualified nurse must demonstrate cognitive capability (C_1), clinical capability (C_2), and professional capability (C_3) as main competency characteristics. These competencies are the abilities that a newly graduated nurses needs before undertaking professional duties. Nurses must also be able to cope with new and unfamiliar nursing situations so they must evaluate and improve existing abilities. These cognitive, clinical, and professional capabilities contribute to a newly graduated nurse's sense of self-efficacy and a sense of practice readiness [31]. These three main capabilities are described in detail in Table 1.

Table 1: The competency framework for newly graduated nurses to become practice-ready.

Aspect	Attribute	Description	Ref.
Cognitive capability (C_1)	Competence (C_{11})	Newly graduated nurses know how to use nursing knowledge in particular contexts or situations	[32, 33]
	Effective clinical and moral reasoning (C_{12})	Newly graduated nurses should exhibit critical thinking and problem-solving skills to make effective clinical judgments in different clinical situations, and they should also be expected to use nursing procedures to plan and evaluate effective nursing measures for patients. Moral reasoning plays a role in the formulation of professional practice.	[28, 29, 34–36]
	Situational awareness (C_{13})	Newly graduated nurses should be able to determine their limitations in different clinical situations and seek help when needed. This cognitive ability of situational awareness is crucial to providing safe nursing.	[34, 37]
Clinical capability (C_2)	Psychomotor skills (C_{21})	Newly graduated nurses must have the necessary mental skills to be psychologically prepared for practical nursing work	[28, 29, 35, 38]
	Clinical assessment (C_{22})	Clinical evaluation is an important skill that reflects clinical ability, including analyzing a patient's condition, detecting changes in a patient's condition, predicting problems, and identifying a patient's needs	[29, 36, 39]
	Care delivery and coordination (C_{23})	Soft skills to provide patient care and coordinate and cooperate with colleagues in the workplace, such as effective interpersonal communication, teamwork, and time management, are also basic requirements for providing effective patient care	[28, 36, 38, 40]
Professional capability (C_3)	Development of professional identity (C_{31})	The development of the professional identity of newly graduated nurses allows them to fulfil the multiple roles of professional nurses, such as administration, clinical and nursing research	[28, 37]
	Self-esteem, self-worth, and self-efficacy (C_{32})	Newly graduated nurses should have self-esteem, self-worth, and the ability to operate independently and be responsible for their nursing practices. The individual's belief in their readiness for practice is also an important element of the various capabilities.	[35–37, 41]
	Ethical practice (C_{33})	Newly graduated nurses must also have a moral and ethical code that is based on moral reasoning components (cognitive abilities).	[29]

2.2. The MADM Studies Involving Newly Graduated Nurses. In the past few decades, the MADM methods have been widely applied to evaluate and improve personnel competency. Studies that use expert knowledge as a decision-making basis can be roughly divided into three categories: (1) crisp MADM methods, (2) uncertain MADM methods, and (3) other advanced hybrid methods.

Crisp MADM methods. This category of research establishes the function of a decision-making model (i.e., selection or improvement) and uses different decision-making viewpoints, such as preference relationship or influence relationship, so it is a crisp MADM model. For example, Chen et al. [20] used the AHP method to evaluate human resources performance in the logistics market. De Moura and Sobral [21] used the ELECTRE method to classify the competency of employees in a call center. El-Santawy and Ahmed [42] addressed the personnel selection problem using the SDV-MOORA method. The basis of these MADM models is independent relationships, which is inconsistent with the actual situation. Studies regard interdependence as the premise for modeling. Ishizaka and Pereira [23] described a management system that uses MADM analysis and visualization technology, including ANP and PROMETHEE methods. Kucukaltan et al. [24] developed a decision model to identify and prioritize key factors in the logistics industry using a balanced scorecard (BSC) and ANP methods. Aksakal and Dağdeviren [22] developed an integrated approach for personnel selection using DEMATEL and ANP methods. This type of studies establishes the functions of the decision-making model and does not consider the fuzziness of human semantic expression.

Uncertain MADM methods. These consider various ambiguous phenomena and extend the decision-making function using crisp MADM models for the fuzzy expression of human semantics. These are termed fuzzy MADM models. Skrzypek and Dąbrowski [43] used the fuzzy AHP approach to organize an effective team from a group of employees. Karabasevic et al. [44] proposed an uncertain personnel selection framework using SWARA and ARAS methods. Gilan et al. [45] used an interval type-2 fuzzy set for selection of personnel in construction companies using a hierarchy of competency. Wu and Wang [46] combined a cross-entropy and neutrosophic set for the selection of middle-level managers. Canós et al. [47] used fuzzy set methods for personnel selection by evaluating competencies. This type of studies takes into account the ambiguity of human semantic expression, so the decision model uses an independent relationship perspective.

Other advanced hybrid methods. These studies solve complex decision-making problems using a combination of methods, including mathematical methods, soft computing, and machine learning methods. The models are also called advanced hybrid decision-making models. Patel and Jha [48]

used an artificial neural network (ANN) to evaluate and predict the work behavior of employees at construction projects. Ahmed et al. [49] used a fuzzy inference system (FIS) to evaluate employee performance. Hanna et al. [50] proposed a general mathematical formula to reliably weigh the capabilities of different project managers in the construction industry. Bohlouli et al. [51] developed a competency evaluation model as a human resource management expert system using mathematical methods. For this type of study, mathematical methods can only process quantitative data. Data mining then requires a large amount of survey data.

2.3. DEMATEL-Based ANP Method. The decision-making trial and evaluation laboratory (DEMATEL) is a structural analysis technique that was developed by Battelle Memorial Institute to solve complex structural problems of society in the real world [52, 53]. The results for the DEMATEL method allow decision-makers to understand the contextual and interdependent relationships between elements in the system effectively [54]. Lee et al. [55] extended its application value and further proposed a new weight analysis method, which is based on the concept and calculation of the ANP method [56], namely, the DEMATEL-based ANP or DANP method. The DANP method provides the causal influence-network diagram (CIND) and influence weight from a systematic perspective. Now it is used for various decision-making problems, such as cloud service applications [57], workload stress problems [58], and design of public open space [59].

2.4. MOORA-AS Method. MOORA began as an evaluation method that is stable. It was developed by Brauers and Zavadskas [60]. This method combines two different analytical viewpoints: the ratio system and the reference point. The former represents the overall performance of an evaluation system, which allows decision-makers understand the performance of all alternatives. The latter represents the worst performance for attributes in the system. The system is stable in terms of performance ranking, so it is also used for supplier selection [61], failure prioritization [62], and financial services [63]. Liou et al. [64] further integrated the concept of an aspiration level to solve the defect of the reference point in the original MOORA method (i.e., picking a good apple from a group of rotten apples), which is the MOORA-AS method. The MOORA-AS method gives robust results in terms of solving the ranking and selecting alternatives but also reduces the performance ratio-gap for each attribute, so it is close to zero and reaches the aspiration level.

3. The Hybrid MADM Model

The proposed newly graduated nurse's competence framework (as shown in Table 1) is used to develop a hybrid MADM evaluation model, which uses the clinical experience of nursing experts (i.e., administrators, head nurses, and instructors of nurses), the interdependent relationship structure, and the weight between attributes in the real world. The modeling process for the proposed hybrid MADM model uses the DEMATEL, DANP, and MOORA-AS methods. There are three main modeling stages, as shown in Figure 1.

The first stage. The DEMATEL method is used to construct the interdependent relationship structure between attributes.

This stage derives the total influence matrix and the causal influence-network diagram (CIND) between attributes using the clinical experience of an expert group. The computing process for this method follows the following steps:

Step 1: constructing an initial influence relation matrix **H** using clinical nursing experience.

Using the competency framework with n attributes, clinical nursing experts use the five-point response scale with language variables (as shown in Table 2) to respond in pairs to the initial influence relation degree between attributes to obtain a nonnegative initial influence relation matrix $\mathbf{H} = [h_{ij}]_{n\times n}$, as shown in equation (2).

$$\mathbf{H} = \left[h_{ij}\right]_{n\times n} = \left[\frac{1}{K}\sum_{k=1}^{K} o_{ij}^{k}\right]_{n\times n}, \quad i, j \in \{1, 2, \ldots, n\}, \tag{1}$$

where k represents the number of clinical nursing experts and o_{ij}^{k} represents the initial influence relation matrix for the k-th expert.

Step 2: calculating the normalized influence relation matrix **S**.

Before calculating each subsequent degree of influence for matrix **H**, the influence boundary/limited scope is defined and converted into the matrix **S**, as shown in equations (2) and (3).

$$\mathbf{S} = \left[s_{ij}\right]_{n\times n} = \left[\frac{h_{ij}}{\varepsilon}\right]_{n\times n}, \tag{2}$$

$$\varepsilon = \max\left[\left(\max_{1\le j\le n}\sum_{j=1}^{n} h_{ij}, \max_{1\le i\le n}\sum_{i=1}^{n} h_{ij}\right)\right]_{n\times n}, \tag{3}$$

where the values in matrix **S** are between 0 and 1 and the maximum sum of a row or a column is always equal to 1.

Step 3: exporting the total influence relation matrix **A**.

Under the condition of matrix **S**, the influence relation degree (direct and indirect) between attributes is continuously calculated using the Markov chain process. If $\lim_{\Omega\longrightarrow\infty}\mathbf{S}^{\Omega} = [0]_{n\times n}$, the total influence relation matrix is obtained, as shown in equation (4).

$$\mathbf{A} = \mathbf{S}^{1} + \mathbf{S}^{2} + \cdots + \mathbf{S}^{\Omega}, \quad \text{when } \lim_{\Omega\longrightarrow\infty}\mathbf{S}^{\Omega} = [0]_{n\times n}, \tag{4}$$

where **I** is an $n \times n$ identity matrix.

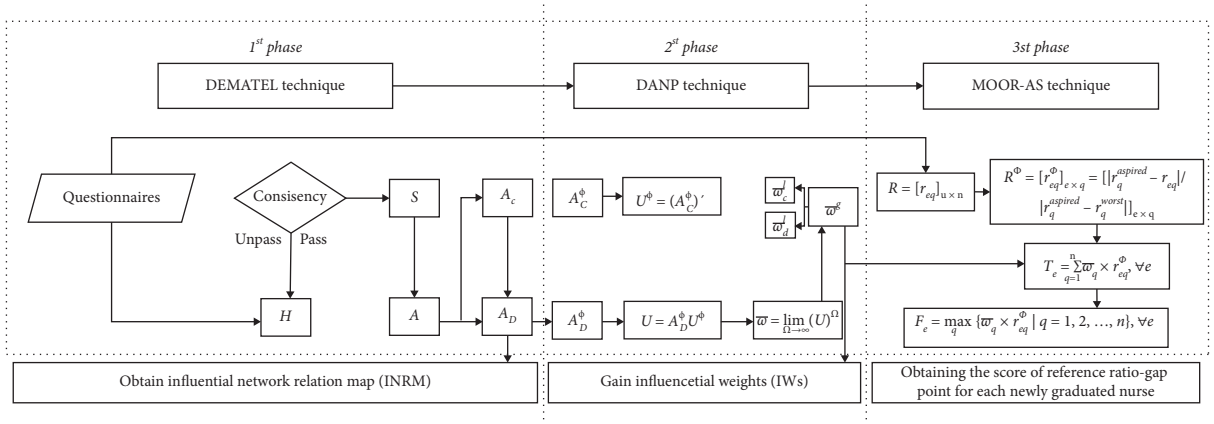


Figure 1: The three main stages of the hybrid MADM model.

Table 2: Five-point response scale using linguistic variables.

Linguistic variables	Crips number
No influence (NI)	0
Weak influence (WI)	1
Medium influence (MI)	2
Strong influence (SI)	3
Very strong influence (VSI)	4

Step 4: creating the causal influence-network diagram (CIND).

Using matrix **A**, four different influence indices are created for each attribute. The equations are shown as (5)–(8), and their corresponding explanations are as follows:

Given influence:

$$p_i = (p_i)_{n\times 1} = \left[\sum_{j=1}^{n} a_{ij}\right]_{n\times 1}. \tag{5}$$

Received influence:

$$y_i = (y_i)_{n\times 1} = \left(y_j\right)_{1\times n}^{\Gamma} = \left[\sum_{i=1}^{n} a_{ij}\right]_{1\times n}^{\Gamma}. \tag{6}$$

Prominence:

$$p_i + y_i. \tag{7}$$

Relation:

$$p_i - y_i, \tag{8}$$

where $i, j \in \{1, 2, \ldots, n\}$; let $i = j$; superscript Γ is denoted transpose.

The "given influence (p_i)" is the total influence degree of the i-th attribute on other attributes; otherwise, the "received influence (y_i)" is the total influence degree of the i-th attribute due to other attributes. "Prominence $(p_i + y_i)$" is the total strength of the given and received influence for the i-th attribute in a clinical nursing environment. "Relation $(p_i - y_i)$" describes the properties of the main influence of the i-th attribute in the clinical nursing environment. If the value of $p_i - y_i$ is positive, then the i-th attribute affects other attributes. If the value is negative, the i-th attribute is influenced by other attributes. A causal influence-network diagram (CIND) is created using "prominence $(p_i + y_i)$" and "relation $(p_i - y_i)$."

The second stage: the DANP method is used to derive the influential weight using the interdependent relationship structure between attributes.

The DANP method is used to derive the influential weights from the results for the total influence relation matrix **A** using the DEMATEL method. If the capability framework is hierarchical, matrix **A** has two levels: the matrix with attribute level $\mathbf{A}_C$ and aspect level $\mathbf{A}_D$. The computing process for this method follows the following steps:

Step 1: transferring the unweighted supermatrix $\mathbf{A}_C^{\Phi *}$ from the total influence relation matrix $\mathbf{A}_C$.

The proportion and transposition of influence relations at different aspect levels mean that this conforms to the ANP method calculation process. The normalized total influence relation matrix $\mathbf{A}_C^{\Phi}$, which is also known as unweighted supermatrix $\mathbf{A}_C^{\Phi *}$, is shown in

$$A_C^{\Phi*} = (A_C)^{\Gamma} = \begin{array}{c} \\ D_1 \\ \vdots \\ D_i \\ \vdots \\ D_m \end{array} \begin{array}{c} \begin{array}{cccccc} D_1 & \cdots & D_j & \cdots & D_m \\ c_{11} \cdots c_{lm_1} & & c_{j1} \cdots c_{jm_j} & & c_{m1} \cdots c_{mm_m} \end{array} \\ \begin{bmatrix} A_{\Phi C}^{11} & \cdots & A_{\Phi C}^{1j} & \cdots & A_{\Phi C}^{1m} \\ \vdots & & \vdots & & \vdots \\ A_{\Phi C}^{i1} & \cdots & A_{\Phi C}^{ij} & \cdots & A_{\Phi C}^{im} \\ \vdots & & \vdots & & \vdots \\ A_{\Phi C}^{m1} & \cdots & A_{\Phi C}^{mj} & \cdots & A_{\Phi C}^{mm} \end{bmatrix}_{n \times m | m < n, \sum_{j=1}^{m} m_j = n} \end{array} \tag{9}$$

where $\mathbf{A}_C^{\Phi}$ represents the normalized influence relation ratio for matrix $\mathbf{A}_C$ within the aspect $\mathbf{A}_D$, superscript Γ is the transpose, n represents the number of attributes, and m represents the number of aspects.

For this step, the normalization action is explained using the submatrix $\mathbf{A}_{\Phi C}^{11}$ in the matrix $\mathbf{A}_C^{\Phi}$ as the sample, as shown in equation (10). Other submatrices $\mathbf{A}_{\Phi C}^{mm}$ are then used to establish a complete unweighted supermatrix by the same actions.

$$A_{\Phi C}^{11} = \begin{array}{c} \\ c_{11} \\ \vdots \\ c_{1i} \\ \vdots \\ c_{lm_1} \end{array} \begin{array}{c} \begin{array}{cccccc} c_{11} & \cdots & c_{1j} & \cdots & c_{1m_1} \end{array} \\ \begin{bmatrix} a_{C^{11}}^{11} / c_1^{11} & \cdots & a_{C^{1j}}^{11} / c_1^{11} & \cdots & a_{C^{1m_1}}^{11} / c_1^{11} \\ \vdots & & \vdots & & \vdots \\ a_{C^{i1}}^{11} / c_i^{11} & \cdots & a_{C^{ij}}^{11} / c_i^{11} & \cdots & a_{C^{im_1}}^{11} / c_i^{11} \\ \vdots & & \vdots & & \vdots \\ a_{C^{m,1}}^{11} / c_{m_1}^{11} & \cdots & a_{C^{m,j}}^{11} / c_{m_1}^{11} & \cdots & a_{C^{m,m_1}}^{11} / c_{m_1}^{11} \end{bmatrix} \end{array} = \begin{array}{c} \\ c_{11} \\ \vdots \\ c_{1i} \\ \vdots \\ c_{lm_1} \end{array} \begin{array}{c} \begin{array}{cccccc} c_{11} & \cdots & c_{1o} & \cdots & c_{1m_1} \end{array} \\ \begin{bmatrix} a_{C^{11}}^{\Phi 11} & \cdots & a_{C^{1j}}^{\Phi 11} & \cdots & a_{C^{1m_1}}^{\Phi 11} \\ \vdots & & \vdots & & \vdots \\ a_{C^{i1}}^{\Phi 11} / c_i & \cdots & a_{C^{ij}}^{\Phi 11} & \cdots & a_{C^{im_1}}^{11} \\ \vdots & & \vdots & & \vdots \\ a_{C^{m,1}}^{\Phi 11} / c_{m_1}^{11} & \cdots & a_{C^{m,j}}^{\Phi 11} & \cdots & a_{C^{m,m_1}}^{11} \end{bmatrix}_{m_1 \times m_1} \end{array} \tag{10}$$

where $c_i^{11} = \sum_{j=1}^{m_1} a_{C^{ij}}^{11}, i = 1, 2, \ldots, m_1$.

Step 2: constructing a weighted supermatrix $\mathbf{U}$ using different influence ratio perspectives.

The unweighted supermatrix $\mathbf{A}_C^{\Phi*}$ only converts numerical values into the proportional form and does not consider the difference in influence proportion between aspects (i.e., the proportion of each aspect is equal). In this step, the total influence relation matrix $\mathbf{A}_D$ is the coefficient (i.e., transposed and normalized matrix $\mathbf{A}_D^{\Phi*}$) to adjust the matrix $\mathbf{A}_C^{\Phi*}$, so the final influence weight is closer to reality. In terms of ideas and methods. It also addresses the defects of the weighted super-matrix for the original ANP method. The equations are

$$\mathbf{A}_D^{\Phi*} = \left(\mathbf{A}_D^{\Phi}\right)^{\Gamma} = \begin{bmatrix} \frac{a_{11}^{D_{11}}}{d_1} & \cdots & \frac{a_{1o}^{D_{1j}}}{d_1} & \cdots & \frac{a_{1m}^{D_{1m}}}{d_1} \\ \vdots & & \vdots & & \vdots \\ \frac{a_{i1}^{D_{i1}}}{d_i} & \cdots & \frac{a_{ij}^{D_{ij}}}{d_i} & \cdots & \frac{a_{im}^{D_{im}}}{d_i} \\ \vdots & & \vdots & & \vdots \\ \frac{a_{m1}^{D_{m1}}}{d_m} & \cdots & \frac{a_{mj}^{D_{mj}}}{d_m} & \cdots & \frac{a_{mm}^{D_{mm}}}{d_m} \end{bmatrix} = \begin{bmatrix} a_{11}^{\Phi_{11}} & \cdots & a_{1j}^{\Phi_{1j}} & \cdots & a_{1m}^{\Phi_{1m}} \\ \vdots & & \vdots & & \vdots \\ a_{i1}^{\Phi_{i1}} & \cdots & a_{ij}^{\Phi_{ij}} & \cdots & a_{im}^{\Phi_{im}} \\ \vdots & & \vdots & & \vdots \\ a_{m1}^{\Phi_{m1}} & \cdots & a_{mj}^{\Phi_{mj}} & \cdots & a_{mm}^{\Phi_{mm}} \end{bmatrix}_{m \times m}, \tag{11}$$

where $d_i = \sum_{j=1}^{m} a_{ij}^{Di\ j}, i = 1, 2, \ldots, m.$

$$\mathbf{U} = \mathbf{A}_D^{\Phi *} \times \mathbf{A}_C^{\Phi *} = \begin{bmatrix} a_{11}^{\Phi_{11}} \times \mathbf{A}_{\Phi C}^{11} & \cdots & a_{i1}^{\Phi_{i1}} \times \mathbf{A}_{\Phi C}^{i1} & \cdots & a_{m1}^{\Phi_{m1}} \times \mathbf{A}_{\Phi C}^{m1} \\ \vdots & & \vdots & & \vdots \\ a_{1j}^{\Phi_{1j}} \times \mathbf{A}_{\Phi C}^{1j} & \cdots & a_{ij}^{\Phi_{ij}} \times \mathbf{A}_{\Phi C}^{ij} & \cdots & a_{mj}^{\Phi_{mj}} \times \mathbf{A}_{\Phi C}^{mj} \\ \vdots & & \vdots & & \vdots \\ a_{1m}^{\Phi_{1m}} \times \mathbf{A}_{\Phi C}^{1m} & \cdots & a_{im}^{\Phi_{im}} \times \mathbf{A}_{\Phi C}^{im} & \cdots & a_{mm}^{\Phi_{mm}} \times \mathbf{A}_{\Phi C}^{mm} \end{bmatrix}_{n \times n | m < n, \sum} \quad j = 1^m m_j = n. \tag{12}$$

Step 3: calculating the influential weights.

The weighted supermatrix $\mathbf{U}$ is multiplied by itself several times to become a steady-state supermatrix using the Markov chain concept and calculation. The limited supermatrix ϖ is then obtained, as shown in

$$\varpi = \lim_{\Omega \longrightarrow \infty} (\mathbf{U})^{\Omega}. \tag{13}$$

The third stage: the MOORA-AS method is used to calculate the performance ratio-gaps on competencies for newly graduated nurses.

The MOORA-AS method is used to calculate the total competency performance ratio-gap and the ratio-gap for each attribute for newly graduated nurses. The steps for this method are as follows:

Step 1: Establishing a performance matrix for newly graduated nurses.

The performance matrix $\mathbf{R}$ uses the evaluation framework (as shown in Table 1), whereby clinical nursing supervisors or managers use an 11-point scale (0 is the worst; 10 is the best) to evaluate the ability of newly graduated nurses, and the ability of each new nurse is calculated using the average method. The performance matrix $\mathbf{R}$ is defined in

$$\mathbf{R} = \left[r_{\text{eq}}\right]_{u \times n}, \quad e \in \{1, 2, \ldots, u\}, j \in \{1, 2, \ldots, n\}, \tag{14}$$

where u represents the number of newly graduated nurses and n represents the number of attributes.

Step 2: normalizing the performance matrix using the aspiration level.

In MADM decision-making problems, the scale for each attribute often lies within a known range, from the maximum value to the minimum value of the scale. When an aspiration level is used, the maximum value for the scale becomes the aspirated solution (i.e., r_q^{aspired}). However, the minimum value is the worst solution (i.e., r_q^{worst}). The normalized matrix is shown in

$$\mathbf{R}^{\Phi} = \left[r_{\text{eq}}^{\Phi}\right]_{e \times q} = \left[\frac{\left|r_q^{\text{aspired}} - r_{\text{eq}}\right|}{\left|r_q^{\text{aspired}} - r_q^{\text{worst}}\right|}\right]_{e \times q}, \tag{15}$$

where r_{eq}^{Φ} represents the performance gap ratio between 0 and 1.

Step 3: building the score for the ratio-gap system for each newly graduated nurse.

The ratio-gap system for newly graduated nurses reflects the overall gap between current performance and aspirational level. The lower the value, the higher the ranking for the new nurse in terms of all new nurse candidates. The calculation is shown as

$$T_e = \sum_{q=1}^{n} \varpi_q \times r_{\text{eq}}^{\Phi}, \quad \forall e, \tag{16}$$

where ϖ_q represents the influential weights, as calculated using the DANP method.

Step 4 : Calculating the score for the reference ratio-gap point for each newly graduated nurse.

The reference point for newly graduated nurses is the worst attribute (the maximum ratio-gap), so this is the maximum ratio-gap for ability, as shown in

$$F_e = \max_q \left\{\varpi_q \times r_{eq}^{\Phi} | q = 1, 2, \ldots, n\right\}, \quad \forall e. \tag{17}$$

4. Empirical Case Study

A third-class and first-class official hospital in China are used to illustrate the use of the hybrid proposed MADM model to evaluate the ability and determine the improvements in newly graduated nurses.

4.1. The Background Description and Data Collection. The case involves a third-class and a first-class hospital that offer medical treatment, scientific research, teaching, and prevention. In terms of medical disciplines, the case hospital has 14 provincial specialized disease centers in the region (there are 20 centers in the region) and 23 key medical disciplines (orthopedics, digestive medicine, and cardiovascular medicine). There are 1,924 beds in the case hospital and an

Table 3: Initial influence relation matrix **H**.

Attribute	C_{11}	C_{12}	C_{13}	C_{21}	C_{22}	C_{23}	C_{31}	C_{32}	C_{33}
C_{11}	0.000	3.000	2.400	2.667	3.067	2.867	2.400	2.267	2.133
C_{12}	3.133	0.000	2.533	2.067	3.467	3.067	2.733	2.400	2.667
C_{13}	2.133	2.200	0.000	2.600	2.267	2.733	1.867	2.133	2.133
C_{21}	2.533	2.200	2.333	0.000	2.000	3.000	2.200	2.600	2.467
C_{22}	3.067	2.667	1.800	2.000	0.000	2.800	2.200	2.400	2.400
C_{23}	2.800	2.733	2.733	2.600	2.800	0.000	1.933	2.200	2.400
C_{31}	2.400	2.200	2.133	2.267	2.333	2.600	0.000	2.533	2.667
C_{32}	2.133	2.200	2.200	2.533	2.133	2.467	2.400	0.000	2.267
C_{33}	2.467	2.467	2.267	2.067	2.200	2.533	2.267	2.333	0.000

Note. Values are calculated from equation (1).

Table 4: The variance of the 15-fold cross-validation.

Number of experts	No. 1	No. 2	No. 3	No. 4	No. 5	No. 6	No. 7	No. 8
Gap	0.016	0.011	0.014	0.014	0.019	0.037	0.023	0.018
Gap (%)	1.6%	1.1%	1.4%	1.4%	1.9%	3.7%	2.3%	1.8%
Number of experts	No. 9	No. 10	No. 11	No. 12	No. 13	No. 14	No. 15	Average
Gap	0.043	0.014	0.017	0.021	0.015	0.016	0.013	0.019
Gap (%)	4.3%	1.4%	1.7%	2.1%	1.5	1.6%	1.3%	1.9%

Note. The significant confidence equation is $(1/9(9-1))\sum_{i=1}^{15}\sum_{j=1}^{15}(|h_{ij}^{15} - h_{ij}^{15-1}|/h_{ij}^{15}) \times 100\% = 1.9\% < 5\%$; i.e., significant confidence is 98.1%.

average of 110,000 hospitalizations each year. In the 2019 self-assessed statistical report, the number of outpatients was over 2.7 million, and the number of discharged patients was about 108,000. The number of medical disciplines and the population in the region mean that the case hospital is one of the major hospitals in the region.

Every year, the hospital employs approximately 150 new nurses. Every month, the hospital gives a series of standardized training courses to track the training and assessment results for new nurses. There is also a series of assessment standards for clinical practice. These measures allow new nurses to integrate into clinical nursing work as soon as possible and provide better nursing services for hospitalized patients. The evaluation of competency includes quantitative indicators and non-quantitative indicators; therefore, in addition to the existing clinical quantitative index evaluation system, the decision-making model for competency allows nursing departments to systematically evaluate the ability of new nurses and to propose improvement strategies, which is the most important management problem in hospitals at present.

This study proposes a hybrid MADM model that uses the newly graduated nurse competency framework (as shown in Table 1) from a past study and MADM methods. For this study, data was collected from 15 clinical nursing experts (2 nursing administrators, 6 head nurses, and 7 nursing instructors). In terms of clinical nursing experience, one expert had more than 2 to 3 years, 1 expert had more than 4 to 5 years, 3 experts had more than 5 to 10 years, and 10 experts had more than 10 years. These head nurses or nurse instructors are nurses with much professional and clinical experience in nursing. They are also aware of the relationship between the abilities that newly graduated nurses should possess and these abilities. Three newly graduated nurses with less than 3 years of employment are the evaluation objects.

4.2. Calculating the Interdependence Relationship between Attributes Using the DEMATEL Method. Based on the practical experience of 15 clinical nursing experts, an initialization influence relationship matrix was established, as shown in Table 3. The quality of this matrix is very important because it affects the results for the structure of a causal relationship and the weight, as well as subsequent evaluation and improvement strategies that are proposed as a consequence. Using the concept of a confidence level and a 15-fold cross-validation process, the average variance consistency for the initialization matrix is between 0.043 (i.e., 4.3%) and 0.011 (i.e., 1.1%), with an average of 0.019 (i.e., 1.9%, less than 5%). The significant confidence level is 98.1% (i.e., greater than 95%), so the experience of this group of clinical nursing experts is consistent, and the subsequent results are very stable, as shown in Table 4.

The initialization matrix **H** describes the total influence relationship degree between attributes using equations (2) to (4), which is the total influence relationship matrix **A**, as shown in Table 5.

Finally, the total influence relationship matrix **A** uses (5) to (8) to calculate the total "given influence (p_i)" and "received influence (y_i)" for each attribute and the "prominence ($p_i + y_i$)" and "relation ($p_i - y_i$),"as shown in Table 6.

A causal influence-network diagram (CIND) allows nursing department decision-makers/managers to understand the interaction between all aspects and attributes for newly graduated nurses by identifying the influence structure between "prominence ($p_i + y_i$)" and "relation ($p_i - y_i$),"as shown in Figure 2.

Table 5: Total influence relation matrix **A**.

Attribute	C_{11}	C_{12}	C_{13}	C_{21}	C_{22}	C_{23}	C_{31}	C_{32}	C_{33}
C_{11}	0.877	0.959	0.886	0.910	0.986	1.042	0.868	0.896	0.902
C_{12}	1.048	0.884	0.932	0.930	1.045	1.097	0.921	0.943	0.964
C_{13}	0.858	0.827	0.692	0.810	0.850	0.923	0.754	0.793	0.802
C_{21}	0.920	0.873	0.831	0.749	0.887	0.983	0.809	0.854	0.859
C_{22}	0.945	0.895	0.815	0.836	0.810	0.981	0.813	0.851	0.861
C_{23}	0.965	0.926	0.876	0.885	0.952	0.901	0.830	0.871	0.889
C_{31}	0.907	0.865	0.816	0.834	0.891	0.960	0.711	0.844	0.859
C_{32}	0.866	0.835	0.791	0.816	0.854	0.923	0.782	0.713	0.816
C_{33}	0.891	0.857	0.804	0.809	0.869	0.938	0.788	0.820	0.734

Note. Values are calculated from equations (1) to (4).

Table 6: The influence indicators for aspects and attributes.

Aspect	p_i	y_i	$p_i + y_i$	$p_i - y_i$	Attribute	p_i	y_i	$p_i + y_i$	$p_i - y_i$
					C_{11}	8.326	8.277	16.603	0.049
C_1	2.711	2.627	5.338	0.084	C_{12}	8.764	7.923	16.687	0.842
					C_{13}	7.309	7.442	14.751	−0.134
					C_{21}	7.764	7.578	15.342	0.186
C_2	2.630	2.719	5.349	−0.089	C_{22}	7.807	8.144	15.951	−0.337
					C_{23}	8.096	8.749	16.845	−0.653
					C_{31}	7.687	7.275	14.962	0.411
C_3	2.510	2.505	5.016	0.005	C_{32}	7.397	7.585	14.982	−0.188
					C_{33}	7.511	7.687	15.198	−0.176

Note. Values are calculated from equations (5) to (8).

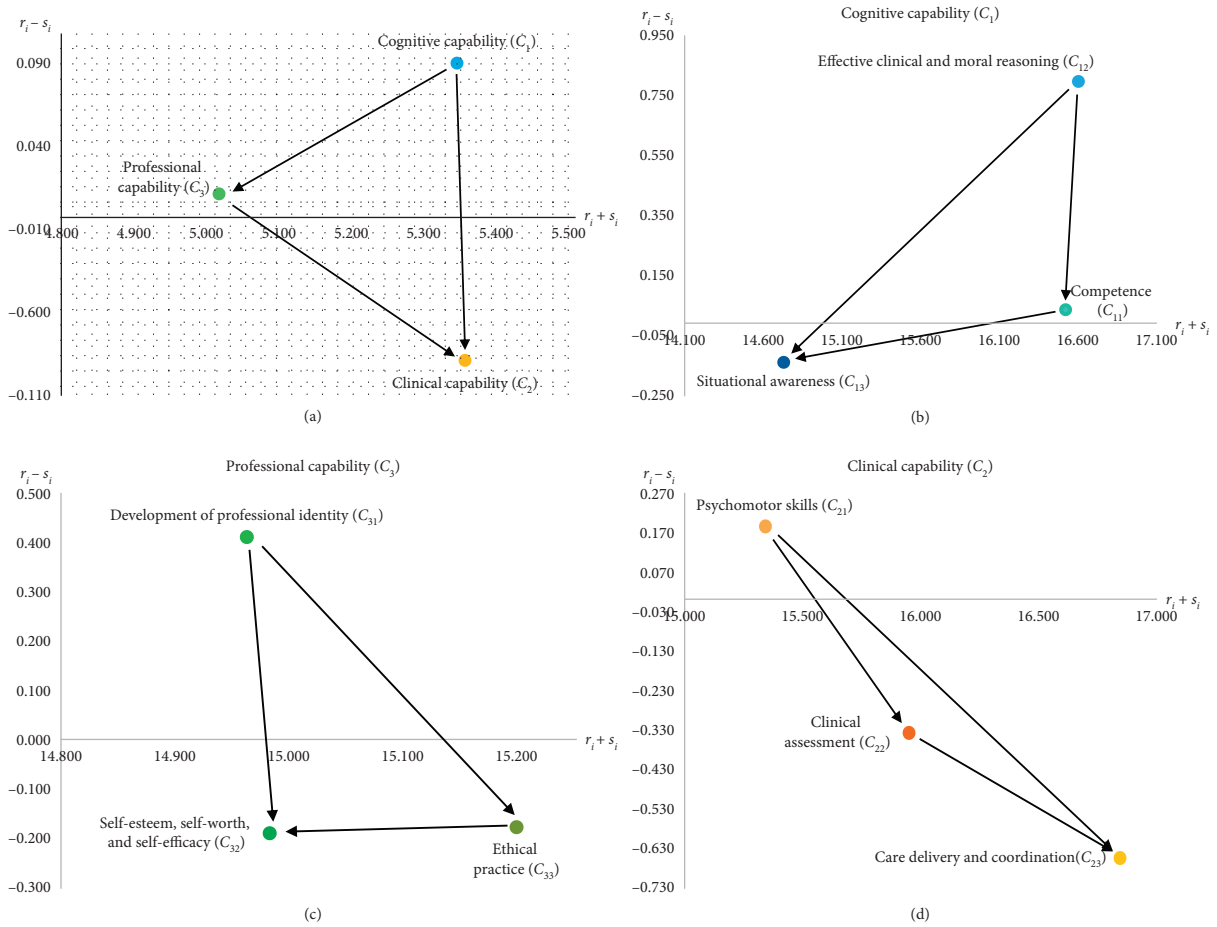


Figure 2: (a) Newly graduated nurse competency framework. (b) Cognitive capability (*C*1). (c) Professional capability (*C*3). (d) Clinical capability (*C*2).

Table 7: Influential weights of attributes.

Aspect	Local weight	Ranking	Attribute	Local weight	Ranking	Global weight	Ranking
C_1	0.335	2	C_{11}	0.350	1	0.117	2
			C_{12}	0.335	2	0.112	4
			C_{13}	0.315	3	0.105	8
C_2	0.346	1	C_{21}	0.310	3	0.107	7
			C_{22}	0.333	2	0.115	3
			C_{23}	0.357	1	0.124	1
C_3	0.319	3	C_{31}	0.323	3	0.103	9
			C_{32}	0.336	2	0.107	6
			C_{33}	0.341	1	0.109	5

Note. Values are calculated using equations (9) to (13).

In Figure 2, in terms of aspect level, "cognitive capability (C_1)" is the basis for the other two abilities ("professional capability (C_3)" and "clinical capability (C_2)"). Only when a newly graduated nurse has a clear understanding and judgment of situations and surrounding things in the clinical environment, are the advantages, disadvantages, and limitations obvious, and only then can professional values, beliefs, and goals be defined. This develops a strong internal driving force to continuously work to achieve goals. Under the influence of this internal drive, nurses study hard and improve their self-competence, professional ability, and clinical ability. "Effective clinical and moral reasoning (C_{12})," "psychomotor skills (C_{21})," and "development of professional identity (C_{31})" are the most critical attributes at the corresponding levels.

4.3. Calculating the Weights of Influence Relationship for Attributes Using DANP. The total influence relation matrix**A**describes the influence weight for each attribute using (9) to (13), which also shows the importance of each attribute to the overall ability of newly graduated nurses, as shown in Table 7. In terms of weight, an improvement in "clinical capability (C_2)" is the premise for newly graduated nurses to be better qualified for clinical nursing work. If there is a rapid improvement in clinical ability, newly graduated nurses can continuously improve their ability to observe and discover problems and their prospective analysis ability, so newly graduated nurses can learn to make correct clinical judgments and decisions and address existing or potential clinical problems. In the process of clinical nursing practice, clinical ability is the most important of these three competencies at aspect level.

4.4. Evaluating and Ranking Newly Graduated Nurses Using the MOORA-AS Method. Three newly graduated nurses were assessed by clinical nursing experts using the newly graduated nurse's ability assessment framework (as shown in Table 1) and an 11-point scale (the worst is 0 points; the best is 10 points) to derive an original performance score using the average method. The original performance score uses equations (14) to (17) for the MOORA-AS method to calculate the performance ratio-gaps for the overall ability of the three newly graduated nurses. The results are shown in Table 8. These results show that "newly graduated nurse *B*" has the smallest ratio-gap with the aspirational level, so this is the best of the three newly graduated nurses.

5. Discussion

Based on these results, the improvement strategies and the causal relationship are further explained and the original MOORA and MOORA-AS methods are compared.

5.1. Systemic Improvement Strategy Using the CIND. Improvement strategies for newly graduated nurses are formulated using the results for the reference ratio-gap ratio (Table 8), the CIND (Figure 2), and the weight (Table 7). The results for newly graduated nurse *A* are shown in Figure 3. "Cognitive capability (C_1)" is the worst attribute (i.e., ratio-gap = 0. 082) at the aspect level, and it is also the most main influence source for the entire capability framework (relation is ranked first). Further analysis at the attribute level shows that "effective clinical and moral reasoning (C_{12})" has the maximum ratio-gap (i.e., 0.084). It is also the main influence source.

Using these results, the nursing department decision-makers/managers might propose appropriate training courses or alternatives to improve "effective clinical and moral reasoning (C_{12})." When "effective clinical and moral reasoning (C_{12})" is improved, the overall ability of newly graduated nurse *A* will improve. This study shows that the quality of standardized training for nurses after graduation is important. Standardized training for nurses uses the basic stages of the growth for new graduated nurses and remains crucial for a nurse's professional development. It is a process that promotes the formation of professional ethics, cultivation of clinical thinking mode, and standardization and improvement of nursing practice for new nurses.

To improve the performance of newly graduated nurse *A* in terms of the C_{12} criteria, a systematic symptom care course for each symptom is necessary: overview; nursing evaluation; nursing measures and case analysis. Newly graduated nurse *A* must attain basic theoretical knowledge and skills for psychological care, health education, and other humanistic care knowledge. Newly graduated nurse *A* must understand the internal relationship between symptoms and

Table 8: Results for the evaluation of newly graduated nurses using the MOORA-AS method.

	Original performance				MOORA-AS		
	Newly graduated nurse A	Newly graduated nurse B	Newly graduated nurse C	Local weight	Newly graduated nurse A	Newly graduated nurse B	Newly graduated nurse C
C_1				0.335	0.082	0.074	0.164
C_{11}	7.667	8.000	5.500	0.350	0.082	0.070	0.157
C_{12}	7.500	7.600	5.250	0.335	0.084	0.081	0.159
C_{13}	7.500	7.800	4.500	0.315	0.079	0.069	0.173
C_2				0.346	0.076	0.067	0.139
C_{21}	8.000	8.600	7.250	0.310	0.062	0.043	0.085
C_{22}	6.833	7.400	4.000	0.333	0.105	0.086	0.200
C_{23}	8.500	8.200	6.750	0.357	0.054	0.064	0.116
C_3				0.319	0.062	0.072	0.104
C_{31}	8.000	6.400	6.500	0.323	0.065	0.116	0.113
C_{32}	7.667	7.800	6.500	0.336	0.078	0.074	0.118
C_{33}	8.500	9.000	7.250	0.341	0.051	0.034	0.094
The ratio-gap system					0.220 (2)	0.212 (1)	0.406 (3)
The reference ratio-gap point					0.082 (2)	0.074 (1)	0.164 (3)
Final rank					2	1	3

Note. The values of the MOORA-AS method are weighted and are calculated using equations (14) to (17).

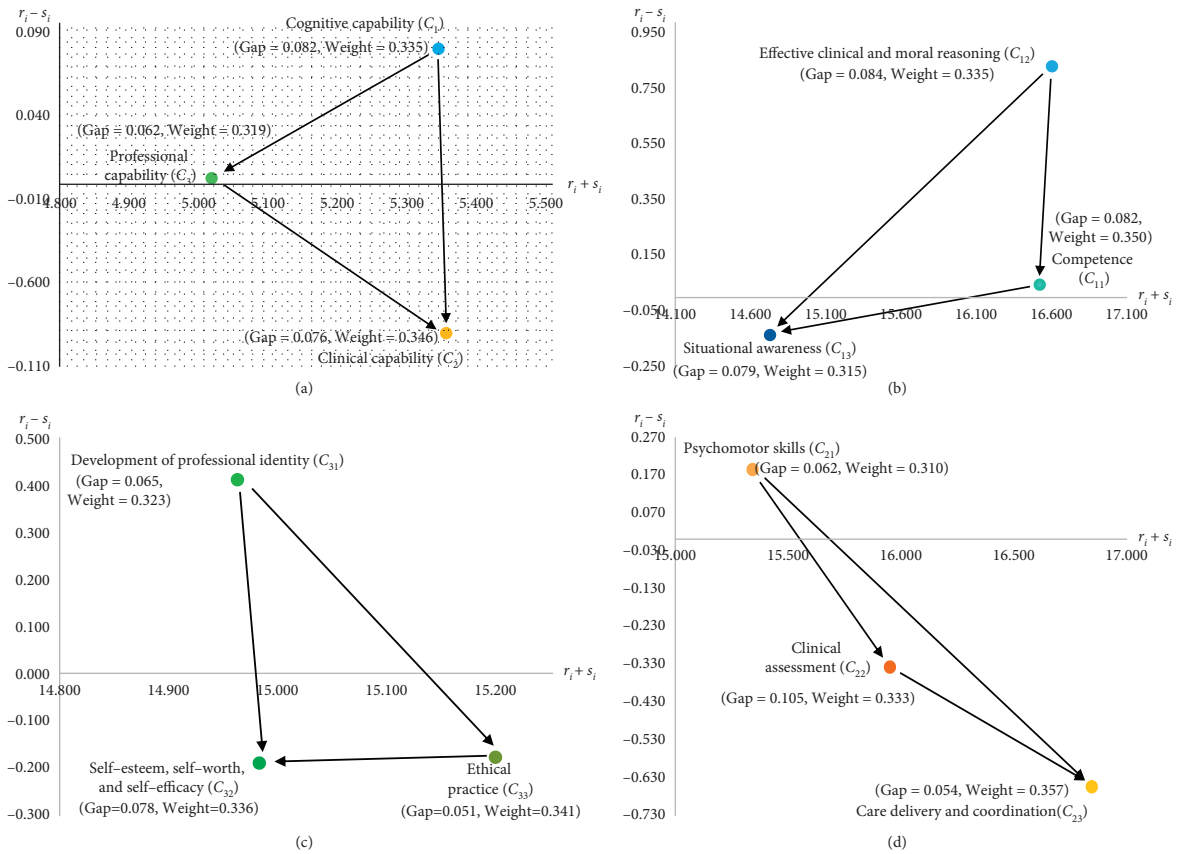


Figure 3: (a) Newly graduated nurse competency framework. (b) Cognitive capability (*C*1). (c) Professional capability (*C*3). (d) Clinical capability (*C*2).

Table 9: Results for the original MOORA and the MOORA-AS methods.

Aspect/ attribute	Local weight	Original MOORA			MOORA-AS		
		Newly graduated nurse A	Newly graduated nurse B	Newly graduated nurse C	Newly graduated nurse A	Newly graduated nurse B	Newly graduated nurse C
C_1	0.335	0.030	0.000	0.335	0.082	0.074	0.164
C_{11}	0.350	0.047	0.000	0.350	0.082	0.070	0.157
C_{12}	0.335	0.014	0.000	0.335	0.084	0.081	0.159
C_{13}	0.315	0.029	0.000	0.315	0.079	0.069	0.173
C_2	0.346	0.067	0.021	0.346	0.076	0.067	0.139
C_{21}	0.310	0.138	0.000	0.310	0.062	0.043	0.085
C_{22}	0.333	0.055	0.000	0.333	0.105	0.086	0.200
C_{23}	0.357	0.000	0.061	0.357	0.054	0.064	0.116
C_3	0.319	0.042	0.103	0.313	0.062	0.072	0.104
C_{31}	0.323	0.000	0.323	0.303	0.065	0.116	0.113
C_{32}	0.336	0.034	0.000	0.336	0.078	0.074	0.118
C_{33}	0.341	0.097	0.000	0.341	0.051	0.034	0.094
The ratio-gap system		0.139 (2)	0.124 (1)	0.994 (3)	0.220 (2)	0.212 (1)	0.406 (3)
The reference ratio-gap point		0.06 (1)	0.103 (2)	0.346 (3)	0.082 (2)	0.074 (1)	0.164 (3)

Note. The values for the original MOORA and MOORA-AS methods are weighted.

disease for symptom analysis, make correct clinical judgments, and better guide clinical practice.

5.2. Difference between Original MOORA and MOORA-AS Methods. Table 9 shows the difference between different methods for the same performance data. The ranking for the reference ratio-gap point is different. The ranking for the original MOORA method is newly graduated nurse $A \succ$ newly graduated nurse $B \succ$ newly graduated nurse C, but for the MOORA-AS method, the ranking is newly graduated nurse $B \succ$ newly graduated nurse $A \succ$ newly graduated nurse C. Using the original MOORA method, the attribute ratio-gap has a large number of zero scores or perfect scores for newly graduated nurse B and newly graduated nurse C, who are the best and worst cases, respectively. Many ratio-gaps = 0 points (i.e., C_1, C_{11}, C_{12}, C_{13}, C_{21}, C_{22}, C_{32}, C_{33}) for newly graduated nurse B, and many proportion-gaps = 1 point (i.e., C_1, C_{11}, C_{12}, C_{13}, C_2, C_{21}, C_{22}, C_{23}, C_{32}, C_{33}) for newly graduated nurse C. This is because the original MOORA method uses relative concepts for ranking, but MOORA-AS uses absolute concepts (i.e., aspiration level) for ranking. The MOORA-AS method gives a more robust ranking for nursing department decision-makers/managers, especially if the measurement scale is known.

6. Conclusions

Professional nurses play a key role in the fight against COVID-19 infection. The shortage of professional nurses has been a problem for the medical system and hospitals because of the high turnover rate for newly graduated nurses. This study proposes a decision-making model to allow nursing department decision-makers/managers to improve the competence of newly graduated nurses and to decrease the ability gaps between newly qualified and professional nurses. A hybrid MADM model is proposed, which uses an evaluation framework that is based on findings for the competency characteristics of new graduate nurses from previous studies, combined with the MADM method, which is based on the system perspective. An empirical study of the hybrid MADM model uses data from a third-class and a first-class hospital in China, including the modeling process, result analysis, and improvement strategy. The highlights of each stage are as follows:

(1) The competency framework for this study describes the competency that is required for a newly graduated nurse before entering a new nursing position and the ability to cope with new and unfamiliar nursing situations by evaluating and improving existing competencies. This competency framework is based on the results of previous studies of competency problems involving recent graduates. The problem is a MADM problem and is the subject of few studies.

(2) The results for the CIND show that cognitive capability (C_1) is a critical factor for the other two abilities: professional capability (C_3) and clinical capability (C_2). This is because the cognitive capability is improved. The professional capability and clinical capability are also improved to a certain extent. Effective clinical and moral reasoning (C_{12}), psychomotor skills (C_{21}), and development of professional identity (C_{31}) are the most critical attributes at the corresponding levels.

(3) The results for influential weight show that an improvement in clinical capability (C_2) is related to newly graduated nurses being better qualified for clinical nursing work, so newly graduated nurses make correct clinical judgments and decisions and finally address the current and potential clinical problems of patients correctly. For the process of

clinical nursing practice, the weight for clinical ability is the most important of these three competencies in terms of aspect level.

(4) In terms of a systematic improvement strategy for newly graduated nurse A, nursing department decision-makers/managers could propose appropriate training courses or programs to improve effective clinical and moral reasoning (C_{12}) because this attribute has the maximum ratio-gap and is also a critical factor. If an improvement is achieved, other abilities will also be improved.

(5) The results for the original MOORA and MOORA-AS methods show that the MOORA-AS method addresses the problem of ranking and is better than the original MOORA method for MADM problems because an absolute concept (i.e., aspiration level) gives more stable results than a relative concept.

This research model allows hospitals to evaluate the ability of newly graduated nurses and to propose improvement strategies from a systematic viewpoint. The nursing department can combine practical training courses to create bespoke training courses for each newly graduated nurse. The results of this study are limited to the experience for a single hospital because different hospitals may have different weights and performance scores for newly graduated nurses. This study does not consider the uncertainty (i.e., fuzzy phenomenon) between competencies in the clinical nursing environment. Future research models might use various other theories and methods (such as intuitionistic fuzzy theory [65], 2-tuple linguistic model [66], and a gray system [67]) to establish decision-making models that have more technical and practical value for newly graduated nurses.

Data Availability

The data used to support the findings of this study are included in the article.

Conflicts of Interest

The authors declare no conflicts of interest.

Authors' Contributions

Fengmin Cheng and Yanjun Jin contributed equally to this study.

Acknowledgments

The authors are very grateful to the nursing experts in the case hospital for assisting with the questionnaire survey. This research was supported by the Graduate Education Innovation Project (GEIP) of Guangdong Province (2019JGXM36) and the Specialized Subsidy Scheme for Macao Higher Education Institutions in the Area of Research in Humanities and Social Sciences (and Specialized Subsidy Scheme for Prevention and Response to Major Infectious Diseases) (HSSMUST-2020-9).

References

[1] T. Peiró, L. Lorente, and M. Vera, "The COVID-19 crisis: skills that are paramount to build into nursing programs for future global health crisis," *International Journal of Environmental Research and Public Health*, vol. 17, no. 18, p. 6532, 2020.

[2] C. Paterson, B. Gobel, T. Gosselin et al., "Oncology nursing during a pandemic: critical reflections in the context of COVID-19," *Seminars in Oncology Nursing*, vol. 36, no. 3, Article ID 151028, 2020.

[3] R. E. Jordan, P. Adab, and K. K. Cheng, "Covid-19: risk factors for severe disease and death," *BMJ*, vol. 368, p. m1198, 2020.

[4] E. Halcomb, S. McInnes, A. Williams et al., "The experiences of primary healthcare nurses during the COVID-19 pandemic in Australia," *Journal of Nursing Scholarship*, vol. 52, no. 5, pp. 553–563, 2020.

[5] World Health Organization (WHO), *Year of the Nurse and the Midwife*, WHO, Geneva, Switzerland, 2020, https://www.who.int/campaigns/year-of-the-nurse-and-the-midwife-2020.

[6] M. Marc, A. Bartosiewicz, J. Burzynska, Z. Chmiel, and P. Januszewicz, "A nursing shortage - a prospect of global and local policies," *International Nursing Review*, vol. 66, pp. 9–16, 2019.

[7] Y. Song and L. L. McCreary, "New graduate nurses' self-assessed competencies: an integrative review," *Nurse Education in Practice*, vol. 45, Article ID 102801, 2020.

[8] P. Griffiths, J. Ball, J. Drennan et al., "Nurse staffing and patient outcomes: strengths and limitations of the evidence to inform policy and practice. A review and discussion paper based on evidence reviewed for the national institute for health and care excellence safe staffing guideline development," *International Journal of Nursing Studies*, vol. 63, pp. 213–225, 2016.

[9] N. N. Solutions, *2020 NSI National Health Care Retention & RN Staffing Report*, NSI Nursing Solutions Inc, East Petersburg, PA, USA, 2020.

[10] D. Nei, L. A. Snyder, and B. J. Litwiller, "Promoting retention of nurses," *Health Care Management Review*, vol. 40, no. 3, pp. 237–253, 2015.

[11] C.-Y. Cheng, H.-M. Tsai, C.-H. Chang, and S.-R. Liou, "New graduate nurses' clinical competence, clinical stress, and intention to leave: a longitudinal study in Taiwan," *The Scientific World Journal*, vol. 2014, p. 1, 2014.

[12] P. C. Beecroft, F. Dorey, and M. Wenten, "Turnover intention in new graduate nurses: a multivariate analysis," *Journal of Advanced Nursing*, vol. 62, no. 1, pp. 41–52, 2008.

[13] L.-H. Odland, T. Sneltvedt, and V. Sörlie, "Responsible but unprepared: experiences of newly educated nurses in hospital care," *Nurse Education in Practice*, vol. 14, no. 5, pp. 538–543, 2014.

[14] M. S. Hezaveh, F. Rafii, and N. Seyedfatemi, "Novice nurses' experiences of unpreparedness at the beginning of the work," *Global Journal of Health Science*, vol. 6, no. 1, pp. 215–222, 2013.

[15] T. Levett-Jones and M. FitzGerald, "A review of graduate nurse transition programs in Australia," *The Australian Journal of Advanced Nursing*, vol. 23, pp. 40–45, 2005.

[16] J. E. B. Duchscher, "Transition shock: the initial stage of role adaptation for newly graduated registered nurses," *Journal of Advanced Nursing*, vol. 65, no. 5, pp. 1103–1113, 2009.

[17] A. Baumann, M. Hunsberger, and M. Crea-Arsenio, "Workforce integration of new graduate nurses: evaluation of a health human resources employment policy," *Healthcare Policy | Politiques de Santé*, vol. 7, no. 2, pp. 47–59, 2011.

[18] A. C. Wolff, S. Regan, B. Pesut, and J. Black, "Ready for what? An exploration of the meaning of new graduate nurses' readiness for practice," *International Journal of Nursing Education Scholarship*, vol. 7, no. 1, 2010.

[19] R. Watson, A. Stimpson, A. Topping, and D. Porock, "Clinical competence assessment in nursing: a systematic review of the literature," *Journal of Advanced Nursing*, vol. 39, no. 5, pp. 421–431, 2002.

[20] Q. Chen, S.-B. Tsai, Y. Zhai et al., "An empirical study on low-carbon: human resources performance evaluation," *International Journal of Environmental Research and Public Health*, vol. 15, no. 1, p. 62, 2018.

[21] M. C. S. de Moura and M. F. F. Sobral, "Evaluation skills with support of multicriteria modeling," *International Business Management*, vol. 10, pp. 1–8, 2016.

[22] E. Aksakal and M. Dağdeviren, "An integrated approach for personel selection with DEMATEL and ANP methods," *Journal of the Faculty of Engineering and Architecture of Gazi University*, vol. 25, pp. 905–913, 2010.

[23] A. Ishizaka and V. E. Pereira, "Portraying an employee performance management system based on multi-criteria decision analysis and visual techniques," *International Journal of Manpower*, vol. 37, no. 4, pp. 628–659, 2016.

[24] B. Kucukaltan, Z. Irani, and E. Aktas, "A decision support model for identification and prioritization of key performance indicators in the logistics industry," *Computers in Human Behavior*, vol. 65, pp. 346–358, 2016.

[25] Y.-C. Chuang, S.-K. Hu, J. J. H. Liou, and G.-H. Tzeng, "A data-driven MADM model for personnel selection and improvement," *Technological and Economic Development of Economy*, vol. 26, no. 4, pp. 751–784, 2020.

[26] S.-S. Weng, Y. Liu, and Y.-C. Chuang, "Reform of Chinese universities in the context of sustainable development: teacher evaluation and improvement based on hybrid multiple criteria decision-making model," *Sustainability*, vol. 11, no. 19, p. 5471, 2019.

[27] Y. Liu, Y. Chen, and G.-H. Tzeng, "Identification of key factors in consumers' adoption behavior of intelligent medical terminals based on a hybrid modified MADM model for product improvement," *International Journal of Medical Informatics*, vol. 105, pp. 68–82, 2017.

[28] K. Casey, R. Fink, C. Jaynes, L. Campbell, P. Cook, and V. Wilson, "Readiness for practice: the senior practicum experience," *Journal of Nursing Education*, vol. 50, no. 11, pp. 646–652, 2011.

[29] C. Bowdoin, *Thinking like a Nurse and Perceived Readiness for Professional Practice: A Mixed Methods Study*, Mercer University, Macon, Georgia, 2014.

[30] N. Mirza, L. Manankil-Rankin, D. Prentice, L.-A. Hagerman, and C. Draenos, "Practice readiness of new nursing graduates: a concept analysis," *Nurse Education in Practice*, vol. 37, pp. 68–74, 2019.

[31] A. Bandura, *Self-efficacy in Changing Societies*, Cambridge University Press, Cambridge, UK, 1995.

[32] G. E. Miller, "The assessment of clinical skills/competence/performance," *Academic Medicine*, vol. 65, no. 9, pp. S63–S67, 1990.

[33] N. A. Mirza, "Effects of abductive reasoning training on hypothesis generation abilities of first and second year baccalaureate nursing students," Master thesis, McMaster University, Hamilton, Canada, 2015.

[34] M. Murray, D. Sundin, and V. Cope, "New graduate registered nurses' knowledge of patient safety and practice: a literature review," *Journal of Clinical Nursing*, vol. 27, no. 1-2, pp. 31–47, 2018.

[35] V. M. Evenson, "The consistency of clinical nursing education with employer expectations: perceptions of registered nurses in Montana," Doctoral dissertation, University of Montana, Missoula, MT, USA, 1989.

[36] K. Missen, L. McKenna, and A. Beauchamp, "Registered nurses' perceptions of new nursing graduates' clinical competence: a systematic integrative review," *Nursing & Health Sciences*, vol. 18, no. 2, pp. 143–153, 2016.

[37] A. Pillai, "Ready, set, practice: a study of generational differences and their impact on senior baccalaureate nursing student perceived readiness to practice," Doctoral Dissertation. Retrieved from ProQuest Dissertations and Theses, 2014.

[38] A. P. Wright, "Preparedness to practice as perceived by the graduate nurse," 2014.

[39] S. Wangensteen, I. S. Johansson, M. E. Björkström, and G. Nordström, "Newly graduated nurses' perception of competence and possible predictors: a cross-sectional survey," *Journal of Professional Nursing*, vol. 28, no. 3, pp. 170–181, 2012.

[40] K.-l. Edward, K. Ousey, J. Playle, and J.-A. Giandinoto, "Are new nurses work ready - the impact of preceptorship. An integrative systematic review," *Journal of Professional Nursing*, vol. 33, no. 5, pp. 326–333, 2017.

[41] J. Reagor, *Perceived Readiness for Practice of Senior Baccalaureate Nursing Students*, University of Kansas, Lawrence, KS, USA, 2010.

[42] M. F. El-Santawy and A. N. Ahmed, "Personnel training selection problem based on SDV-MOORA," *Life Sciences Journal*, vol. 9, pp. 152–154, 2012.

[43] K. Skrzypek and K. Dąbrowski, "Selection of employees in the metal industry based on competences, on the example of a designer position," *Foundations of Management*, vol. 7, no. 1, pp. 217–224, 2015.

[44] D. Karabasevic, E. K. Zavadskas, Z. Turskis, and D. Stanujkic, "The framework for the selection of personnel based on the SWARA and ARAS methods under uncertainties," *Informatica*, vol. 27, no. 1, pp. 49–65, 2016.

[45] S. S. Gilan, M. H. Sebt, and V. Shahhosseini, "Computing with words for hierarchical competency based selection of personnel in construction companies," *Applied Soft Computing*, vol. 12, pp. 860–871, 2012.

[46] X.-H. Wu and J.-Q. Wang, "Cross-entropy measures of multivalued neutrosophic sets and its application in selecting middle-level manager," *International Journal for Uncertainty Quantification*, vol. 7, no. 2, pp. 155–176, 2017.

[47] L. Canós, T. Casasús, V. Liern, and J. C. Pérez, "Soft computing methods for personnel selection based on the valuation of competences," *International Journal of Intelligent Systems*, vol. 29, no. 12, pp. 1079–1099, 2014.

[48] D. A. Patel and K. N. Jha, "Evaluation of construction projects based on the safe work behavior of co-employees through a neural network model," *Safety Science*, vol. 89, pp. 240–248, 2016.

[49] I. Ahmed, I. Sultana, S. K. Paul, and A. Azeem, "Employee performance evaluation: a fuzzy approach," *International Journal of Productivity and Performance Management*, vol. 62, no. 7, pp. 718–734, 2013.

[50] A. S. Hanna, M. W. Ibrahim, W. Lotfallah, K. A. Iskandar, and J. S. Russell, "Modeling project manager competency: an integrated mathematical approach," *Journal of Construction Engineering and Management*, vol. 142, 2016.

[51] M. Bohlouli, N. Mittas, G. Kakarontzas, T. Theodosiou, L. Angelis, and M. Fathi, "Competence assessment as an expert system for human resource management: a mathematical approach," *Expert Systems with Applications*, vol. 70, pp. 83–102, 2017.

[52] A. Gabus and E. Fontela, *World Problems, an Invitation to Further Thought within the Framework of DEMATEL*, Battelle Institute, Geneva Research Center, Geneva, Switzerland, 1972.

[53] E. Fontela and A. Gabus, *The DEMATEL Observer*, Battelle Geneva Research Center, Geneva, Switzerland, 1976.

[54] X. Zhou, Y. Shi, X. Deng, and Y. Deng, "D-DEMATEL: a new method to identify critical success factors in emergency management," *Safety Science*, vol. 91, pp. 93–104, 2017.

[55] W. S. Lee, G. H. Tzeng, and C. H. Cheng, "Using novel MCDM methods based on Fama-French three-factor model for probing the stock selection," in *Proceedings of the 10th Asia-Pacific Industrial Engineering and Management Systems Conference*, pp. 1460–1474, Kitakyushu, Japan, October 2019.

[56] T. L. Saaty, *Decision Making with Dependence and Feedback: The Analytic Network Process: The Organization and Prioritization of Complexity*, RWS Publications, Pittsburgh, PA, USA, 2001.

[57] M.-H. Yang, C.-H. Su, and W.-C. Wang, "Use of hybrid MCDM model in evaluation for cloud service application improvement," *EURASIP Journal on Wireless Communications and Networking*, vol. 2018, no. 1, p. 98, 2018.

[58] M. F. Bongo, K. M. S. Alimpangog, J. F. Loar, J. A. Montefalcon, and L. A. Ocampo, "An application of DEMATEL-ANP and PROMETHEE II approach for air traffic controllers' workload stress problem: a case of Mactan Civil Aviation Authority of the Philippines," *Journal of Air Transport Management*, vol. 68, pp. 198–213, 2018.

[59] B. W. Zhu, J. R. Zhang, G. H. Tzeng, S. L. Huang, and L. Xiong, "Public open space development for elderly people by using the DANP-V model to establish continuous improvement strategies towards a sustainable and healthy aging society," *Sustainability*, vol. 9, pp. 1–29, 2017.

[60] W. K. M. Brauers and E. K. Zavadskas, "The MOORA method and its application to privatization in a transition economy," *Control and Cybernetics*, vol. 35, pp. 445–469, 2006.

[61] A. Arabsheybani, M. M. Paydar, and A. S. Safaei, "An integrated fuzzy MOORA method and FMEA technique for sustainable supplier selection considering quantity discounts and supplier's risk," *Journal of Cleaner Production*, vol. 190, pp. 577–591, 2018.

[62] S. J. Ghoushchi, S. Yousefi, and M. Khazaeili, "An extended FMEA approach based on the Z-MOORA and fuzzy BWM for prioritization of failures," *Applied Soft Computing*, vol. 81, Article ID 105505, 2019.

[63] H. Dinçer, S. Yüksel, and L. Martínez, "Interval type 2-based hybrid fuzzy evaluation of financial services in E7 economies with DEMATEL-ANP and MOORA methods," *Applied Soft Computing*, vol. 79, pp. 186–202, 2019.

[64] J. J. H. Liou, Y.-C. Chuang, E. K. Zavadskas, and G.-H. Tzeng, "Data-driven hybrid multiple attribute decision-making model for green supplier evaluation and performance improvement," *Journal of Cleaner Production*, vol. 241, Article ID 118321, 2019.

[65] L. Xiao, S. Zhang, G. Wei et al., "Green supplier selection in steel industry with intuitionistic fuzzy Taxonomy method," *Journal of Intelligent & Fuzzy Systems*, vol. 39, no. 5, pp. 7247–7258, 2020.

[66] G. Wei, M. Lu, F. E. Alsaadi, T. Hayat, and A. Alsaedi, "Pythagorean 2-tuple linguistic aggregation operators in multiple attribute decision making," *Journal of Intelligent & Fuzzy Systems*, vol. 33, no. 2, pp. 1129–1142, 2017.

[67] G. Wei, F. Lei, R. Lin et al., "Algorithms for probabilistic uncertain linguistic multiple attribute group decision making based on the GRA and CRITIC method: application to location planning of electric vehicle charging stations," *Economic Research-Ekonomska Istraživanja*, vol. 33, no. 1, pp. 828–846, 2020.

Hindawi
Journal of Healthcare Engineering
Volume 2021, Article ID 5559529, 10 pages
https://doi.org/10.1155/2021/5559529

Research Article

Complex Entropy and Its Application in Decision-Making for Medical Diagnosis

Fuyuan Xiao[1] **and Xiao-Guang Yue**[2]

[1]*School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu, China*
[2]*Department of Computer Science and Engineering, School of Sciences, European University Cyprus, Nicosia 1516, Cyprus*

Correspondence should be addressed to Fuyuan Xiao; complexbpa@163.com

Received 6 January 2021; Revised 20 January 2021; Accepted 16 February 2021; Published 25 February 2021

Academic Editor: Chi-Hua Chen

In decision-making systems, how to measure uncertain information remains an open issue, especially for information processing modeled on complex planes. In this paper, a new complex entropy is proposed to measure the uncertainty of a complex-valued distribution (CvD). The proposed complex entropy is a generalization of Gini entropy that has a powerful capability to measure uncertainty. In particular, when a CvD reduces to a probability distribution, the complex entropy will degrade into Gini entropy. In addition, the properties of complex entropy, including the nonnegativity, maximum and minimum entropies, and boundedness, are analyzed and discussed. Several numerical examples illuminate the superiority of the newly defined complex entropy. Based on the newly defined complex entropy, a multisource information fusion algorithm for decision-making is developed. Finally, we apply the decision-making algorithm in a medical diagnosis problem to validate its practicability.

1. Introduction

Uncertainty is inevitable in the applications of decision-making systems [1–4]. Considerable attention has addressed uncertainty in the past few decades [5, 6]. How to express the knowledge involved in sources of uncertain information still remains an open issue [7, 8]. Hence, researchers have attempted to model and measure uncertain information using extended soft sets [9], evidence theory [10], reasoning [11–13], belief structures [14, 15], D numbers [16, 17], Z numbers [18, 19], and other hybrid methods [20–23].

One successful alternative uncertain information measure is Gini entropy [24], which is simple to implement and has received a substantial amount of attention from researchers. Inspired by Gini entropy [24], Yager and Petry [25] recently devise an intelligent quality-based approach for fusing multisource information [26]. Bouhamed et al.[27] extend it to combine multisource possibilistic information. Later, researchers generalized the Gini entropy-based information quality to belief functions to measure uncertainty. The method of Li et al.[28, 29] is an example that has been well applied in various fields. Although Gini entropy [24] can be used to measure uncertainty, it can only be used for probability distributions.

The complex-valued model has potential expressional properties, especially for the modeling of uncertainty [30, 31]. Therefore, the complex-valued model was widely investigated and applied in various fields, such as medical diagnosis [32], decision-making [33, 34], and predicting interference effects [35, 36]. Given that the complex-valued representation model is well suited for certain applications, how can Gini entropy be generalized to complex planes to provide a more powerful capability to measure uncertainty?

In this paper, to address the abovementioned issue, a generalized entropy is proposed for measuring the uncertainty of CvDs. When CvDs reduce to probability distributions, the newly defined entropy degrades into Gini entropy. Specifically, vector expressions of CvDs are first proposed to model knowledge in complex planes. After that, a novel complex entropy called Xiao entropy is defined to measure uncertainties of CvDs. Then, the properties of complex entropy, including nonnegativity,

maximum and minimum entropies, and boundedness, are analyzed and discussed. Based on the newly defined complex entropy, a multisource information fusion algorithm for decision-making is devised. Finally, we apply the decision-making algorithm in a medical diagnosis problem to verify its practicability.

The contributions of this work are summarized as follows:

- A novel complex entropy, called Xiao entropy, which has the properties of nonnegativity, maximum and minimum entropies, and boundedness, is defined for the CvD
- The multisource information fusion algorithm based on the newly defined entropy can be well applied to support decision-making
- This study provides a new perspective of complex-valued representation for uncertain information and offers a promising and generalized solution in terms of uncertainty measurements

The preliminaries are introduced in Section 2. In Section 3, CvD vectors are defined. In Section 4, a complex entropy is defined to measure the uncertainty of CvDs. In Section 5, several numerical examples illustrate the properties of complex entropy. In Section 6, an algorithm for decision-making is designed on the basis of the newly defined entropy. Then, the decision-making algorithm is used in a medical diagnosis. Section 7 concludes this work.

2. Preliminaries

In this section, some essential concepts of uncertainty measures related to this work are introduced.

Definition 1 (Gini entropy). Let $P = [p_1, \ldots, p_j, \ldots, p_n]$ be a probability distribution vector. The Gini entropy of P is defined by [24]

$$G(P) = 1 - \sum_{j=1}^{n} p_j^2. \tag{1}$$

Definition 2 (Pennecchi and Oberto's uncertainty measures). Let $\mathbb{C} = [\mathfrak{c}_1, \ldots, \mathfrak{c}_j, \ldots, \mathfrak{c}_n]$ be a CvD vector, where $\mathfrak{c}_j = a_j + b_j i$. Pennecchi and Oberto's modulus estimations of $\mathbb{C}$ are defined by [37]

$$\begin{aligned} \langle \|\overrightarrow{c}\| \rangle_a &= \frac{1}{n} \sum_{j=1}^{n} \sqrt{x_j^2 + y_j^2}, \\ \langle \|\overrightarrow{c}\| \rangle_b &= \sqrt{\left(\frac{1}{n} \sum_{j=1}^{n} a_j\right)^2 + \left(\frac{1}{n} \sum_{j=1}^{n} b_j\right)^2}, \end{aligned} \tag{2}$$

where $1 - \langle \|\overrightarrow{c}\| \rangle_a$ and $1 - \langle \|\overrightarrow{c}\| \rangle_b$ can be used as uncertainty measures.

3. Vector Representation of CvD

Modeling uncertainty has attracted a substantial amount of attention in a variety of areas [38]. Many methods have been proposed and applied in various fields, such as failure and risk analysis [39], classification [40, 41], information fusion [42], and decision-making [43, 44]. Here, a vector representation of CvD is presented for expressing uncertainty in a complex plane. In addition, the norm of CvD is also defined and analyzed.

Definition 3 (CvD vector). Let $\mathbb{C}_k$ be a CvD vector on the frame of discernment (FOD) $\Psi = \{\psi_1, \ldots, \psi_j, \ldots, \psi_n\}$, denoted by

$$\mathbb{C}_k = \left[\mathfrak{c}_{k1}, \ldots, \mathfrak{c}_{kj}, \ldots, \mathfrak{c}_{kn}\right], \tag{3}$$

where $\mathfrak{c}_{kj}$ is the complex value with regard to the occurrence of ψ_j:

$$\mathfrak{c}_{kj} = a_{kj} + b_{kj} i, \tag{4}$$

where a_{kj} and b_{kj} are real numbers and i is the imaginary unit, satisfying $i^2 = -1$.

$\mathfrak{c}_{kj}$ in equation (4) satisfies

$$\begin{aligned} &\sqrt{a_{kj}^2 + b_{kj}^2} \in [0, 1], \\ &\sum_{j=1}^{n} \left|\mathfrak{c}_{kj}\right| = 1, \end{aligned} \tag{5}$$

where $|\mathfrak{c}_{kj}|$ is the modulus of $\mathfrak{c}_{kj}$.

Equation (4) is also expressed as follows:

$$\mathfrak{c}_{kj} = r_{kj} e^{i\theta_{kj}}, \tag{6}$$

with

$$r_{kj} = \sqrt{a_{kj}^2 + b_{kj}^2}, \tag{7}$$

where $r_{kj} = |\mathfrak{c}_{kj}| \geq 0$ and $\theta_{kj} \in [-\pi, \pi]$ denotes an angle (phase) of $\mathfrak{c}_{kj}$.

Definition 4 (norm of CvD). Let $\mathbb{C}_k$ be a CvD vector on FOD Ψ. Norm of CvD vector, $\|\mathbb{C}_k\|$, is defined by

$$\left\|\mathbb{C}_k\right\| = \sqrt{\sum_{j=1}^{n} \left|\mathfrak{c}_{kj}\right|^2} = \sqrt{\sum_{j=1}^{n} \left(a_{kj}^2 + b_{kj}^2\right)}. \tag{8}$$

Consider properties of CvD vector in Definition 3, where for each $\mathfrak{c}_{kj}$, $a_{kj}^2 + b_{kj}^2 \in [0, 1]$ and $\sum_{j=1}^{n} |\mathfrak{c}_{kj}| = 1$, we observe the following:

Case 1. The maximal value of $\|\mathbb{C}_k\|$, denoted by $\max[\|\mathbb{C}_k\|]$, is generated, when

$$\begin{cases} \mathfrak{c}_{kj} = 1, & \text{for one } j, \\ \mathfrak{c}_{kj} = 0, & \text{for others } j, \end{cases} \tag{9}$$

such that

$$\max\left[\left\|\mathbb{C}_k\right\|\right] = 1. \tag{10}$$

Case 2. When $\mathfrak{c}_{kj}$ degrades into real numbers, i.e., $\mathfrak{c}_{kj} = a_{kj}$ ($b_{kj} = 0$), the minimum value of $\|\mathbb{C}_k\|$, denoted by $\min[\|\mathbb{C}_k\|]$, is generated, when

$$\mathfrak{c}_{kj} = \frac{1}{n}, \quad 1 \le j \le n, \tag{11}$$

where

$$\min\left[\left\|\mathbb{C}_k\right\|\right] = \sqrt{\sum_{j=1}^{n}\left(\frac{1}{n}\right)^2} = \sqrt{\frac{1}{n}} = \frac{1}{\sqrt{n}}. \tag{12}$$

In summary,

$$\frac{1}{\sqrt{n}} \le \left\|\mathbb{C}_k\right\| \le 1, \tag{13}$$

where $\|\mathbb{C}_k\|$ has a maximum value of 1 with $\mathfrak{c}_{kj} = 1$ for one ψ_j and others $\mathfrak{c}_{kj} = 0$; $\|\mathbb{C}_k\|$ has a minimal value of $1/\sqrt{n}$ with all $\mathfrak{c}_{kj} = (1/n)$.

4. Entropy for CvD

Entropy is useful for measuring uncertainty [45–47], where many kinds of entropies, such as Tsallis entropy [48], fuzzy entropy [49, 50], Deng entropy [51–53], and cross-entropy [54], are presented for different aspects [55–59]. Among them, Shannon and Gini entropies are very popular. The greater the uncertainty is, the greater the entropy is; the lesser the uncertainty is, the lesser the entropy is [60]. We make use of the concept of Gini entropy [24] to measure the uncertainty of CvD.

Definition 5 (complex entropy). Let $\mathbb{C}_k$ be a CvD vector on FOD Ψ. The complex entropy of $\mathbb{C}_k$, denoted as $E_{\mathcal{X}}(\mathbb{C}_k)$, is defined as

$$E_{\mathcal{X}}(\mathbb{C}_k) = 1 - \left\|\mathbb{C}_k\right\|^2 = 1 - \sum_{j=1}^{n}\left|\mathfrak{c}_{kj}\right|^2. \tag{14}$$

When a CvD reduces to a probability distribution, where $b_{kj} = 0$ and $\mathfrak{c}_{kj} = a_{kj}$, then $E_{\mathcal{X}}(\mathbb{C}_k)$ can be expressed as follows:

$$E_{\mathcal{X}}(\mathbb{C}_k) = 1 - \sum_{j=1}^{n}\left|\mathfrak{c}_{kj}\right|^2 = 1 - \sum_{j=1}^{n} a_{kj}^2, \tag{15}$$

which is equal to equation (1).

Property 1. $E_{\mathcal{X}}$ is a generalized model of Gini entropy [24]. Specifically, when a CvD becomes a probability distribution, $E_{\mathcal{X}}$ degrades into Gini entropy [24].

According to equation (13), because $(1/\sqrt{n}) \le \|\mathbb{C}_k\| \le 1$, we have

$$\frac{1}{n} \le \left\|\mathbb{C}_k\right\|^2 \le 1, \tag{16}$$

such that

$$0 \le 1 - \left\|\mathbb{C}_k\right\|^2 \le 1 - \frac{1}{n}. \tag{17}$$

Thus, $E_{\mathcal{X}}(\mathbb{C}_k)$ has the boundedness of $[0, (1/n)]$.
It is inferred that

$E_{\mathcal{X}}(\mathbb{C}_k)$ reaches its maximal value $E_{\mathcal{X}}(\mathbb{C}_k) = 1 - (1/n)$ when $\mathfrak{c}_{kj} = (1/n)$ for $1 \le j \le n$. When $n \longrightarrow +\infty$ and $(1/n) \longrightarrow 0$, $\mathcal{X}(\mathbb{C}_k)$ reaches the maximum value 1.

$E_{\mathcal{X}}(\mathbb{C}_k)$ reaches its minimal value $E_{\mathcal{X}}(\mathbb{C}_k) = 0$ when one $\mathfrak{c}_{kj} = 1$ and others $\mathfrak{c}_{kj} = 0$.

Remark 1. Notably, the larger $E_{\mathcal{X}}(\mathbb{C}_k)$ is, the larger the uncertainty in CvD $\mathbb{C}_k$ is, which results in lower certainty.

Definition 6 (the completely certain CvD). CvD $\mathbb{C}_k$ is completely certain when $E_{\mathcal{X}}(\mathbb{C}_k) = 0$.

Definition 7 (the completely uncertain CvD).
CvD $\mathbb{C}_k$ is completely uncertain when $E_{\mathcal{X}}(\mathbb{C}_k) = 1$.

Theorem 1. *$E_{\mathcal{X}}$ has the desired properties of the entropy of the CvD, including nonnegativity, maximum and minimum entropies, and boundedness.*

Property 2. Let $\mathbb{C}_k$ be an arbitrary CvD:

P 2.1 Nonnegativity: $E_{\mathcal{X}}(\mathbb{C}_k) \ge 0$

P 2.2 Maximum entropy: $E_{\mathcal{X}}(\mathbb{C}_k) \le \max[E_{\mathcal{X}}(\mathbb{C}_k)]$

P 2.3 Minimum entropy: $E_{\mathcal{X}}(\mathbb{C}_k) \ge \min[E_{\mathcal{X}}(\mathbb{C}_k)]$

P 2.4 Boundedness: $0 \le E_{\mathcal{X}}(\mathbb{C}_k) \le 1$

Proof. The proofs are trivial. □

5. Numerical Examples

In this section, several examples are presented to illustrate the entropy for CvD.

Example 1. Consider a CvD $\mathbb{C}$ in the FOD $\Psi = \{\psi_1, \psi_2\}$:

$$\mathbb{C} = [x, 1 - x]. \tag{18}$$

In Example 1, $\mathbb{C}$ changes as parameter x varies, where x is set within [0,1], such that $\mathbb{C}$ reduces to a probability distribution.

By leveraging the Gini entropy G and Xiao entropy $E_{\mathcal{X}}$, the corresponding entropy measures are shown in Figure 1. Clearly, $E_{\mathcal{X}}$ is the same as G entropy, which verifies that when a CvD reduces to a probability distribution, $E_{\mathcal{X}}$ degrades into Gini entropy. Additionally, when $x = 0$ or $x = 1$, such that $\mathbb{C} = [0, 1]$ or $\mathbb{C} = [1, 0]$, $G(\mathbb{C})$ and $E_{\mathcal{X}}(\mathbb{C})$ achieve the minimum entropy of 0, because in this case, $\mathbb{C}$ is the completely certain CvD. By contrast, only when $x = 0.5$,

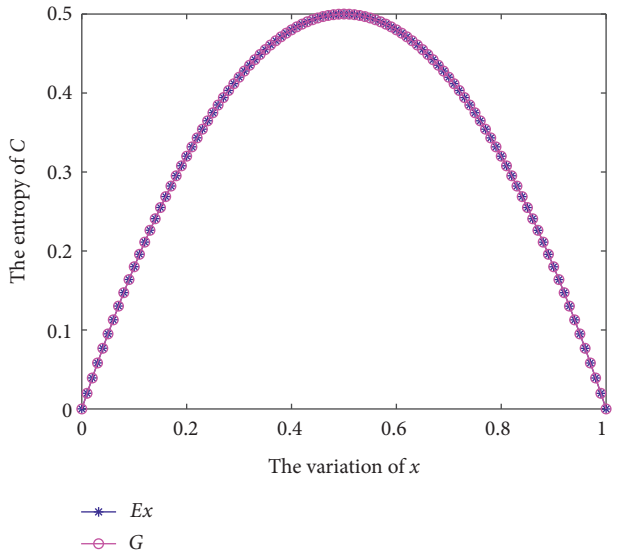


Figure 1: G and $E_{\mathcal{X}}$ in Example 1.

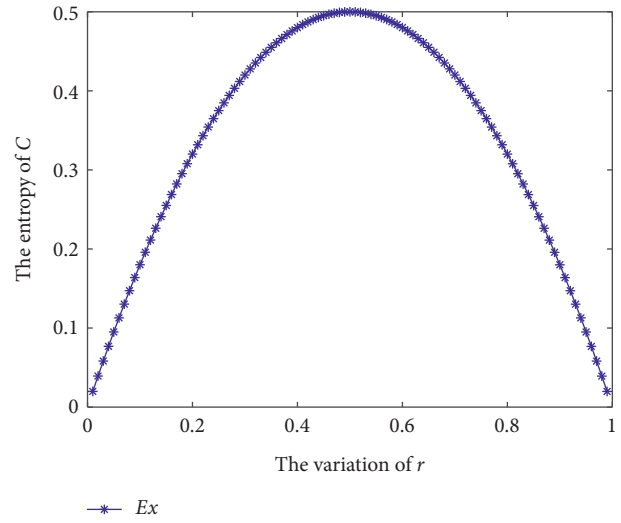


Figure 2: $E_{\mathcal{X}}$ in Example 2.

such that $\mathbb{C} = [0.5, 0.5]$, can $G(\mathbb{C})$ and $E_{\mathcal{X}}(\mathbb{C})$ achieve a maximum entropy of 0.5.

Example 2. Consider a CvD $\mathbb{C}$ in the FOD $\Psi = \{\psi_1, \psi_2\}$:

$$\mathbb{C} = \left[re^{((\pi/2)i)}, (1-r)e^{((\pi/9)i)}\right]. \tag{19}$$

In Example 2, $\mathbb{C}$ changes as modulus r varies, where r is set within [0.01,0.99].

Because $\mathbb{C}$ consists of complex numbers, Gini entropy is not applicable. The result of $E_{\mathcal{X}}$ entropy is shown in Figure 2. As r increases from 0.01 to 0.5, $E_{\mathcal{X}}$ entropy increases from 0.0198 to 0.5, while as r increases from 0.5 to 0.99, $E_{\mathcal{X}}$ entropy gradually decreases to 0.0198. This result shows a similar trend as the entropy measures in Figure 1.

A comparison of the results in Examples 1 and 2 shows that the proposed $E_{\mathcal{X}}$ entropy is a more capable uncertainty measure than Gini entropy.

Example 3. Consider a CvD $\mathbb{C}$ in the FOD $\Psi = \{\psi_1, \ldots, \psi_j, \ldots, \psi_x\}$:

$$\mathbb{C} = \left[\frac{1}{\alpha}, \ldots, \frac{1}{\alpha}, \ldots, \frac{1}{\alpha}\right]. \tag{20}$$

In Example 3, we set six different scales of α, namely, $\alpha \in [1, 10]$, $[1, 10^2]$, $[1, 10^3]$, $[1, 10^4]$, $[1, 10^5]$, and $[1, 10^6]$, to measure the variation of $G(\mathbb{C})$ and $E_{\mathcal{X}}(\mathbb{C})$.

Figures 3(a)–3(f) depict the results of $G(\mathbb{C})$ and $E_{\mathcal{X}}(\mathbb{C})$ with regard to six different cases, respectively. Particularly, as α varies within [1, 10], $E_{\mathcal{X}}(\mathbb{C})$ has a maximum value of $\max[E_{\mathcal{X}}(\mathbb{C})] = 0.9$ and a minimum value of $\min[E_{\mathcal{X}}(\mathbb{C})] = 0$. When α changes within $[1, 10^2]$, $\max[E_{\mathcal{X}}(\mathbb{C})] = 0.99$ and $\min[E_{\mathcal{X}}(\mathbb{C})] = 0$. When α varies within $[1, 10^3]$, $\max[E_{\mathcal{X}}(\mathbb{C})] = 0.999$ and $\min[E_{\mathcal{X}}(\mathbb{C})] = 0$. When α changes within $[1, 10^4]$, $\max[E_{\mathcal{X}}(\mathbb{C})] = 0.9999$ and $\min[E_{\mathcal{X}}(\mathbb{C})] = 0$. When α varies within $[1, 10^5]$, $\max[E_{\mathcal{X}}(\mathbb{C})] = 1$ and $\min[E_{\mathcal{X}}(\mathbb{C})] = 0$. When α changes within $[1, 10^6]$, $\max[E_{\mathcal{X}}(\mathbb{C})] = 1$ and $\min[E_{\mathcal{X}}(\mathbb{C})] = 0$. Thus, when a CvD becomes a completely certain distribution, i.e., a probability distribution, in which $\mathfrak{c}_{kj} = a_{kj} = 1$ for one j and other $\mathfrak{c}_{kj} = 0$, it has a minimum entropy of $\min[E_{\mathcal{X}}(\mathbb{C})] = 0$. On the other hand, when $\alpha \longrightarrow +\infty$, $\max[E_{\mathcal{X}}(\mathbb{C})]$ is close to 1, because in this case $\mathbb{C}$ is completely uncertain.

Example 4. Assume that there is a CvD $\mathbb{C}$ in the FOD $\Psi = \{\psi_1, \ldots, \psi_j, \ldots, \psi_x\}$:

$$\mathbb{C} = \left[re^{(\xi\pi i)}, (1-r)e^{(\xi\pi i)}\right]. \tag{21}$$

In Example 4, $\mathbb{C}$ changes as r and ξ vary. Here, we set r within [0,1] and ξ within [-1,1], as shown in Figure 4(a). The entropy measure of $E_{\mathcal{X}}(\mathbb{C})$ is presented in Figure 4(b), which shows how the variations in the modulus and angle of the elements in $\mathbb{C}$ impact $E_{\mathcal{X}}(\mathbb{C})$.

$E_{\mathcal{X}}(\mathbb{C})$ changes as r varies, whereas the variation in angle $\theta = \xi\pi$ has no effect on $E_{\mathcal{X}}(\mathbb{C})$. This result is reasonable because $r_{kj}^2 = |\mathfrak{c}_{kj}|^2 = a_{kj}^2 + b_{kj}^2$ is related to the modulus r rather than θ.

Example 5. Consider Example 2.

In Example 5, r is set within [0,1]. We compare the proposed $E_{\mathcal{X}}$ with related works, that is, Pennecchi and Oberto's uncertainty measures $1 - \langle\|\overrightarrow{c}\|\rangle_a$ and $1 - \langle\|\overrightarrow{c}\|\rangle_b$.

By comparing the results of $E_{\mathcal{X}}$, $1 - \langle\|\overrightarrow{\mu}\|\rangle_a$, and $1 - \langle\|\overrightarrow{\mu}\|\rangle_b$ shown in Figure 5, we can see that $1 - \langle\|\overrightarrow{c}\|\rangle_a$ remains 0.5 and cannot accurately measure the uncertainty. However, $1 - \langle\|\overrightarrow{\mu}\|\rangle_b$ provides a better measure of the uncertainty compared to $1 - \langle\|\overrightarrow{c}\|\rangle_a$ because as r increases from 0.01 to 0.5, it increases from 0.2929 to 0.4208, while as r increases from 0.5 to 0.99, it gradually decreases to 0.2929. Nevertheless, the proposed $E_{\mathcal{X}}$ has better discrimination as an uncertainty measurement and is superior to other methods.

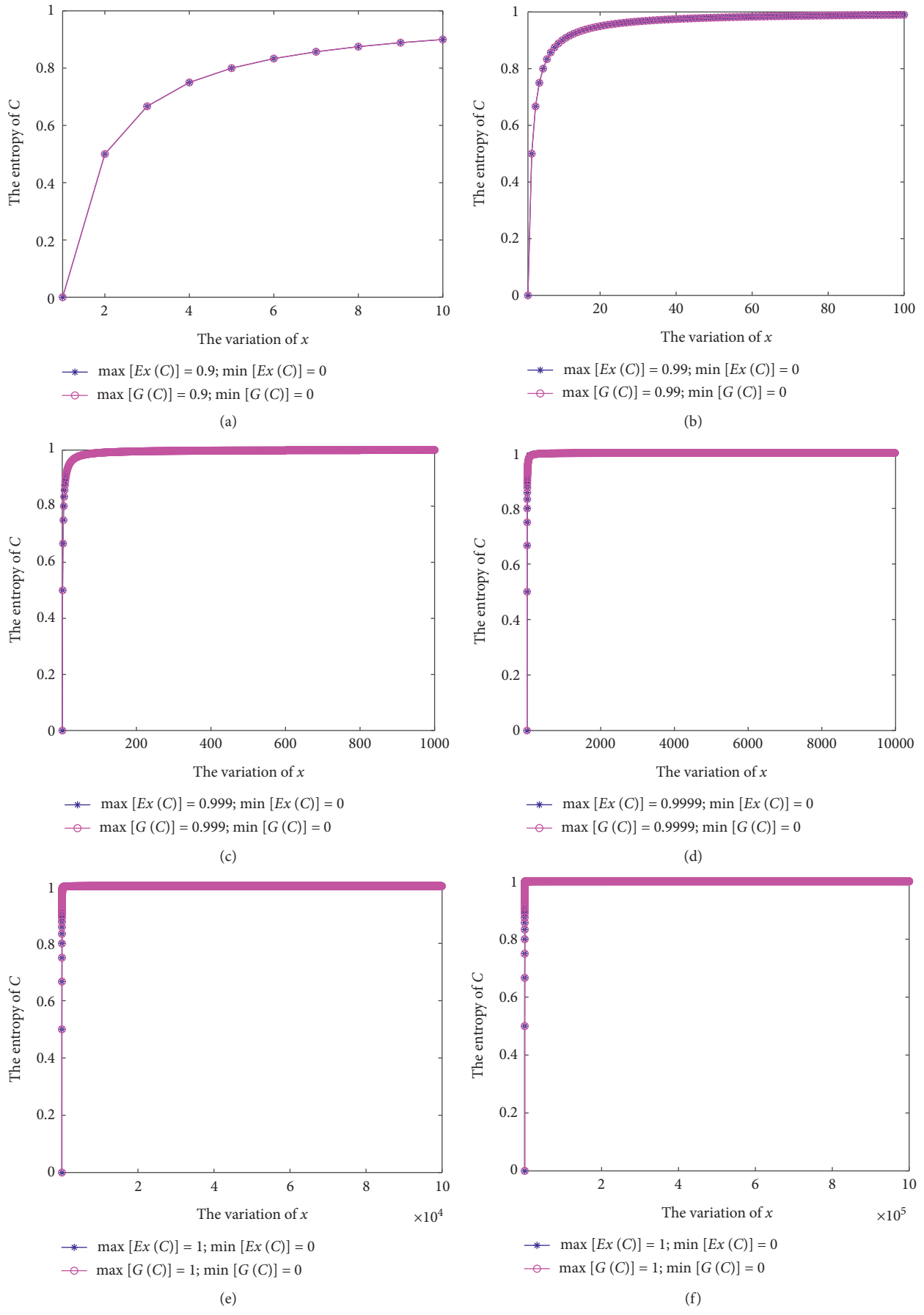


FIGURE 3: The entropy measures of $G(\mathbb{C})$ and $E_{\mathcal{X}}(\mathbb{C})$ in Example 3. (a)$G(\mathbb{C})$ and $E_{\mathcal{X}}(\mathbb{C})$: $1 \leq \alpha \leq 10$. (b)$G(\mathbb{C})$ and $E_{\mathcal{X}}(\mathbb{C})$: $1 \leq \alpha \leq 10^2$. (c)$G(\mathbb{C})$ and $E_{\mathcal{X}}(\mathbb{C})$: $1 \leq \alpha \leq 10^3$. (d)$G(\mathbb{C})$ and $E_{\mathcal{X}}(\mathbb{C})$: $1 \leq \alpha \leq 10^4$. (e)$G(\mathbb{C})$ and $E_{\mathcal{X}}(\mathbb{C})$: $1 \leq \alpha \leq 10^5$. (f)$G(\mathbb{C})$ and $E_{\mathcal{X}}(\mathbb{C})$: $1 \leq \alpha \leq 10^6$.

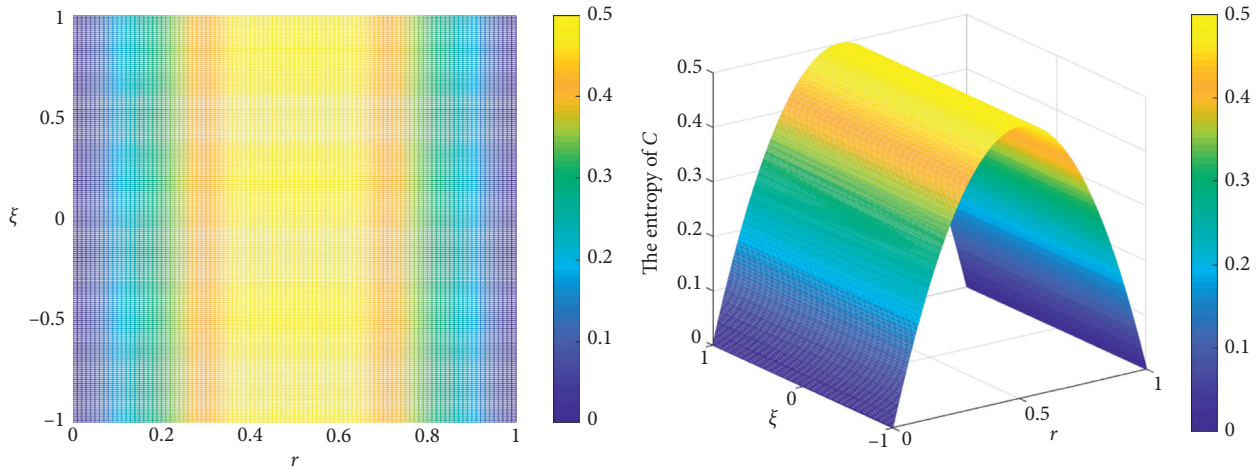


Figure 4: The entropy measure in Example 4. (a) The variation of r and ξ. (b)$E_{\mathcal{X}}$

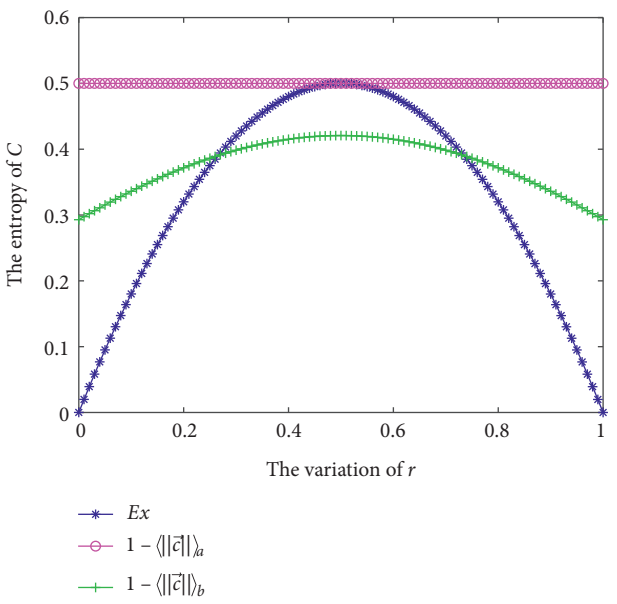


Figure 5: Comparison of different uncertainty measures.

6. Algorithm and Application

How to deal with decision-making problems has attracted much attention [61–65], especially for complex-valued expressed information [66, 67]. In this section, we first design a multisource information fusion algorithm for decision-making based on the proposed entropy. Then, we apply the decision-making algorithm in medical diagnosis to validate its practicability.

6.1. A Multisource Information Fusion Algorithm for Decision-Making. Problem statement: let Ψ be a FOD with a set of objectives $\{\psi_1, \ldots, \psi_j, \ldots, \psi_n\}$ to be recognized. Suppose there are t CvDs: $\mathbb{C} = \{\mathbb{C}_1, \ldots, \mathbb{C}_k, \ldots, \mathbb{C}_t\}$ where $\mathbb{C}_k = [\mathfrak{c}_{k1}, \ldots, \mathfrak{c}_{kj}, \ldots, \mathfrak{c}_{kn}]$ and $\mathfrak{c}_{kj} = a_{kj} + b_{kj}i$. The decision-making algorithm is to identify the target from $\{\psi_1, \ldots, \psi_j, \ldots, \psi_n\}$ by combining multiple CvDs $\{\mathbb{C}_1, \ldots, \mathbb{C}_k, \ldots, \mathbb{C}_t\}$.

The specific steps are given as follows:

Step 1: For $1 \le k \le t$, its corresponding entropy of CvD $\mathbb{C}_k$, denoted by $E_{\mathcal{X}}(\mathbb{C}_k)$, can be generated as follows:

$$E_{\mathcal{X}}(\mathbb{C}_k) = 1 - \left\|\mathbb{C}_k\right\|^2. \tag{22}$$

Step 2: For $1 \le k \le t$, its corresponding information volume of CvD $\mathbb{C}_k$, denoted by $IV(\mathbb{C}_k)$, can be measured by

$$IV(\mathbb{C}_k) = e^{E_{\mathcal{X}}(\mathbb{C}_k)}. \tag{23}$$

Step 3: The information volume $IV(\mathbb{C}_k)$ is normalized by

$$\overline{IV}(\mathbb{C}_k) = \frac{IV(\mathbb{C}_k)}{\sum_{h=1}^{t} IV(\mathbb{C}_h)}, \quad 1 \le k \le t. \tag{24}$$

Step 4: According to the normalized information volumes, the weighted average CvD, denoted as $\widetilde{\mathbb{C}}$, is defined by

$$\widetilde{\mathbb{C}} = \sum_{i=1}^{k} \left[\overline{IV}(\mathbb{C}_k) \times \left|\mathbb{C}_k\right|\right], \quad 1 \le k \le t, \tag{25}$$

where $\left|\mathbb{C}_k\right| = [|\mathfrak{c}_{k1}|, \ldots, |\mathfrak{c}_{kj}|, \ldots, |\mathfrak{c}_{kn}|]$ and $|\mathfrak{c}_{kj}| = \sqrt{x_{kj}^2 + y_{kj}^2}$.

Step 5: $\widetilde{\mathbb{C}}$ is fused via the complex Dempster's combination rule [68] by $t-1$ times:

$$\widetilde{\widetilde{\mathbb{C}}}_{t-1} = \left((\widetilde{\mathbb{C}} \oplus \widetilde{\mathbb{C}})_1 \oplus \cdots \oplus \widetilde{\mathbb{C}}\right)_{t-1}. \tag{26}$$

Step 6: For $\widetilde{\widetilde{\mathbb{C}}}_{t-1}(\psi_j)$, the ψ_δ with the maximum absolute value is chosen:

$$\widetilde{\widetilde{\mathbb{C}}}_{t-1}(\psi_\delta) = \max_{1 \le j \le n}\left\{\widetilde{\widetilde{\mathbb{C}}}_{t-1}\left(\psi_j\right)\right\}, \quad 1 \le j \le n. \tag{27}$$

Input: A FOD $\Psi = \{\psi_1, \ldots, \psi_j, \ldots, \psi_n\}$;
A set of CvDs:$\mathbb{C} = \{\mathbb{C}_1, \ldots, \mathbb{C}_k, \ldots, \mathbb{C}_t\}$
(1) **for** $1 \le k \le t$ **do**
(2) Calculate the CvD entropy $E_{\mathcal{X}}(\mathbb{C}_k)$ by equation (22)
(3) Measure the CvD information volume $IV(\mathbb{C}_k)$ by equation (23)
(4) **end**
(5) Calculate the normalized information volume $\overline{IV}(\mathbb{C}_k)$ by equation (24)
(6) Generate the weighted average CvD $\widetilde{\mathbb{C}}$ by equation (25)
(7) Obtain the fused $\widehat{\widetilde{\mathbb{C}}}_{t-1}$ via the complex Dempster's combination rule by equation (26)
(8) Choose the maximum absolute value $\widehat{\widetilde{\mathbb{C}}}_{t-1}(\psi_\delta)$ by equation (27)
(9) **if** $\widehat{\widetilde{\mathbb{C}}}_{t-1}(\psi_\delta) \ge \lambda$ **then**
(10) $\delta = \text{argmax}_{1 \le j \le n}\left\{\widehat{\widetilde{\mathbb{C}}}_{t-1}(\psi_j)\right\}$
(11) Target $\leftarrow \psi_\delta$.
(12) **else**
(13) Cannot be determined.
(14) **end**

Algorithm 1: Complex entropy-based multisource information fusion algorithm for decision-making.

Table 1: The evaluated data for patient modeled as CvDs.

Experts	CvDs	Diseases			
		Viral fever: D_1	Malaria: D_2	Typhoid: D_3	Stomach problem: D_4
E_1	$\mathbb{C}_{E_1}$	$0.65e^{0.2i}$	$0.10e^{0.3i}$	$0.10e^{0.3i}$	$0.15e^{0.2i}$
E_2	$\mathbb{C}_{E_2}$	$0.10e^{0.3i}$	$0.60e^{0.2i}$	$0.10e^{0.3i}$	$0.20e^{0.2i}$
E_3	$\mathbb{C}_{E_3}$	$0.40e^{0.3i}$	$0.10e^{0.3i}$	$0.30e^{0.4i}$	$0.20e^{0.4i}$
E_4	$\mathbb{C}_{E_4}$	$0.50e^{0.2i}$	$0.20e^{0.2i}$	$0.10e^{0.3i}$	$0.20e^{0.3i}$
E_5	$\mathbb{C}_{E_5}$	$0.55e^{0.2i}$	$0.10e^{0.3i}$	$0.15e^{0.2i}$	$0.20e^{0.3i}$

Step 7: Let λ be a threshold value for decision-making, which can be set in advance according to specific applications. If $\widehat{\widetilde{\mathbb{C}}}_{t-1}(\psi_\delta) \ge \lambda$, the ψ_δ can be identified as the target by

$$\delta = \arg\max_{1 \le j \le n}\left\{\widehat{\widetilde{\mathbb{C}}}_{t-1}\left(\psi_j\right)\right\}, \quad \text{Target} \leftarrow \psi_\delta. \tag{28}$$

If $\widehat{\widetilde{\mathbb{C}}}_{t-1}(\psi_\delta) < \lambda$, it cannot be determined.
The corresponding pseudocode is given in Algorithm 1.

6.2. Application in Medical Diagnosis. In this section, the proposed decision-making method is applied in medical diagnosis to demonstrate its practicability. The scenario and data of the application are based on [32].

Considering a medical diagnosis problem, where for a patient P, P suffers with the most possible disease from $D = \{D_1$: viral fever, D_2: malaria, D_3: typhoid, D_4: stomach problem$\}$. To clarify which disease the patient may suffer, five experts diagnose the patient's condition, in which the evaluation data are modeled as CvDs in Table 1. The threshold λ is set as 0.80 for this application to make a decision. We try to diagnose the patient P by integrating the evaluations from the five experts.

Then, the decision-making algorithm is applied to medical diagnosis by the following steps:

Step 1: The entropy values of CvD $\mathbb{C}_{E_k}$ $(1 \le k \le 5)$ are calculated by equation (22), as shown in Table 2.

Step 2: The information volumes of CvD $\mathbb{C}_{E_k}$ $(1 \le k \le 5)$ are calculated by equation (23), as shown in Table 2.

Step 3: The information volume $IV(\mathbb{C}_{E_k})$ $(1 \le k \le 5)$ is normalized by equation (24), as shown in Table 2.

Step 4: The weighted average CvD $\widetilde{\mathbb{C}}$ is generated by equation (25), as shown in Table 3.

Step 5: By gradually fusing the weighted average CvD with 4 times, their corresponding results are generated by equation (26), as shown in Table 3.

Step 6: The maximal absolute value of $\widehat{\widetilde{\mathbb{C}}}_{t-1}(D_j)$ is marked with the correct color in Table 3.

Step 7: Patient P is diagnosed as most likely to suffer the disease D_1:

$$\widehat{\widetilde{\mathbb{C}}}_4(D_\delta) = \max_{1 \le j \le 4}\left\{\widehat{\widetilde{\mathbb{C}}}_4\left(D_j\right)\right\} \ge \theta, \quad \theta = 0.80$$
$$\delta = \arg\max_{1 \le j \le 4}\left\{\widehat{\widetilde{\mathbb{C}}}_4\left(D_j\right)\right\} = 1, \tag{29}$$
$$P \leftarrow D_1.$$

6.3. Discussion. As shown in Table 1, we see that $|\mathbb{C}_{E_1}(D_1)| = 0.65$, $|\mathbb{C}_{E_3}(D_1)| = 0.4$, $|\mathbb{C}_{E_4}(D_1)| = 0.5$, and $|\mathbb{C}_{E_5}(D_1)| = 0.55$, which all support viral fever: D_1 disease.

Table 2: The results in terms of entropy, information volume, and normalized information volume.

Results	CvDs				
	$\mathbb{C}_{E_1}$	$\mathbb{C}_{E_2}$	$\mathbb{C}_{E_3}$	$\mathbb{C}_{E_4}$	$\mathbb{C}_{E_5}$
$E_{\mathcal{X}}(\mathbb{C}_{E_k})$	0.5350	0.5800	0.7000	0.6600	0.6250
$IV(\mathbb{C}_{E_k})$	1.7074	1.7860	2.0138	1.9348	1.8682
$\overline{IV}(\mathbb{C}_{E_k})$	0.1834	0.1918	0.2163	0.2078	0.2007

Table 3: The weighted average CvD and fused results obtained by the complex Dempster's combination rule.

Results	Diseases				Diagnosis results
	Viral fever: D_1	Malaria: D_2	Typhoid: D_3	Stomach problem: D_4	
$\widetilde{\mathbb{C}}$	0.4392	0.2167	0.1533	0.1908	Cannot be determined
$\widehat{\mathbb{C}}_1$	0.6435	0.1567	0.0784	0.1215	Cannot be determined
$\widehat{\mathbb{C}}_2$	0.8034	0.0965	0.0342	0.0659	Viral fever
$\widehat{\mathbb{C}}_3$	0.9011	0.0534	0.0134	0.0321	Viral fever
$\widehat{\mathbb{C}}_4$	0.9525	0.0279	0.0049	0.0148	Viral fever

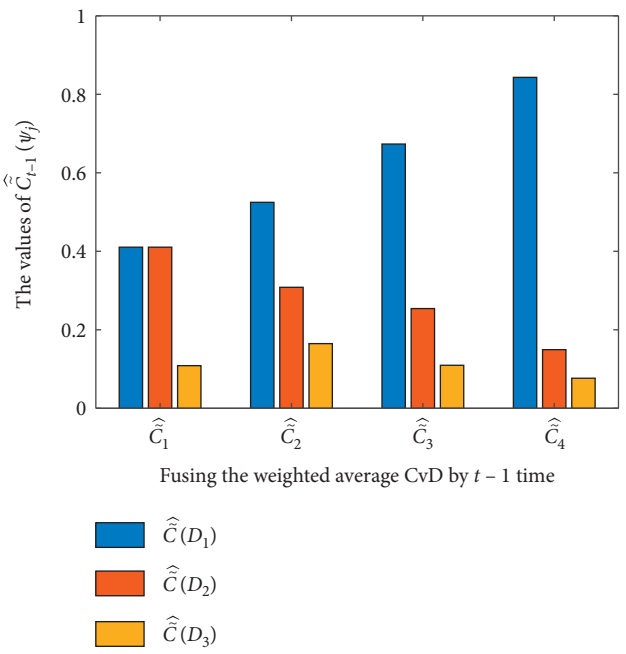


Figure 6: The fusion results of the weighted average CvD.

However, $|\mathbb{C}_{E_2}(D_3)| = 0.6$ supports malaria: D_2 disease. Hence, $\mathbb{C}_{E_2}$ conflicts with $\mathbb{C}_{E_1}$, $\mathbb{C}_{E_3}$, $\mathbb{C}_{E_4}$, and $\mathbb{C}_{E_5}$. By only using Table 1, it is difficult to make an accurate decision because a conflict exists among the experts. It is necessary to fuse the data collected from different experts to better support decision-making. There are five evaluations from five experts. To illuminate the effectiveness of the proposed decision-making algorithm, we gradually fuse the weighted average CvD, and the results are given in Table 3.

When the weighted average CvD is fused by 1 time, we obtain the result that $\widetilde{\mathbb{C}}_1(D_1)$ has the largest value of 0.6435. Because 0.6435 is smaller than the threshold $\lambda = 0.80$, the patient's disease cannot be determined. When the weighted average CvD is fused by 2 times, it is calculated that $\widetilde{\mathbb{C}}_2(D_1)$ has the largest value of 0.8034. Because 0.8034 is larger than the threshold $\lambda = 0.80$, the patient is diagnosed with viral fever: D_1. When the weighted average CvD is fused by 3 and 4 times, it is easy to see that $\widetilde{\mathbb{C}}_3(D_1)$ and $\widetilde{\mathbb{C}}_4(D_1)$ have increasingly large values of 0.9011 and 0.9525 to better support decision-making. Finally, the patient is diagnosed as most likely to suffer viral fever: D_1. Consequently, the value in terms of disease D_1 is increased for decision-making from 0.6435 to 0.8034 to 0.9011 and then to 0.9525 as shown in Figure 6. As a result, the proposed decision-making algorithm is effective to address medical diagnosis problem.

7. Conclusions

In this paper, a complex entropy, called Xiao entropy, is proposed to measure the uncertainty of complex-valued distributions (CvDs). The complex entropy is a generalized model of Gini entropy. Specifically, when the CvD turns into a probability distribution, the proposed entropy degrades into Gini entropy. Furthermore, we study the properties of complex entropy, including nonnegativity, maximum and minimum entropies, and boundedness. Several numerical examples compare the proposed complex entropy with related works. The results illuminate the superiority of the proposed complex entropy. Based on the complex entropy, a multisource information fusion algorithm for decision-making is devised. Finally, we apply the decision-making algorithm in a medical diagnosis problem to validate its practicability.

The main contributions are that this study provides a new perspective of complex-valued representation for uncertain information; the newly defined complex entropy has a powerful capability to measure uncertainty. Additionally, it offers a promising application in decision theory. In the future work, we intend to apply this complex entropy to handle more complex decision-making problems, such as the analyzing and processing of image and physiological signals.

Data Availability

The data used to support the findings of this study are provided in the article.

Conflicts of Interest

The author states that there are no conflicts of interest.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (nos. 62003280 and 61902189).

References

[1] R. R. Yager and M. Z. Reformat, "Selecting an action to satisfy multiple aspects of a system based on uncertain granular observations," *Expert Systems with Applications*, vol. 126, pp. 1–8, 2019.

[2] Z. Liu, G. Li, G. Mercier, Y. He, and Q. Pan, "Change detection in heterogenous remote sensing images via homogeneous pixel transformation," *IEEE Transactions on Image Processing*, vol. 27, no. 4, pp. 1822–1834, 2017.

[3] B. Kang, P. Zhang, Z. Gao, G. Chhipi-Shrestha, K. Hewage, and R. Sadiq, "Environmental assessment under uncertainty using Dempster-Shafer theory and Z-numbers," *Journal of Ambient Intelligence and Humanized Computing*, vol. 11, no. 5, pp. 2041–2060, 2020.

[4] L. Fei, Y. Feng, and L. Liu, "On Pythagorean fuzzy decision making using soft likelihood functions," *International Journal of Intelligent Systems*, vol. 34, no. 12, pp. 3317–3335, 2019.

[5] F. Xiao, "Evidential fuzzy multicriteria decision making based on belief entropy," *IEEE Transactions on Fuzzy Systems*, vol. 28, no. 7, pp. 1477–1491, 2020.

[6] L. Fei and Y. Feng, "An attitudinal nonlinear integral and applications in decision making," *International Journal of Fuzzy Systems*, 2020.

[7] Y. Xue, Y. Deng, and H. Garg, "Uncertain database retrieval with measure - based belief function attribute values under intuitionistic fuzzy set," *Information Sciences*, vol. 546, pp. 436–447, 2021.

[8] X. Wang and Y. Song, "Uncertainty measure in evidence theory with its applications," *Applied Intelligence*, vol. 48, no. 7, pp. 1672–1688, 2018.

[9] F. Feng, J. Cho, W. Pedrycz, H. Fujita, and T. Herawan, "Soft set based association rule mining," *Knowledge-Based Systems*, vol. 111, pp. 268–282, 2016.

[10] R. R. Yager, "Generalized dempster-shafer structures," *IEEE Transactions on Fuzzy Systems*, vol. 27, no. 3, pp. 428–435, 2019.

[11] Z.-G. Liu, Q. Pan, J. Dezert, and A. Martin, "Combination of classifiers with optimal weight based on evidential reasoning," *IEEE Transactions on Fuzzy Systems*, vol. 26, no. 3, pp. 1217–1230, 2018.

[12] L. Fei and Y. Feng, "A novel retrieval strategy for case-based reasoning based on attitudinal Choquet integral," *Engineering Applications of Artificial Intelligence*, vol. 94, Article ID 103791, 2020.

[13] M. Zhou, X. Liu, and J. Yang, "Evidential reasoning approach for MADM based on incomplete interval value," *Journal of Intelligent & Fuzzy Systems*, vol. 33, no. 6, pp. 3707–3721, 2017.

[14] X. Xu, H. Xu, C. Wen, J. Li, P. Hou, and J. Zhang, "A belief rule-based evidence updating method for industrial alarm system design," *Control Engineering Practice*, vol. 81, pp. 73–84, 2018.

[15] C. Fu, M. Xue, D.-L. Xu, and S.-L. Yang, "Selecting strategic partner for tax information systems based on weight learning with belief structures," *International Journal of Approximate Reasoning*, vol. 105, pp. 66–84, 2019.

[16] X. Deng and W. Jiang, "A total uncertainty measure for D numbers based on belief intervals," *International Journal of Intelligent Systems*, vol. 34, no. 12, p. 3302, 2019.

[17] B. Liu and Y. Deng, "Risk evaluation in failure mode and effects analysis based on D numbers theory," *International Journal of Computers Communications & Control*, vol. 14, no. 5, pp. 672–691, 2019.

[18] W. Jiang, Y. Cao, and X. Deng, "A novel Z-network model based on Bayesian network and Z-number," *IEEE Transactions on Fuzzy Systems*, vol. 28, no. 8, p. 1585, 2020.

[19] Y. Tian, L. Liu, X. Mi, and B. Kang, "Zslf:a new soft likelihood function based on Z-numbers and its application in expert decision system," *IEEE Transactions on Fuzzy Systems*, vol. 22, no. 7, pp. 2333–2349, 2020.

[20] L. Fei, Y. Feng, and L. Liu, "Evidence combination using OWA-based soft likelihood functions," *International Journal of Intelligent Systems*, vol. 34, no. 9, pp. 2269–2290, 2019.

[21] C.-H. Chen, "An arrival time prediction method for bus system," *IEEE Internet of Things Journal*, vol. 5, no. 5, pp. 4231-4232, 2018.

[22] W. Jiang, K. Huang, J. Geng, and X. Deng, "Multi-scale metric learning for few-shot learning," *IEEE Transactions on Circuits and Systems for Video Technology*, p. 1, 2020.

[23] C.-H. Chen, F. Song, F.-J. Hwang, and L. Wu, "A probability density function generator based on neural networks," *Physica A: Statistical Mechanics and Its Applications*, vol. 541, Article ID 123344, 2020.

[24] C. Gini, "Variabilità e mutabilità,," in *Reprinted in Memorie di Metodologica Statistica*, E. Pizetti and T. Salvemini, Eds., Libreria Eredi Virgilio Veschi, Rome, 2005.

[25] R. R. Yager and F. Petry, "An intelligent quality-based approach to fusing multi-source probabilistic information," *Information Fusion*, vol. 31, pp. 127–136, 2016.

[26] R. R. Yager and F. E. Petry, *Using Quality Measures in the Intelligent Fusion of Probabilistic Information," in Information Quality In Information Fusion And Decision Making*, pp. 51–77, Springer, Berlin, Germany, 2019.

[27] S. A. Bouhamed, I. K. Kallel, R. R. Yager, É. Bossé, and B. Solaiman, "An intelligent quality-based approach to fusing multi-source possibilistic information," *Information Fusion*, vol. 55, pp. 68–90, 2020.

[28] D. Li, Y. Deng, and X. Gao, "A generalized expression for information quality of basic probability assignment," *IEEE Access*, vol. 7, no. 1, pp. 174 734–174 739, 2019.

[29] D. Li and Y. Deng, "A new correlation coefficient based on generalized information quality," *IEEE Access*, vol. 7, no. 1, pp. 175 411–175 419, 2019.

[30] H. Garg and D. Rani, "New generalised Bonferroni mean aggregation operators of complex intuitionistic fuzzy information based on Archimedean t-norm and t-conorm," *Journal of Experimental & Theoretical Artificial Intelligence*, vol. 32, no. 1, pp. 81–109, 2020.

[31] X. Gao and Y. Deng, "Quantum model of mass function," *International Journal of Intelligent Systems*, vol. 35, no. 2, pp. 267–282, 2020.

[32] F. Xiao, ""Generalization of Dempster–Shafer theory: a complex mass function," *Applied Intelligence*, vol. 50, no. 10, pp. 3266–3275, 2019.
[33] Z. He, F. T. S. Chan, and W. Jiang, "A quantum framework for modelling subjectivity in multi-attribute group decision making," *Computers & Industrial Engineering*, vol. 124, pp. 560–572, 2018.
[34] H. Garg and D. Rani, "A robust correlation coefficient measure of complex intuitionistic fuzzy sets and their applications in decision-making," *Applied Intelligence*, vol. 49, no. 2, pp. 496–512, 2019.
[35] F. Xiao, "CEQD: a complex mass function to predict interference effects," *IEEE Transactions on Cybernetics*, p. 1, 2021.
[36] J. Dai and Y. Deng, "A new method to predict the interference effect in quantum-like Bayesian networks," *Soft Computing*, vol. 24, no. 14, pp. 10 287–310 294, 2020.
[37] F. Pennecchi and L. Oberto, "Uncertainty evaluation for the estimate of a complex-valued quantity modulus," *Metrologia*, vol. 47, no. 3, p. 157, 2010.
[38] X. Deng, "Analyzing the monotonicity of belief interval based uncertainty measures in belief function theory," *International Journal of Intelligent Systems*, vol. 33, no. 9, pp. 1869–1879, 2018.
[39] F. Xiao, Z. Cao, and A. Jolfaei, "A novel conflict measurement in decision making and its application in fault diagnosis," *IEEE Transactions on Fuzzy Systems*, vol. 29, no. 1, pp. 186–197, 2020.
[40] Z. Liu, Q. Pan, J. Dezert, J.-W. Han, and Y. He, "Classifier fusion with contextual reliability evaluation," *IEEE Transactions on Cybernetics*, vol. 48, no. 5, pp. 1605–1618, 2018.
[41] X. Xu, J. Zheng, J.-b. Yang, D.-l. Xu, and Y.-w. Chen, "Data classification using evidence reasoning rule," *Knowledge-Based Systems*, vol. 116, pp. 144–151, 2017.
[42] F. Xiao, "Evidence combination based on prospect theory for multi-sensor data fusion," *ISA Transactions*, vol. 106, pp. 253–261, 2020.
[43] C. Fu, W. Chang, and S. Yang, "Multiple criteria group decision making based on group satisfaction," *Information Sciences*, vol. 518, pp. 309–329, 2020.
[44] M. Zhou, X.-B. Liu, Y.-W. Chen, and J.-B. Yang, "Evidential reasoning rule for MADM with both weights and reliabilities in group decision making," *Knowledge-Based Systems*, vol. 143, pp. 142–161, 2018.
[45] C. Wang, Z. X. Tan, Y. Ye, L. Wang, K. H. Cheong, and N.-G. Xie, "A rumor spreading model based on information entropy," *Scientific Reports*, vol. 7, no. 1, pp. 1–14, 2017.
[46] H. Yan and Y. Deng, "An improved belief entropy in evidence theory," *IEEE Access*, vol. 8, no. 1, pp. 57 505–557 516, 2020.
[47] S. Babajanyan, A. Allahverdyan, and K. H. Cheong, "Energy and entropy: path from game theory to statistical mechanics," *Physical Review Research*, vol. 2, no. 4, Article ID 043055, 2020.
[48] J. Zhang, R. Liu, J. Zhang, and B. Kang, "Extension of Yager's negation of a probability distribution based on Tsallis entropy," *International Journal of Intelligent Systems*, vol. 35, no. 1, pp. 72–84, 2020.
[49] Z. Cao, C.-T. Lin, K.-L. Lai et al., "Extraction of SSVEPs-based inherent fuzzy entropy using a wearable headband EEG in migraine patients," *IEEE Transactions on Fuzzy Systems*, vol. 28, no. 1, p. 14, 2020.
[50] P. Liu, X. Zhang, and Z. Wang, "An extended VIKOR method for multiple attribute decision making with linguistic D numbers based on fuzzy entropy," *International Journal of Information Technology & Decision Making*, vol. 19, no. 01, pp. 143–167, 2020.
[51] Y. Deng, "Uncertainty measure in evidence theory," *Science China Information Sciences*, vol. 63, no. 11, p. 210201, 2020.
[52] F. Liu, X. Gao, J. Zhao, and Y. Deng, "Generalized belief entropy and its application in identifying conflict evidence," *IEEE Access*, vol. 7, no. 1, pp. 126 625–126 633, 2019.
[53] X. Gao and Y. Deng, "The pseudo-pascal triangle of maximum Deng entropy," *International Journal of Computers Communications & Control*, vol. 15, no. 1, p. 1006, 2020.
[54] Y. Song, Q. Fu, Y.-F. Wang, and X. Wang, "Divergence-based cross entropy and uncertainty measures of Atanassov's intuitionistic fuzzy sets with their application in decision making," *Applied Soft Computing*, vol. 84, Article ID 105703, 2019.
[55] Y. Deng, "Information volume of mass function," *International Journal of Computers Communications & Control*, vol. 15, no. 6, p. 3983, 2020.
[56] J. Abellan and E. Bosse, "Critique of recent uncertainty measures developed under the evidence theory and belief intervals," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 50, no. 3, pp. 1186–1192, 2020.
[57] L. Pan and Y. Deng, "Probability transform based on the ordered weighted averaging and entropy difference," *International Journal of Computers Communications & Control*, vol. 15, no. 4, p. 3743, 2020.
[58] H. Cui, Q. Liu, J. Zhang, and B. Kang, "An improved deng entropy and its application in pattern recognition," *IEEE Access*, vol. 7, pp. 18 284–18 292, 2019.
[59] Y. Deng and Y. Deng, 2021 Entropy Measure of Quantum Entanglement.
[60] F. Xiao, "GIQ: a generalized intelligent quality-based approach for fusing multi-source information," *IEEE Transactions on Fuzzy Systems*, p. 1, 2020.
[61] F. Xiao, "A distance measure for intuitionistic fuzzy sets and its application to pattern classification problems," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2019.
[62] H. Liao, Z. Ren, and R. Fang, "A Deng-entropy-based evidential reasoning approach for multi-expert multi-criterion decision-making with uncertainty," *International Journal of Computational Intelligence Systems*, vol. 13, no. 1, pp. 1281–1294, 2020.
[63] L. Fei, J. Lu, and Y. Feng, "An extended best-worst multi-criteria decision-making method by belief functions and its applications in hospital service evaluation," *Computers & Industrial Engineering*, vol. 142, Article ID 106355, 2020.
[64] H. Garg, "A new possibility degree measure for interval-valued q-rung orthopair fuzzy sets in decision-making," *International Journal of Intelligent Systems*, vol. 36, no. 1, pp. 526–557, 2021.
[65] M. Tang, H. Liao, X. Mi, X. Xu, and F. Herrera, "Dynamic subgroup-quality-based consensus in managing consistency, nearness, and evenness quality indices for large-scale group decision making under hesitant environment," *Journal of the Operational Research Society*, pp. 1–14, 2020.
[66] F. Xiao, "CED: a distance for complex mass functions," *IEEE Transactions on Neural Networks and Learning Systems*, p. 1, 2020.
[67] H. Garg and D. Rani, "Complex interval-valued intuitionistic fuzzy sets and their aggregation operators," *Fundamenta Informaticae*, vol. 164, no. 1, pp. 61–101, 2019.
[68] F. Xiao, "Generalized belief function in complex evidence theory," *Journal of Intelligent & Fuzzy Systems*, vol. 38, no. 4, pp. 3665–3673, 2020.