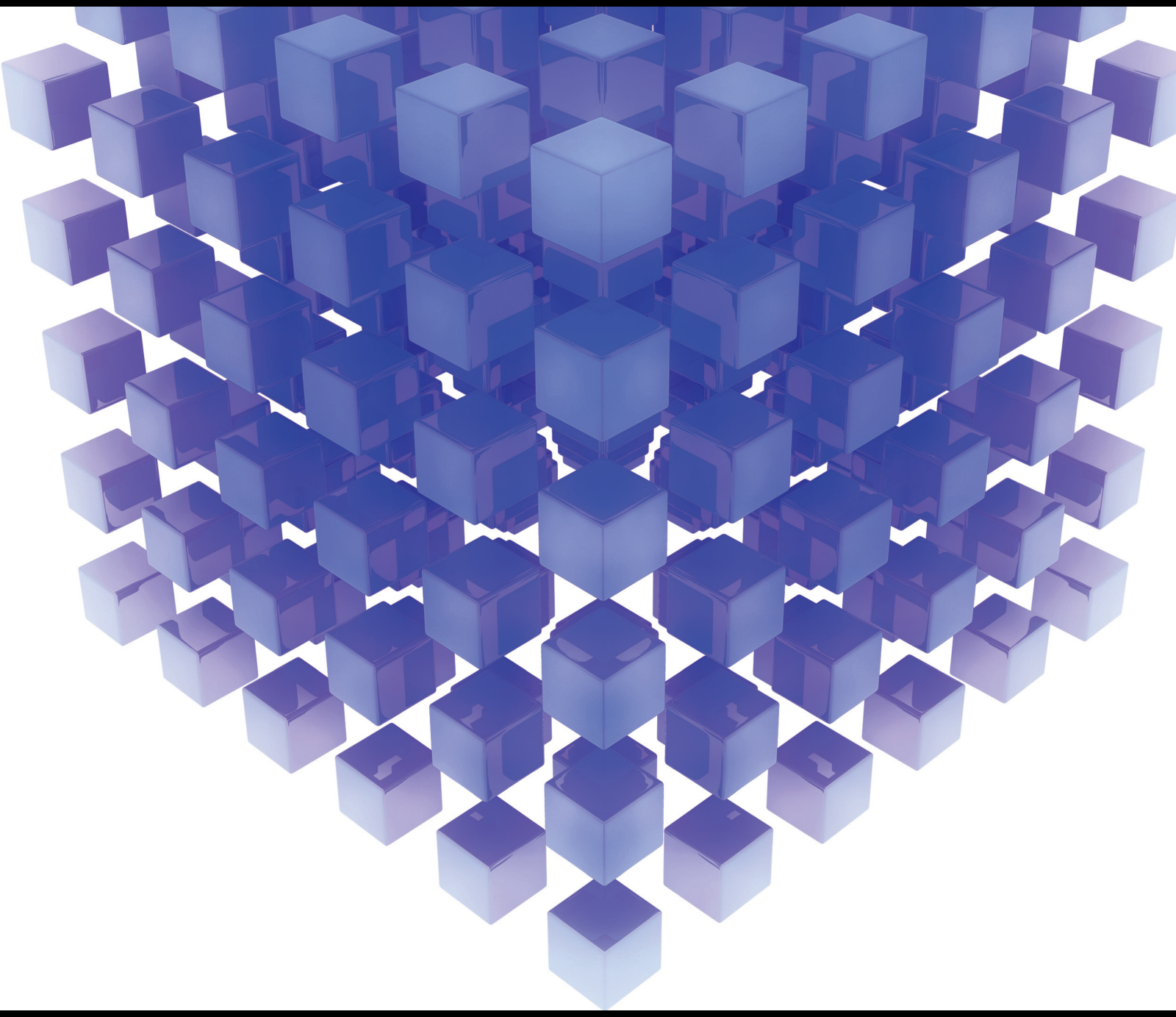
Mathematical Problems in Engineering

Data Intelligence-enabled Industry Applications

Lead Guest Editor: Long Wang
Guest Editors: Jenq-Haur Wang and Chao Huang



Data Intelligence-enabled Industry Applications

Mathematical Problems in Engineering

Data Intelligence-enabled Industry Applications

Lead Guest Editor: Long Wang
Guest Editors: Jenq-Haur Wang and Chao Huang

Contents

Hindawi
Mathematical Problems in Engineering
Volume 2023, Article ID 2919259, 8 pages
https://doi.org/10.1155/2023/2919259

Research Article

Exploring the Clinical Benefits of Mixed-Reality Technology for Breast Lumpectomy

Yongfeng Zhang,[1] Yuanyuan Lu,[1] Junlai Li,[1] Bangjun Huang,[2] Xuan He,[1] and Ruoxiu Xiao[3]

[1]*Department of Ultrasound, Chinese PLA General Hospital, Beijing 100853, China*
[2]*Outpatient Department of the Third Comprehensive Service Guarantee Center, Beijing 100039, China*
[3]*School of Computer and Communication Engineering, University of Science and Technology Beijing, Beijing 100083, China*

Correspondence should be addressed to Junlai Li; li_jl@yeah.net

Received 26 May 2022; Revised 10 October 2022; Accepted 11 October 2022; Published 23 May 2023

Academic Editor: A. M. Bastos Pereira

This study determined the value of mixed-reality (MR) technology for doctor-patient communication, preoperative planning, intraoperative navigation, and tumor localization in treating patients with breast cancer. Fifty-eight patients with breast space-occupying lesions (16 benign and 42 malignant) who underwent breast lumpectomy at the People's Liberation Army General Hospital of China were included in this study. The patients were randomly divided into the MR group and the computed tomography (CT) group. In the MR group, a 3D reconstruction of whole-breast ultrasound was used to localize the spatial position of the breast lesion and was combined with the 3D reconstruction of breast MRI to determine the lesion boundaries. To improve the preciseness of surgery, a postoperative survey was conducted. The MR group exhibited a higher level of patient knowledge regarding the disease, treatment, and diagnosis (26.207 ± 1.698 points) than the CT group (19.228 ± 4.889 points) ($T = 7.033$; $P < 0.01$), and patient satisfaction with surgical results, treatment confidence, and communication methods (4.448 ± 0.572) was also higher than that of the CT group (3.172 4.448 ± 0.572) 0.602) ($P < 0.05$). In addition, doctors were significantly more satisfied with surgical planning and intraoperative localization when the MR technique was used ($T = 8.273$; $P < 0.01$). The use of MR technology in lumpectomy has improved patients' understanding of surgical procedures and surgical results and has achieved positive results. This technique may provide clinical benefits.

1. Introduction

Breasts are important organs that are of great significance to women [1, 2]. However, breast cancer is a serious threat to women's health. According to the World Health Organization's International Agency for Research on Cancer [3], the number of new breast cancer cases worldwide reached 2.26 million in 2020, surpassing lung cancer to become the most commonly diagnosed cancer in the world. Although substantial progress has been achieved in the treatment of breast cancer, patients express significant concern regarding the disease. In general, most younger patients with breast cancer meet the requirements for breast-conserving surgery. However, due to a lack of patient knowledge regarding breast lumpectomy, plastic and reconstructive surgery, and surgical outcomes, several patients undergo radical mastectomy or total mastectomy, leading to irreversible changes [4–6]. Surgical planning based on mixed-reality (MR) technology allows for the accurate navigation of breast lesions and the visualization of the anatomical structure of the breast [7–12]. This technology allows surgeons to perform more precise operations and patients to gain a better understanding of their own condition, increasing their confidence in treatment, which can help patients in choosing surgical treatment methods [13–15].

MR technology has been applied to breast lumpectomy in the past. Amini and Kersten-Oertel [16] developed an MR system that uses Microsoft HoloLens to project a three-

dimensional (3D) hologram of images created using breast magnetic resonance imaging (MRI) onto the patient, allowing the surgeon to accurately identify tumor location during the surgical process. Gouveia et al. [12] were the first to use a digital, noninvasive method for the intraoperative localization of tumors in patients with breast cancer. Their 3D, digital breast model and augmented reality technology were combined to guide breast-conserving surgery with the use of preoperative markings using carbon tattooing. Allison et al. [17] proposed Breast3D, a fully functional mammographic X-ray image analysis system that reconstructs MRI and computed tomography data in extended reality, establishes visualization models, and is portable to different MR head-mounted displays including Magic Leap. Invernizzi et al. [18] designed a set of guidelines for the early diagnosis of breast cancer-related lymphedema. These guidelines integrate augmented reality tools with clinical examination findings for breast cancer-related lymphedema and use a 3D laser scanner and tablet to conduct a digitally assisted assessment. MR technology has gradually developed into an important research field with potential applications in clinical settings.

In this study, whole-breast ultrasound images were fused with breast MRI and applied to specific surgical cases using MR technology. This study determines the value of this clinical application and discusses its clinical uses.

2. Materials and Methods

2.1. Patients. Fifty-eight patients with space-occupying breast lesions who were treated at the Breast Surgery Department of the People's Liberation Army General Hospital between January 2021 and May 2021 were included in this study. The patient inclusion criteria were as follows: aged between 18 and 50 years, underwent surgical treatment of the breast at the study institution, had a maximum transverse nodule diameter of <50 mm, underwent preoperative MRI and ultrasound (using an automated breast volume scanner (ABVS)), communicated clearly, were fully informed (as verified by the patients' families), and provided written informed consent prior to ABVS. Male patients and those who had breast hyperplasia or cystic, inflammatory, or other nonsolid space-occupying lesions were excluded from this study. In addition, patients who did not undergo preoperative MRI and ABVS or those who underwent needle biopsy were excluded from the study. Patients who were enrolled in the study but were subsequently found to not meet the inclusion criteria or those whose actions or surgery did not meet the study protocol were rejected from the study. The mean patient age was 38.2 years (range, 25–45 years), and the transverse diameters of the nodules were 8–48 mm. Invasive ductal carcinoma was diagnosed in 38 patients (65.5%), fibroadenoma in 12 patients (20.8%), and ductal carcinoma in situ in 8 patients (13.7%). Five patients with multifocal lesions and two patients with diffuse disease were excluded from the study.

The patients were randomly divided into the conventional group and the MR group according to the natural order of visits. The conventional group received the traditional abstract oral narration and description of diagnosis and treatment (preoperative communication, preoperative planning, and intraoperative positioning). Patients in the MR group received three-dimensional reconstruction and MR technology based on ultrasound and MRI images, which were applied to preoperative communication, preoperative planning, and intraoperative navigation and positioning of patients (doctors relied on three-dimensional images to display the condition, explain the diagnosis and treatment of patients, and visualize the body surface in the three-dimensional space during surgery). The number of patients, mean age, ratio of left to right breasts, tumor size (including length, width, and depth), educational level, annual income, and marital status were not significantly different between the computed tomography (CT) and MR groups (Table 1).

This study was conducted in accordance with the Declaration of Helsinki. This study was approved by the Ethics Committee of the People's Liberation Army General Hospital. This study was registered at clinicaltrials.gov (number) on January 1, 2019.

2.2. ABVS Ultrasound System. Siemens S2000 ABVS was used to perform automated continuous tomography with a transducer length of 20 cm, a scanning depth of 20 cm, and a slice thickness of 0.525 mm. A sealed silicone sink with adjustable stands was designed and developed [19]. This sink uses the ultrasonic permeability and deformation properties of water to ensure that the unilateral breast being examined could maintain contact with the transducer attached to the ABVS ultrasonic robot arm with the help of the water surface. This ensures that the scanning range of the transducer covered the entire unilateral breast and that the breast was exposed to as little external pressure as possible. Lesion deformation due to external pressure may result in changes in tumor location and morphology. A silicone film at the bottom of the sink provided a seal, while the outer layer of the silicone film and the breast surface were coated with a coupling agent to maintain the connection. Image segmentation, reconstruction, and head-mounted displays were developed by Beijing Weizhuo Zhiyuan Technology Co., Ltd (Beijing, China).

2.3. MR-Based Lumpectomy. Tumor segmentation was performed using the patient's MRI data to create a 3D reconstruction of the tumor. Feature points were marked on the skin of the patient's breast. Relationships between the tumor and feature points were obtained using ABVS and used to construct the 3D model and its corresponding feature points. Based on the transformation matrix obtained with the feature points, point cloud registration of the 3D model was performed, and reality and virtual reality data were fused. The MR-based lumpectomy process is outlined in Figure 1.

2.4. 3D Reconstruction and Localization of the Breast Tumor. Tumor segmentation was performed using the random walker algorithm and preoperative breast MRI data. The random walker graph was established using the imaging data

Table 1: Patient characteristics.

	CT group ($n = 29$)	MR group ($n = 29$)	T value	P value
Age (years)	41.046 ± 6.835	38.863 ± 6.081	0.795	0.456
Left breast/right breast	11/11	13/9	0.487	0.784
Tumor length (mm)	18.514 ± 6.546	20.514 ± 8.160	0.578	0.564
Tumor width (mm)	15.268 ± 4.356	16.786 ± 4.546	0.865	0.426
Tumor depth (mm)	15.641 ± 4.611	18.741 ± 6.475	2.583	0.084
Educational level (years)	13.000 ± 2.794	12.863 ± 2.550	0.719	0.491
Annual income (×10,000)	12.318 ± 20.993	8.909 ± 8.372	0.393	0.677
Marital status (married/unmarried/divorced)	18/1/3	16/3/3	1.118	0.891

Data are presented as the mean ± standard deviation or number. CT, conventional treatment; MR, mixed reality.

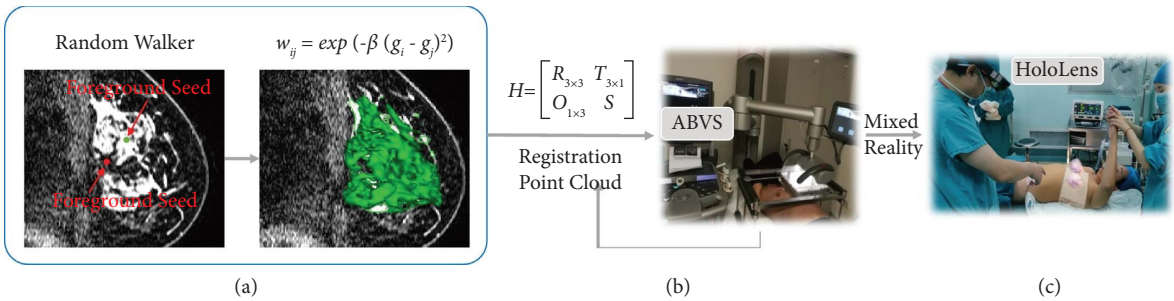


Figure 1: The process of MR-based lumpectomy: (a) the process of a random walker; (b) the registration process of the point cloud; (c) the operation process of mixed reality through HoloLens glasses.

and defined as $G = (V, E)$, where V is the set of vertices in the graph $v \in V$ and E is the undirected edge set of vertices in the graph $e \in E \subseteq V \times V$. e_{ij} is the connection relation between vertices v_i and v_j. The edge weight reflects the degree of similarity between adjacent pixels. As tumors exhibit significant contrast, the edge weights were defined using the Gaussian weighting function based on the grey value [20] given by

$$w_{ij} = \exp\left(-\beta\left(g_i - g_j\right)^2\right), \tag{1}$$

where g_i is the grey value at the vertex v_i and is adjusted by the β parameter.

Using the edge weight, the probability of walking was obtained from unlabeled pixels to each seed. Each pixel was then reassigned with the label to which it had the highest probability of arriving to achieve image segmentation. As previously described [21], the process of solving these probabilities was transformed into a Dirichlet problem:

$$D[x] = \frac{1}{2}(Ax)^T C(Ax) = \frac{1}{2}\sum_{e_{ij}\in E} w_{ij}\left(x_i - x_j\right)^2, \tag{2}$$

where A is the incidence matrix of the edges and vertices, C is the diagonal matrix, and the diagonal elements are the corresponding edge weights. The discrete harmonic function x that minimizes $D[x]$ must be satisfied. As $A^T CA$ is a positive semidefinite matrix, $D[x]$ has a unique minimum and is decomposed as follows:

$$\begin{aligned} D[x_u] &= \frac{1}{2}\left[x_s^T x_u^T\right]\begin{bmatrix} L_M & B \\ B^T & L_U \end{bmatrix}\begin{bmatrix} x_s \\ x_u \end{bmatrix} \\ &= \frac{1}{2}\left(x_s^T L_u x_s + 2x_u^T B^T x_s + x_u^T L_u x_u\right), \end{aligned} \tag{3}$$

where x_s and x_u correspond to the probabilities of the labeled and unlabeled pixels, respectively. The differential of $D[x_u]$ with respect to x_u was solved, and the extrema was found using zero: $L_u x_u = -B^T x_s$. Thus, the solution to the Dirichlet problem was given by $L_U X = -B^T M$, where the sum of all probabilities satisfying any vertex was one, such as $\sum_s x_i^s = 1, \forall v_i \in V$.

In addition, tumor localization was performed using the Siemens S2000 AVBS breast transducer. The transducer scanning range was set to 20 × 20 cm, which includes most unilateral breast sizes in a normal human body. The sealed silicone sink described above was used to account for the mismatch between the flat surface of the ultrasound transducer, the convex surface of the breast, the deformation of the breasts, and certain space-occupying tissues due to the

pressure applied by using the ABVS transducer during examination. This allowed for level contact with the transducer attached to the ABVS ultrasonic robot arm and ensured that the scanning range of the transducer covered the entire unilateral breast while minimizing the changes in tumor location and morphology caused by external pressure.

2.5. Point Cloud Registration. The 3D reconstruction model and breast tumor registration are important prerequisites for MR. In this study, the point cloud method was used for the 3D reconstruction model and a coordinate system conversion was implemented through the construction of the H conversion matrix. In general, after ABVS-based tumor localization, only rotations and translations existed between point cloud data points. Therefore, the H matrix was defined as follows:

$$H = \begin{bmatrix} R_{3\times3} & T_{3\times1} \\ O_{1\times3} & S \end{bmatrix}, \tag{4}$$

where scale factor $S = 1$; $R_{3\times3}$ was the rotation matrix, which was defined as

$$\begin{aligned} R_{3\times3} &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & \sin\alpha \\ 0 & -\sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} \cos\beta & 0 & -\sin\beta \\ 0 & 1 & 0 \\ \sin\beta & 0 & \cos\beta \end{bmatrix} \begin{bmatrix} \cos\gamma & \sin\gamma & 0 \\ -\sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} \cos\beta\cos\gamma & \cos\beta\sin\gamma & -\sin\beta \\ -\cos\alpha\sin\gamma - \sin\alpha\sin\beta\cos\gamma & \cos\alpha\cos\gamma + \sin\alpha\sin\beta\sin\gamma & \sin\alpha\cos\beta \\ \sin\alpha\sin\gamma + \cos\alpha\sin\beta\cos\gamma & -\sin\alpha\cos\gamma - \cos\alpha\sin\beta\sin\gamma & \cos\alpha\cos\beta \end{bmatrix}. \end{aligned} \tag{5}$$

The translation matrix $T_{3\times1}$ was given by

$$T_{3\times1} = \begin{bmatrix} t_x & t_y & t_z \end{bmatrix}^T. \tag{6}$$

Therefore, X was defined as the spatial coordinates of the 3D model, and the transformed matrix X' must satisfy the following equation:

$$X' = R_{3\times3}X + T_{3\times1}. \tag{7}$$

After eight groups of feature points were marked on the patient's skin, ABVS was used to measure the patient's tumor location. The corresponding eight groups of matching points were constructed on the 3D model. The translation and rotation parameters of the H matrix were obtained through the transformation of the matching points. The H matrix was then used to fully convert the point cloud of the model to the patient's physical signs.

2.6. Outcomes. The main outcomes of this study were doctor-patient communication, preoperative planning and discussion, and localization of the tumor using MR technology. Doctor-patient communication was assessed using a postoperative survey regarding communication time (time used by doctors to explain individual patient cases and answer patient questions), patient knowledge (knowledge of illness, treatment plans, prognosis, and potential complications), patient acceptance of surgical outcomes, patient confidence in treatment, and patient satisfaction with communication methods. Preoperative planning and discussion were assessed using postoperative survey items regarding the preoperative planning discussion (specific location and length of the surgical incision) and the doctor's satisfaction with the preoperative planning.

In the MR group, tumor localization was achieved using markers. Localization was achieved by comparing the location of the center point on the actual intraoperative breast tumor image with the surface marking of the tumor central point based on 2D pre-operative ultrasound. Differences up to 1 mm were considered accurate. For all the patients, the patients had reached clinical recovery after follow-up one month after surgery, and no related complications occurred.

2.7. Statistical Analysis. All statistical analyses were performed using SPSS version 19.0 statistical software (manufacturer name and location) and GraphPad statistical graphing software (manufacturer name and location). Normally distributed continuous data are presented as the mean ± standard deviation. The t-test was used to compare these data. Statistical significance was indicated as $P < 0.05$.

The testing methods above were all completed by unchanging examining physicians at our hospital, while manual segmentation and the measurement of images were performed by designated imaging physicians and 3D reconstruction technicians.

In order to evaluate the accuracy of registration and fusion of the MR system, we took ten reconstructed breast models for 3D printing and selected five marker points on each printed model. The marker point on the nipple, M0, was the center of the circle, and four marker points, M1, M2, M3, and M4, were placed on the arc every 90°, with an equal spacing of 3 cm, as shown in Figure 2. Since the fusion error of the tumor could not be measured directly, we used the fusion error of these five marker points for systematic evaluation. The five marker points were used for point cloud registration with the outer surface of the reconstructed breast model. After registration, the minimum distance

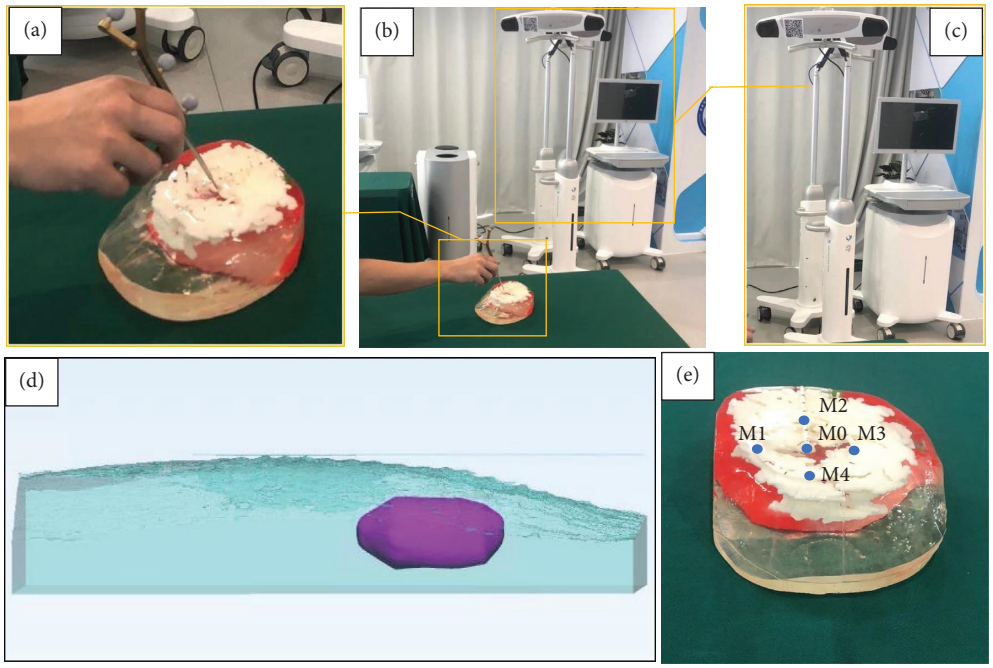


Figure 2: Evaluation process of the fusion error using the breast 3D printing model and the optical tracking and positioning system: (a) selection of the marker points on the breast 3D printing model with the optical tracking and positioning system; (b) overall appearance of the evaluation system; (c) the optical tracking and positioning system; (d) reconstructed breast 3D surface and tumor; (e) five marker points on the breast 3D printing model, the marker point M0 was selected on the nipple, and other four marker points, M1, M2, M3, and M4, are placed on the arc every 90 with an equal spacing of 3 cm.

Table 2: Fusion error measured on selected 5 marker points on the 10 tested patients using 3D printing breast models (unit: mm).

	Error of marker point 1	Error of marker point 2	Error of marker point 3	Error of marker point 4	Error of marker point 5
Data1	1.45	1.83	1.96	0.26	0.61
Data2	0.90	0.82	0.42	1.19	3.34
Data3	0.13	0.87	0.39	0.33	2.43
Data4	0.80	0.40	1.47	1.61	1.06
Data5	0.37	1.17	1.67	0.95	0.96
Data6	0.29	0.83	1.53	1.05	0.95
Data7	1.13	1.66	0.82	0.49	3.08
Data8	1.83	1.57	0.86	1.05	1.94
Data9	1.08	0.55	0.27	1.00	0.33
Data10	1.81	0.92	1.35	1.38	1.05
Average	0.98	1.06	1.07	0.93	1.57
SD	0.60	0.48	0.59	0.44	1.05

SD, standard deviation.

between the marker points and the reconstructed skin was measured as the fusion error. Table 2 shows the obtained fusion error results of the five marker points through 3D printing, registration, and fusion. The average fusion error of our proposed method is between 0.93 mm and 1.57 mm, and the standard deviation (SD) is between 0.44 mm and 1.05 mm.

3. Results

3.1. Doctor-Patient Communication. The time used by doctors to explain individual patient cases was not significantly different between the groups (survey scores: CT group, 3.724 ± 1.437 vs. MR group, 4.241 ± 0.786; $T = -1.701$; $P = 0.096$).

The time used by doctors to answer patient questions was rated as significantly higher in the MR group (survey score: 4.379 ± 1.374) than in the CT group (survey score: 2.483 ± 1.122) ($T = -5.759$; $P < 0.01$) (Table 3).

Similarly, the time used by doctors to respond to patient questions was significantly longer in the MR group (survey score: 4.448 ± 1.429) than in the CT group (survey score: 2.103 ± 0.976) ($T = -7.296$; $P < 0.01$) (Table 3). The total communication time was significantly higher in the MR group (survey score: 12.069 ± 2.103) than in the CT group (survey score: 8.379 ± 2.499) ($T = -6.084$; $P < 0.01$).

Table 3: Doctor-patient communication scores.

		CT group ($n = 22$)	MR group ($n = 22$)	T value	P value
Communication time (min)	Doctor interpretation of the individual case	3.724 ± 1.437	4.241 ± 0.786	−1.701	0.096
	Patient questions	2.483 ± 1.122	4.379 ± 1.374	−5.759	<0.01
	Doctor response	2.103 ± 0.976	4.448 ± 1.429	−7.296	<0.01
	Total time	8.379 ± 2.499	12.069 ± 2.103	−6.084	<0.01
Knowledge level	Disease	3.448 ± 0.827	4.517 ± 0.575	−5.714	<0.01
	Treatment	3.172 ± 0.693	4.724 ± 0.455	−6.623	<0.01
	Prognosis	3.000 ± 0.02	4.448 ± 0.572	−7.917	<0.01
Satisfaction with communication	Surgical outcomes	3.207 ± 0.819	4.276 ± 0.797	−5.038	<0.01
	Treatment confidence	2.759 ± 0.739	3.862 ± 0.789	−5.494	<0.01
	Communication methods	3.138 ± 0.693	4.724 ± 0.455	−10.34	<0.01
Total score		19.448 ± 4.889	26.207 ± 1.698	−7.033	<0.01

Data are presented as the mean ± standard deviation. CT, conventional treatment; MR, mixed reality.

Table 4: Preoperative planning scores.

		CT group ($n = 22$)	MR group ($n = 22$)	T value	P value
Depth of preoperative planning	Incision location	2.607 ± 0.875	4.551 ± 0.572	−9.891	<0.01
	Incision length	3.655 ± 0.669	4.276 ± 0.591	−3.742	<0.01
Doctor's satisfaction		3.172 ± 0.602	4.448 ± 0.572	−8.273	<0.01

Data are presented as the mean ± standard deviation. CT, conventional treatment; MR, mixed reality.

Patients in the MR group exhibited significantly more knowledge regarding their illness (MR group, 4.517 ± 0.575 vs. CT group, 3.448 ± 0.827; $T = -5.714$; $P < 0.01$), the treatment plan (MR group, 4.724 ± 0.455 vs. CT group, 3.172 ± 0.693; $T = -6.623$; $P = < 0.01$), and prognosis and subsequent treatments (MR group, 4.448 ± 0.572 vs. CT group, 3.000 ± 0.02; $T = -7.917$; $P < 0.01$). The reported satisfaction with the surgical outcomes was not significantly different between the two groups (MR group, 4.276 ± 0.797 vs. CT group, 3.207 ± 0.819; $T = -5.038$; $P < 0.01$). Patients in the MR group had significantly higher postoperative treatment confidence (survey score: 3.862 ± 0.789) than patients in the CT group (survey score: 2.759 ± 0.739) ($T = -5.494$; $P < 0.01$) (Table 3). Patients in the MR group reported a higher level of satisfaction with communication methods (survey score: 4.724 ± 0.455) than patients in the CT group (survey score: 3.138 ± 0.693) ($T = -10.34$; $P < 0.01$).

The total patient knowledge and satisfaction score was significantly higher in the MR group (26.207 ± 1.698) than in the CT group (19.448 ± 4.889) ($T = -7.033$; $P < 0.01$).

3.2. Preoperative Planning. The actual location of the surgical incision was significantly closer to the location determined in the preoperative plan in the MR group (survey score: 4.551 ± 0.572) than in the CT group (survey score: 2.607 ± 0.875) ($T = -9.891$; $P < 0.01$) (Table 4). The actual length of the surgical incision was significantly closer to the length determined in the preoperative plan in the MR group (survey score: 4.276 ± 0.591) than in the CT group (survey score: 3.655 ± 0.669) ($T = -3.742$; $P < 0.01$).

The doctors' overall satisfaction with preoperative planning was significantly higher when MR was used (survey score: 4.448 ± 0.572)than when CT was used (3.172 ± 0.602) ($T = -8.273$; $P < 0.01$).

4. Discussion

The use of MR technology during breast lumpectomy surgery was explored in this study. The MR system improved preoperative doctor-patient communication, surgical planning, and image visualization for patients with breast cancer compared to CT.

The results of this study indicate that more time was devoted to doctor-patient communication in the MR group. Although a certain amount of time is added, this time is relatively fixed and limited. In addition, the MR method resulted in a higher level of patient knowledge, increased patient confidence in the treatment, and enhanced patient understanding of future treatments and complications, all of which are important factors for patients with breast cancer. The resection of diseased tissue during breast cancer surgery does not equate to the end of treatment or patient suffering; it may be interpreted by the patient as the first of a series of painful treatments and suffering, as pathological and psychological rehabilitation are needed after surgery. Therefore, it is necessary to increase awareness regarding the disease, which will enable patients to determine the optimal treatment plan for their individual conditions and needs based on a complete understanding of the relationship between the disease and the treatment method. Unlike European, American, and other Western countries, the rate of breast-conserving surgery for women in some countries is relatively low, although it has increased recently. Currently, breast cancer patients rely on doctors' advice when deciding a surgical method to undergo. However, the patient's subjective intentions can only be achieved when patients are able to make rational and autonomous decisions regarding surgical procedures. This study highlights the need to improve patient disease knowledge in the clinical setting and enhance the expression of the patient's subjective goals of care.

The doctors' survey responses indicate that the three dimensionality of the anatomical structure and the visualization of spatial properties provided by MR positively impacted preoperative planning. As the 3D reconstruction was based on ultrasound images, the virtual 3D image could only display unilateral breast structures. While satisfactory intraoperative navigation was achieved regarding tumor location due to high agreement with the anatomical body position, the resulting visual experience lacked images of the contralateral breast tissues. Therefore, the 3D construction of breast MRI images was fused with ultrasound images to improve patient understanding of the visual images and the intraoperative accuracy for the localization of the tumor. Based on the results of this study, this technique allowed for the maximum presentation of available imaging data.

This study is not without limitations. Positional deviations due to the HoloLens equipment may have occurred. Positional deviation refers to the deviation of measurements of stationary objects when viewed from different angles using HoloLens. This deviation is approximately 2 mm, which limits the precision that can be achieved during surgery. However, a positional deviation of 2 mm is within the acceptable range for breast surgery.

Although this technique is currently being developed, it may revolutionize surgical methods. Appropriate standards and quality measures must be developed to evaluate the application of this technique and to enhance its clinical value.

Data Availability

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

Conflicts of Interest

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

Authors' Contributions

Yongfeng Zhang and Yuanyuan Lu contributed equally to this work.

Acknowledgments

This work was supported in part by grants from the National Natural Science Foundation of China (81771835), the Non-Profit Central Research Institute Fund of Chinese Academy of Medical Sciences (2020-JKCS-008), and the Fundamental Research Funds for the Central Universities (FRF-DF-20-05).

References

[1] T. Mertzanidou, J. H. Hipwell, S. Reis et al., "3D volume reconstruction from serial breast specimen radiographs for mapping between histology and 3D whole specimen imaging," *Medical Physics*, vol. 44, no. 3, pp. 935–948, 2017.

[2] W. Y. Lee, M. J. Kim, D. H. Lew, S. Y. Song, and D. W. Lee, "Three-dimensional surface imaging is an effective tool for measuring breast volume: a validation study," *Arch Plast Surg*, vol. 43, no. 05, pp. 430–437, 2016.

[3] International Agency for Research on Cancer, *Latest Global Cacer Data: Cancer Burden Rises to 19.3 Million New Cases and 10.0 Million Cancer Deaths in 2020*, International Agency for Research on Cancer, Lyon, France, 2020, https://www.iarc.fr/faq/latest-global-cancer-data-2020-qa/.

[4] S. Mehta, N. Byrne, N. Karunanithy, and J. Farhadi, "3D printing provides unrivalled bespoke teaching tools for autologous free flap breast reconstruction," *Journal of Plastic, Reconstructive & Aesthetic Surgery*, vol. 69, no. 4, pp. 578–580, 2016.

[5] H. Kim, G. H. Mun, E. S. Wiraatmadja et al., "Preoperative magnetic resonance imaging-based breast volumetry for immediate breast reconstruction," *Aesthetic Plastic Surgery*, vol. 39, no. 3, pp. 369–376, 2015.

[6] M. Mahan, R. F. Spetzler, and P. Nakaji, "Electromagnetic stereotactic navigation for external ventricular drain placement in the intensive care unit," *Journal of Clinical Neuroscience*, vol. 20, no. 12, pp. 1718–1722, 2013.

[7] T. Sielhorst, M. Feuerstein, and N. Navab, "Advanced medical displays: a literature review of augmented reality," *Journal of Display Technology*, vol. 4, pp. 451–467, 2008.

[8] A. Fedorov, R. Beichel, J. Kalpathy-Cramer et al., "3D Slicer as an image computing platform for the quantitative imaging network," *Magnetic Resonance Imaging*, vol. 30, no. 9, pp. 1323–1341, 2012.

[9] C. Karmonik, T. B. Boone, and R. Khavari, "Workflow for visualization of neuroimaging data with an augmented reality device," *Journal of Digital Imaging*, vol. 31, no. 1, pp. 26–31, 2018.

[10] I. E. Sutherland, "A head-mounted three dimensional display," in *Proceedings of the December 9–11 1968, Fall Joint Computer Conference*, ACM, San Francisco California, December 1968.

[11] Y. Sato, M. Nakamoto, Y. Tamaki et al., "Image guidance of breast cancer surgery using 3-D ultrasound images and augmented reality visualization," *IEEE Transactions on Medical Imaging*, vol. 17, no. 5, pp. 681–693, 1998.

[12] P. Gouveia, J. Costa, P. Morgado et al., "Breast cancer surgery with augmented reality," *The Breast*, vol. 56, pp. 14–17, 2021.

[13] E. Y. Rha, I. K. Choi, and G. Yoo, "Accuracy of the method for estimating breast volume on three-dimensional simulated magnetic resonance imaging scans in breast reconstruction," *Plastic and Reconstructive Surgery*, vol. 133, no. 1, pp. 14–20, 2014.

[14] M. P. Chae, D. J. Hunter-Smith, R. T. Spychal, and W. M. Rozen, "3D volumetric analysis for planning breast reconstructive surgery," *Breast Cancer Research and Treatment*, vol. 146, no. 2, pp. 457–460, 2014.

[15] Y. Özbek and W. Freysinger, "Stereo Augmented Reality in einem navigierten Operationsmikroskop (AROSCOPE)," in *Proceedings of the 13th Annual Conference of the German Society for Computer- and Robotic Assisted Surgery (CURAC)*, September 2014.

[16] S. Amini and M. Kersten-Oertel, "Kersten-Oertel M Augmented reality mastectomy surgical planning prototype using the HoloLens template for healthcare technology letters," *Healthcare Technology Letters*, vol. 6, no. 6, pp. 261–265, 2019.

[17] B. Allison, X. Ye, and F. Janan, "Breast3D: an augmented reality system for breast CT and MRI," in *Proceedings of the IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)*, December 2020.

[18] M. Invernizzi, L. Runza, A. D. Sire et al., "Integrating augmented reality tools in breast cancer related lymphedema prognostication and diagnosis," *Journal of Visualized Experiments: JoVE*, vol. 6, 2020.

[19] Y. Lu, J. Li, X. Zhao, J. Li, J. Feng, and E. Fan, "Breast cancer research and treatment reconstruction of unilateral breast structure using three-dimensional ultrasound imaging to assess breast neoplasm," *Breast Cancer Research and Treatment*, vol. 176, no. 1, pp. 87–94, 2019.

[20] L. Grady and G. Funka-Lea, *Multi-label Image Segmentation for Medical Applications Based on Graph-Theoretic Electrical Potentials*, Springer, Berlin, Germany, 2010.

[21] C. Chen, R. Xiao, T. Zhang et al., "Pathological lung segmentation in chest CT images based on improved random walker," *Computer Methods and Programs in Biomedicine*, vol. 200, Article ID 105864, 2021.

Hindawi
Mathematical Problems in Engineering
Volume 2022, Article ID 2402989, 24 pages
https://doi.org/10.1155/2022/2402989

Research Article

Sensitivity Analysis of Stationarity Tests' Outcome to Time Series Facets and Test Parameters

Advait Amol Bawdekar,[1] **B. Rajanarayan Prusty**,[2] **and Kishore Bingi**[3]

[1]*School of Mechanical Engineering, Vellore Institute of Technology, Vellore, India*
[2]*Department of Electrical and Electronics Engineering, Alliance University, Bengaluru, India*
[3]*School of Electrical Engineering, Vellore Institute of Technology, Vellore, India*

Correspondence should be addressed to B. Rajanarayan Prusty; b.r.prusty@ieee.org

Received 22 July 2022; Accepted 27 September 2022; Published 7 October 2022

Academic Editor: Chao Huang

Time series stationarity is vital for the effective implementation of forecasting models. Time series of renewable generation rich power system input variables such as photovoltaic generations, wind power generations, load power, and ambient temperature have inherent time series facets such as trend, seasonality, and volatility. These inherent facets, combined with the length of the time series or time series clustering, have a propensity to bias the stationarity tests' outcome. This research conducts a rigorous comparative analysis to assess the tests' sensitivity to different time series facets. A seasonal, nonstationary load power time series and its derived time series with synthetically embedded trend and volatility effects are used to help the study capture tests' sensitivity to the above time series facets. This comprehensive analysis and discussion, via a set of well-delineated figures and tables, are expected to assist novice researchers in choosing a group of suitable tests for checking time series stationarity.

1. Introduction

Modern-day power systems are encouraged to accommodate renewable generations' high penetrations at transmission and distribution levels [1–4]. The renewable generations have major characterizing features such as (i) uncertainty (unexpected change), (ii) intermittency (unplanned unavailability), and (iii) uncontrollability (power output is not under the control of the system management) [2, 3]. The above features, alongside load variability, necessitate new methodologies to analyze system flexibility to accommodate higher renewable penetrations [5]. Developing prediction models for the above power system input variables is helpful to most power system studies [3]. On this note, stationarity is vital as many applicable forecasting models rely on stationary time series for easy modeling and obtaining accurate results [3, 6]. For a stationary time series, statistical properties do not change over time. The raw time series collected from sources are often nonstationary and methods such as differencing [7] and transformation [8], help yield a stationary time series. Stationarity tests check whether the application of these methods has succeeded in producing a stationary time series. ADF, KPSS, PP, Breitung, MK, Levene's, KW, KS, and SW tests are a few well-established stationarity tests used in the literature [9–16]. The outcome of these tests is solely based on the time series facets, such as trend, seasonality, and volatility effects, as well as the test's computational steps. Data analysts are always interested in a more effective test or a suitable combination of tests for a given application. Hence, a deliberate comparison study of the well-established tests is the need of the hour; therefore, it is carried out in this paper, taking into account various pertinent issues while applying to different time series.

The above tests are well-documented in the literature. These tests are often used to check data stationarity in power system analyses with renewable generations. ADF test was the first test establishing the unit root concept. The presence of a unit root can be detected using unit root tests. Because of its higher reliability, the ADF test is of choice in the applications such as solar irradiance/PV power [17–20] and

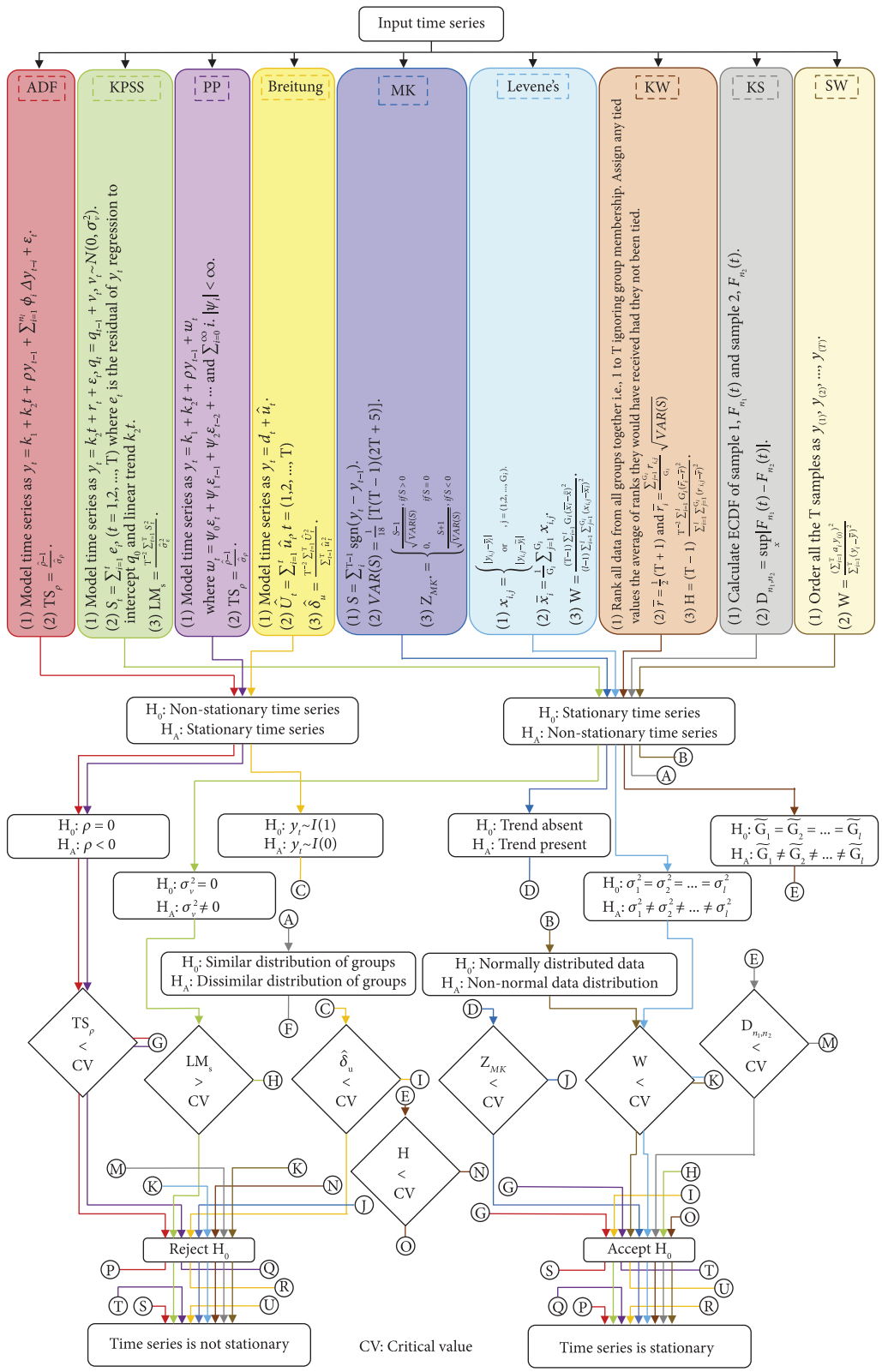


Figure 1: Pictorial representation of the thorough execution steps of well-established stationarity tests.

wind speed/wind power [21–23] time series to check data stationarity. The unit root concept was further used in the PP test, attempting to curb the disadvantages of the ADF test. The main distinction between the ADF and PP tests is how the serial correlation of error terms [10] is handled. The pairs of ADF and PP tests are commented on concrete data stationarity decisions in [24, 25]. In addition to this pair, various normality tests such as Jarque Bera, Liliefors,

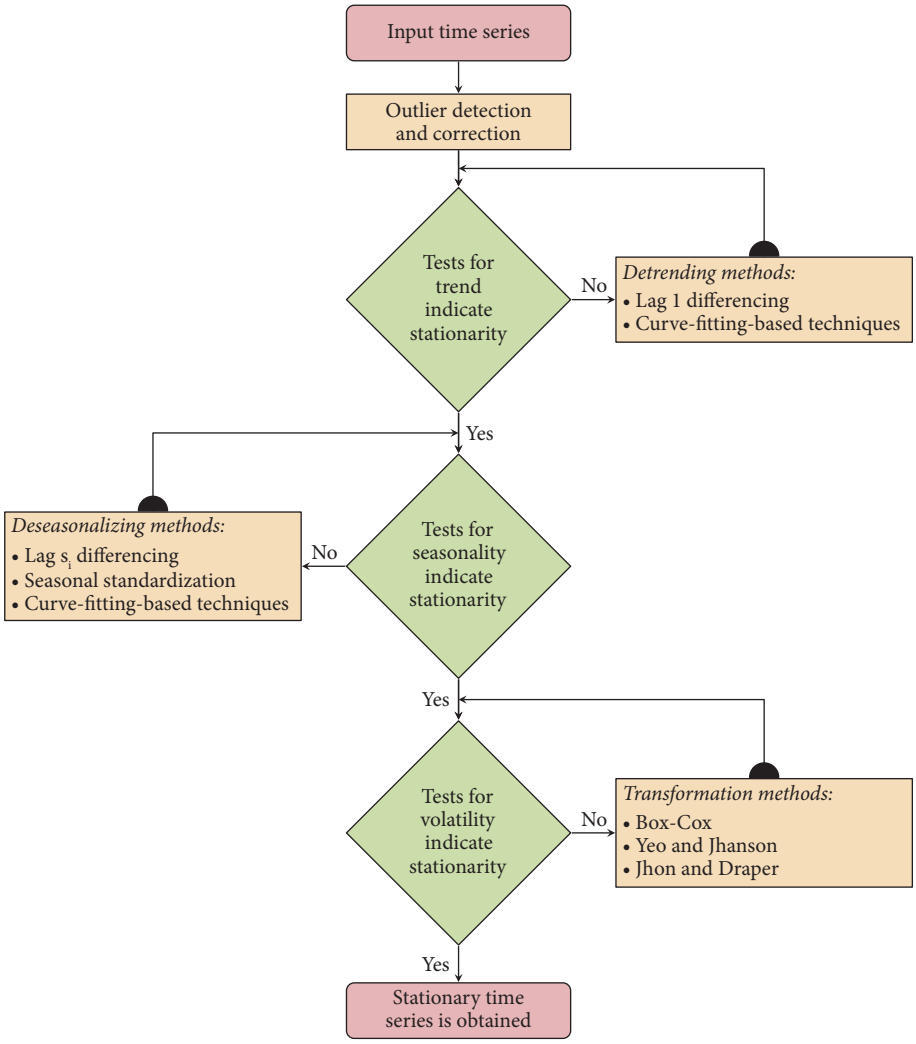


Figure 2: The flow diagram for effective stationarization of time series.

one-way KS, and SW tests were employed to assess stationarity by checking the normality of the time series [11]. The SW test is the most reliable of these tests as it has the highest power [26]. Levene's test was utilized to check the equality of variances to ensure time-series stationarity after using unit root tests such as ADF, PP, and DF-GLS tests [12]. Besides, the nonparametric tests, KPSS, and Breitung tests also assess a time series by determining the unit root. To further support the test results of the ADF, PP, or ADF-PP duo, the KPSS test (which checks the existence of stationarity around a deterministic trend) is readily chosen [3, 27–29]. MK test is applied along with the ADF-KPSS pair to analyze trend effects effectively. Besides, various normality tests discussed above are also chosen to confirm the effective stationarity of time series [14]. The KW test was used after applying ADF, PP, and KPSS tests in [30] to assess the seasonal behavior of the time series. Similarly, one-way ANOVA-based tests and the KPSS test are suggested to evaluate seasonal effects [13]. Breitung test [31] was also used to test data stationarity in wind power applications. Lastly, a two-way KS test is proposed to analyze time series stationarity, highlighting the limitations of ADF and KPSS tests [15]. Furthermore, a comparison study of these tests is also made, considering some of the crucial aspects [10]. The important information about the unit root tests was noted. The effect of time series length on the tests is analyzed in [9] with the help of the power of a test. The reliability of all the nine tests discussed in this paper is analyzed by calculating power [16]. Authors in [32, 33] compared and noted the behavior of specific stationarity tests. Further, the drawbacks of unit root tests were highlighted in [34].

Although the computational steps for tests and their characteristics are well documented in the literature, the question arises of which test or combination of tests to use and when. Further, a comprehensive comparison among the tests by examining their sensitivity towards various time series facets and the effect of other critical factors can better investigate the selection of particular tests for a specific application. Although a systematic summarization and comparison of unit root tests are made in [10], the impacts of time series length/time series clustering, parameters associated with various tests, and time series

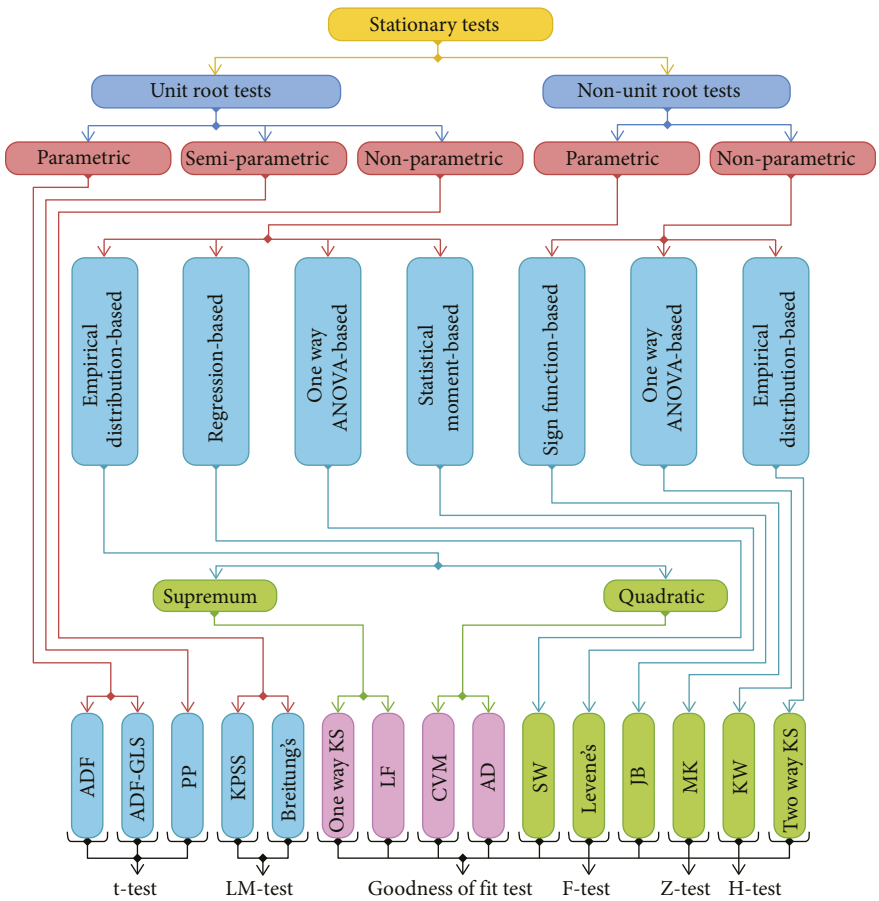


Figure 3: A novel classification of well-established stationarity tests.

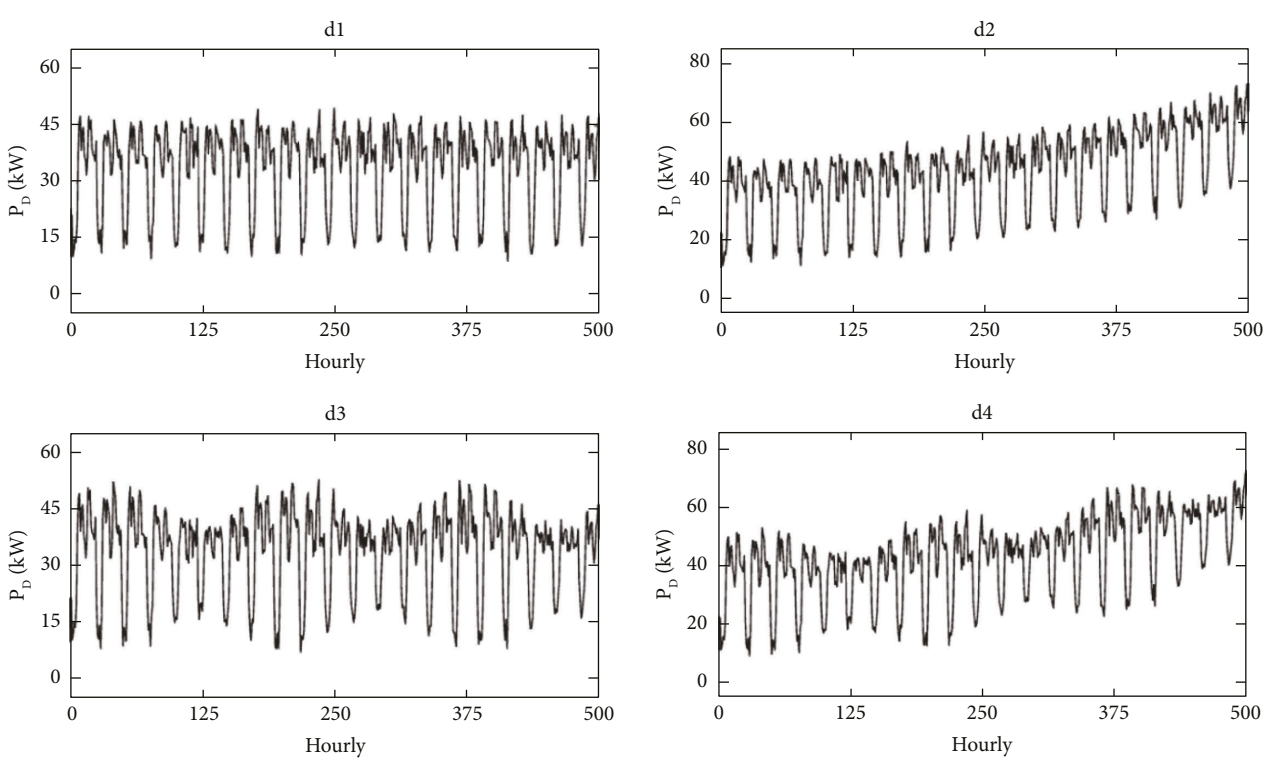


Figure 4: Time series of four synthetic datasets used for simulation.

Table 1: Datasets used for result analyses.

Type	Notation		Characteristics	Sample size	Time step
Load power, P_D (clean)	d1	Synthetic data	—	500	Hourly
	d2		Seasonal, non-linear trend	500	Hourly
	d3		Seasonal, volatile	500	Hourly
	d4		Seasonal, non-linear trend, volatile	500	Hourly

Table 2: Impact of drift and trend component consideration and time series facets on test results and critical values of ADF test.

Model equation (for ADF)	d1		d2		d3		d4	
	Test result	Critical value	Test result	Critical value	Test result	Critical value	Test result	Critical value
Without drift and trend components	−0.6790	−1.9412	−0.0399	−1.9412	−0.7335	−1.9412	−0.0200	−1.9412
With only drift component	−16.7321	−2.8683	−2.6612	−2.8683	−13.7298	−2.8683	−2.7889	−2.8683
With both drift and trend components	−16.7263	−3.4199	−10.6335	−3.4199	−13.7848	−3.4199	−9.8550	−3.4199

Table 3: Impact of trend component consideration and time series facets on test results and critical values of KPSS test.

Model equation (for KPSS)	d1		d2		d3		d4	
	Test result	Critical value	Test result	Critical value	Test result	Critical value	Test result	Critical value
With only drift component	0.0578	0.4630	15.9197	0.4630	0.0314	0.4630	14.8379	0.4630
With both drift and trend components	0.0242	0.1460	0.2979	0.1460	0.0293	0.1460	0.2922	0.1460

Table 4: Impact of drift and trend component consideration and time series facets on test results and critical values of PP test.

Model equation (for PP)	d1		d2		d3		d4	
	Test result	Critical value	Test result	Critical value	Test result	Critical value	Test result	Critical value
Without drift and trend components	−2.1086	−1.9411	−1.5034	−1.9411	−2.1505	−1.9411	−1.5446	−1.9411
With only drift component	−7.6076	−2.8680	−6.0787	−2.8680	−13.7298	−7.5627	−6.1654	−2.8680
With both drift and trend components	−7.6002	−3.4193	−7.4642	−3.4193	−13.7848	−7.5546	−7.4239	−3.4193

facets on test outcomes through a detailed analysis are overlooked. The time series have facets that alter the test results in many ways, as highlighted in [9, 16, 32, 33]. Still, only a few selective tests that belong to a specific category are analyzed and compared. Thus, a study examining these facets' impact on well-established stationarity tests becomes necessary.

The major contributions of the paper are as highlighted underneath

(i) Firstly, the working of the existing tests is pictorially portrayed in this paper, assisting readers in figuring out how different the approach of a particular test is from others.

(ii) Further, a stationarity test's attributes are highlighted, and the impacts of time series length/clustering and test parameters alongside time series facets on the test outcomes are examined.

(iii) The suitability of a test concerning these effects is also highlighted, and the capability of a stationarity test to characterize all properties of the time series and accordingly yield a fair outcome is thoroughly studied.

(iv) The appropriate incorporation of conclusions/suggestions at each analysis stage is expected to allow the readers to understand the tests' suitability for various power system applications.

The rest of the paper is organized as follows: Section 2 meticulously summarizes the nine well-established stationarity tests on a common platform. Their working steps are suitably explained, and an effective comparison is provided

Table 5: Impact of drift and trend component consideration and time series facets on test results and critical values of Breitung test.

Model equation (for breitung)	d1		d2		d3		d4	
	Test result	Critical value	Test result	Critical value	Test result	Critical value	Test result	Critical value
Without drift and trend components	0.3012	0.01986	0.2584	0.01986	0.3016	0.01986	0.2595	0.01986
With only drift component	0.0701	0.01046	0.0240	0.01046	0.0711	0.01046	0.0255	0.01046
With both drift and trend components	0.0603	0.00355	0.0230	0.00355	0.0611	0.00355	0.0244	0.00355

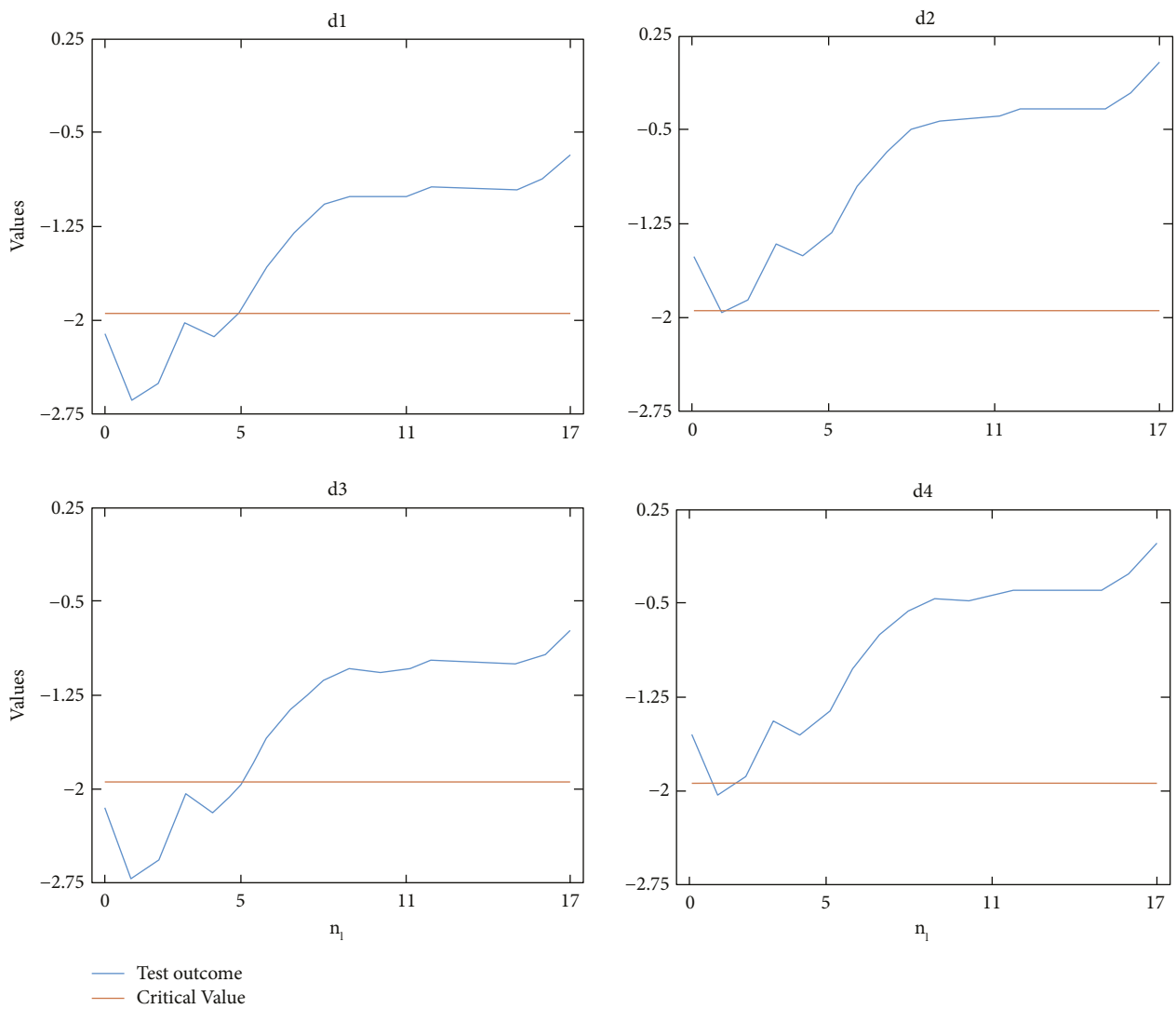


Figure 5: Impact of lag number of ADF test on test results and critical values.

for the tests. Various time series facets are highlighted, and the importance of understanding the effect of these facets on stationarity test outcomes is elaborated in Section 3. Comprehensive result analysis is carried out in Section 4 to compare the impacts of time series length/time series clustering, tests' parameters alongside the time series facets on test results to select tests' suitability for various applications astutely. Finally, the critical findings are concluded in Section 5 and provided with some future scopes.

2. Stationarity Tests

The stationarity tests analyze a time series and mark whether it is stationary through test statistics such as t-test, Z-test, LM

Table 6: Impact of seasonality consideration and time series effects on test results and critical values of MK test.

Parameter	d1		d2		d3		d4	
	Test result	Critical value	Test result	Critical value	Test result	Critical value	Test result	Critical value
Seasonality not considered	0.3480	1.9600	14.4893	1.9600	−0.6888	1.9600	13.6453	1.9600
Seasonality considered	11.7552	1.9600	19.3845	1.9600	11.7552	1.9600	19.7024	1.9600

Table 7: Impact of mean analysis and median analysis with time series effects on test results and critical values of Levene's test.

Parameter	d1		d2		d3		d4	
	Test result	Critical value	Test result	Critical value	Test result	Critical value	Test result	Critical value
Using mean	0.5836	0.0500	0.7196	0.0500	0.0152	0.0500	0.0223	0.0500
Using median	0.8327	0.0500	0.8846	0.0500	0.1307	0.0500	0.1238	0.0500

Table 8: Impact of time series effects on test results and critical values of KW, KS, and SW tests.

Test	d1		d2		d3		d4	
	Test result	Critical value	Test result	Critical value	Test result	Critical value	Test result	Critical value
KW	0.9168	0.0500	0.0000	0.0500	0.8173	0.0500	0.0000	0.0500
KS	0.5784	0.0500	0.0000	0.0500	0.1012	0.0500	0.0000	0.0500
SW	0.0000	0.0500	7.8276×10^{-8}	0.0500	1.1102×10^{-16}	0.0500	8.8698×10^{-9}	0.0500

test, and ANOVA. The test results for these tests are apportioned in various statistical distributions. For a 5% level of significance, the entire region under the distribution curve is divided into the acceptance region, constituting 95% of the whole region, and the critical region, making the rest 5%. When the result value lies in the acceptance region, the null hypothesis is accepted, whereas the null hypothesis is rejected when the test result value lies in the critical region. If the critical region lies at both ends of the curve, the test is two-tailed. When only one end contains the critical region, the test is one-tailed. The value of the test result for separating the critical region and acceptance region is called the critical value [35, 36]. The critical values mark the ends of the acceptance region, i.e., the region under which the null hypothesis is accepted. The tests are either one-tailed or two-tailed, meaning that the critical values signify one or both ends of the significance level, respectively. The working steps of nine well-established stationarity tests are portrayed in Figure 1.

The ADF test is a one-tailed t-test that checks the presence of a unit root [37]. Firstly, a model for the given time series is built (refer to Figure 1) and then tested for stationarity. The critical value of the test changes based on significance level and for cases whether k_1, k_2 are zero or nonzero. If k_1, k_2 components are nonzero, the respective components are eliminated using OLS detrending [10]. The value of n_l for the test is to be chosen according to the time series, which should not exceed a particular value given as, $n_{l,\max} = 12[T/100]^{1/4}$ [9]. The t-test here is conducted using the obtained ρ value. A t-test is an inferable statistic that evaluates the significance of the difference between the mean values of two groups and their relationship. Nextly, the KPSS test is also a one-tailed unit root test. Still, it uses LM statistics, which involves an examination of constraints on various statistical parameters using the gradient of the likelihood function. The critical value depends on the significance level based on constant and trend terms. When $r_0 = 0$ and $k_2 = 0$, e_t denotes y_t and when $r_0 = 0$ and $k_2 \neq 0$, e_t denotes residual of y_t regression to the nonzero intercept term [10]. Except for the difference in asymptotic theory and no need for a decision of lag number, the PP test is the same as the ADF test [10]. Asymptotic theory, or large sample theory, is a framework for evaluating the properties of estimators and statistical tests. The sample size is generally considered to grow indefinitely in this paradigm. The properties of estimators and tests are then examined under the limit of sample size tending to infinity. In reality, a limit analysis is thought to be roughly valid for large finite sample sizes. The Breitung test also follows a one-tailed LM statistic, and the critical values depend on significance level and length of time series [38].

The MK test is a two-tailed test capable of detecting a trend in time series [14]. This test follows Z-statistics, and the critical values depend on the significance level. A Z-test is a statistical test to assess whether the mean values are different when the variances of two populations are known. Like the t-test, the Z-test is a univariate test based on the standard normal distribution; however, the population variance in the t-test is estimated from the sample with a known mean. Further, the seasonal MK test is applied if the time series has a seasonal component. Levene's test can be employed for testing the equality of variances [12, 39]. It is a one-tailed test following the F statistics; hence, critical values are dependent on the level of significance and corresponding degrees of freedom. As F-test determines whether the samples taken come from the same population, they assess whether the statistical properties of the samples considered are significantly similar. Besides, the KW test is a one-tailed H-test

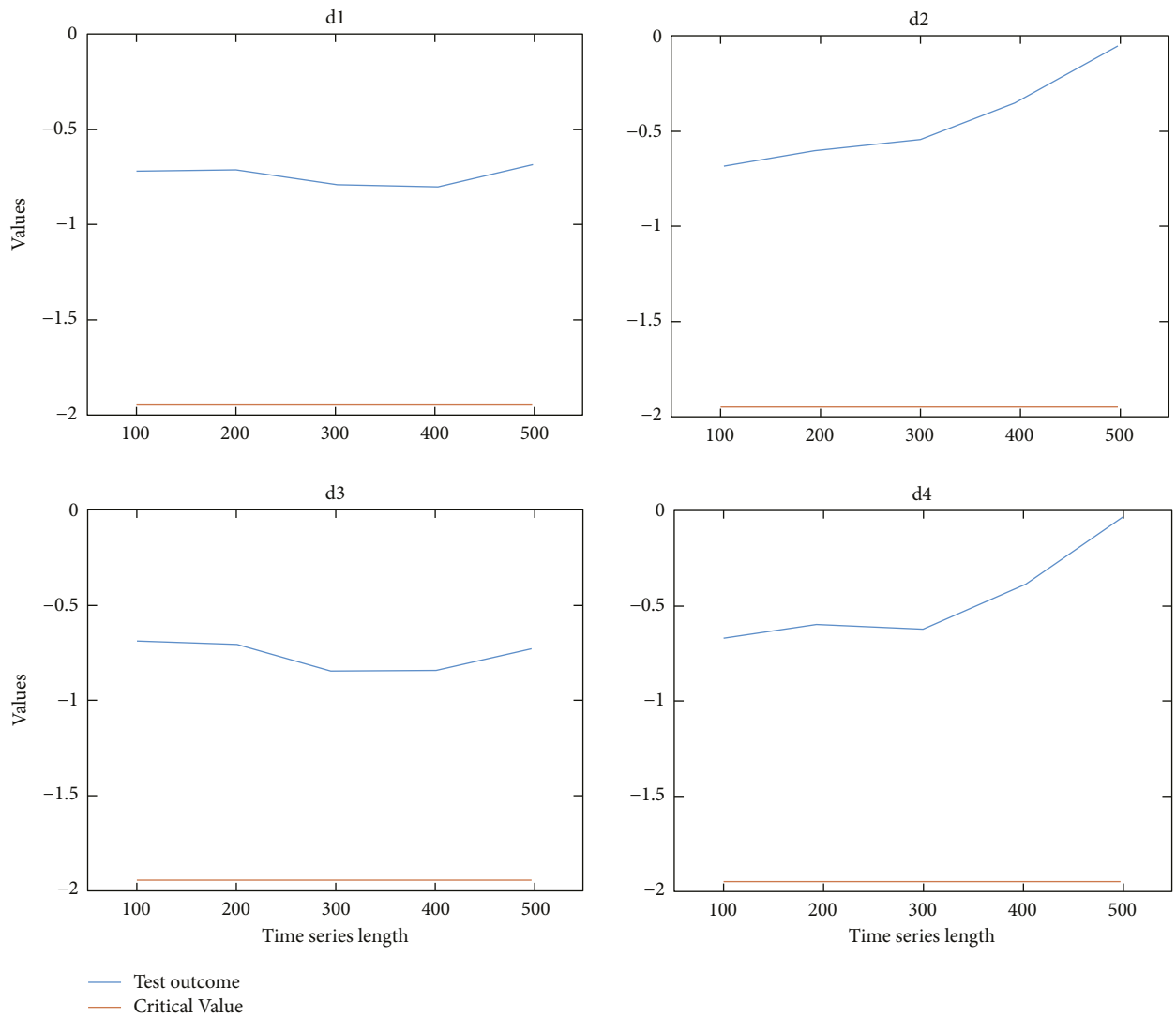


Figure 6: Impact of time series length and time series effects on test results and critical values of ADF test.

known as one-way ANOVA on ranks as it classifies its data into ranks and performs a one-way ANOVA test. The critical values for this test also depend on the level of significance and corresponding degrees of freedom [40]. Further, the two-way KS test is also one-tailed, comparing ECDFs effectively [15]. The critical value for this test is calculated as follows: $D_{m_1,m_2} > \sqrt{-\ln(c_\alpha/2)1 + m_2/m_1/2m_2}$ [15], and thus it is dependent on group size and level of significance. Lastly, the SW test is also based on one-way ANOVA. Therefore it is one-tailed with critical values dependent on the significance level and corresponding degrees of freedom [11, 26]. The formulations of $\widehat{\sigma}_\rho$ and $\widehat{\sigma}_\varepsilon^2$ are provided in [10], and those for calculating a_i are given in [26].

The existing test combos in the literature include, (i) ADF and KPSS [27], (ii) ADF and PP [24], (iii) KPSS and PP [13], (iv) ADF, KPSS, and PP [30], (v) ADF-GLS, PP, and normality tests [11], (vi) ADF, PP, DF-GLS, and Levene's tests [12], (vii) ADF, KPSS, PP, and KW tests,, and (viii) ADF, KPSS, seasonal MK and normality tests [14]. Breitung test [31] and two-way KS test [15] are not used with other tests. The unit root tests, i.e., ADF, KPSS, PP, and Breitung tests, and the MK test, can analyze the trend component in time series. The seasonal MK test is considered to avoid any discrepancies in the test results due to time series seasonality. Similarly, the Fisher, Levene, and KW tests help characterize seasonal behavior in time series. Two-way KS test analyses stationarity by comparing ECDFs in various time series fragments. Therefore, it should be capable of characterizing all the time series components, which account for its nonstationarity if the sizes of fragments are correctly chosen. Lastly, the SW normality test can analyze time series stationarity as normally distributed data are bound to be stationary. These stationarity tests cannot detect all the nonstationarizing time series factors such as trend, seasonality, and volatility. Further, the available stationarizing methods cannot remove all the nonstationary time series components. So a sequence of techniques must be employed for an effective stationarization. Therefore, deciding on a series of steps in which the stationarity components are required to be removed is essential and is suggested in

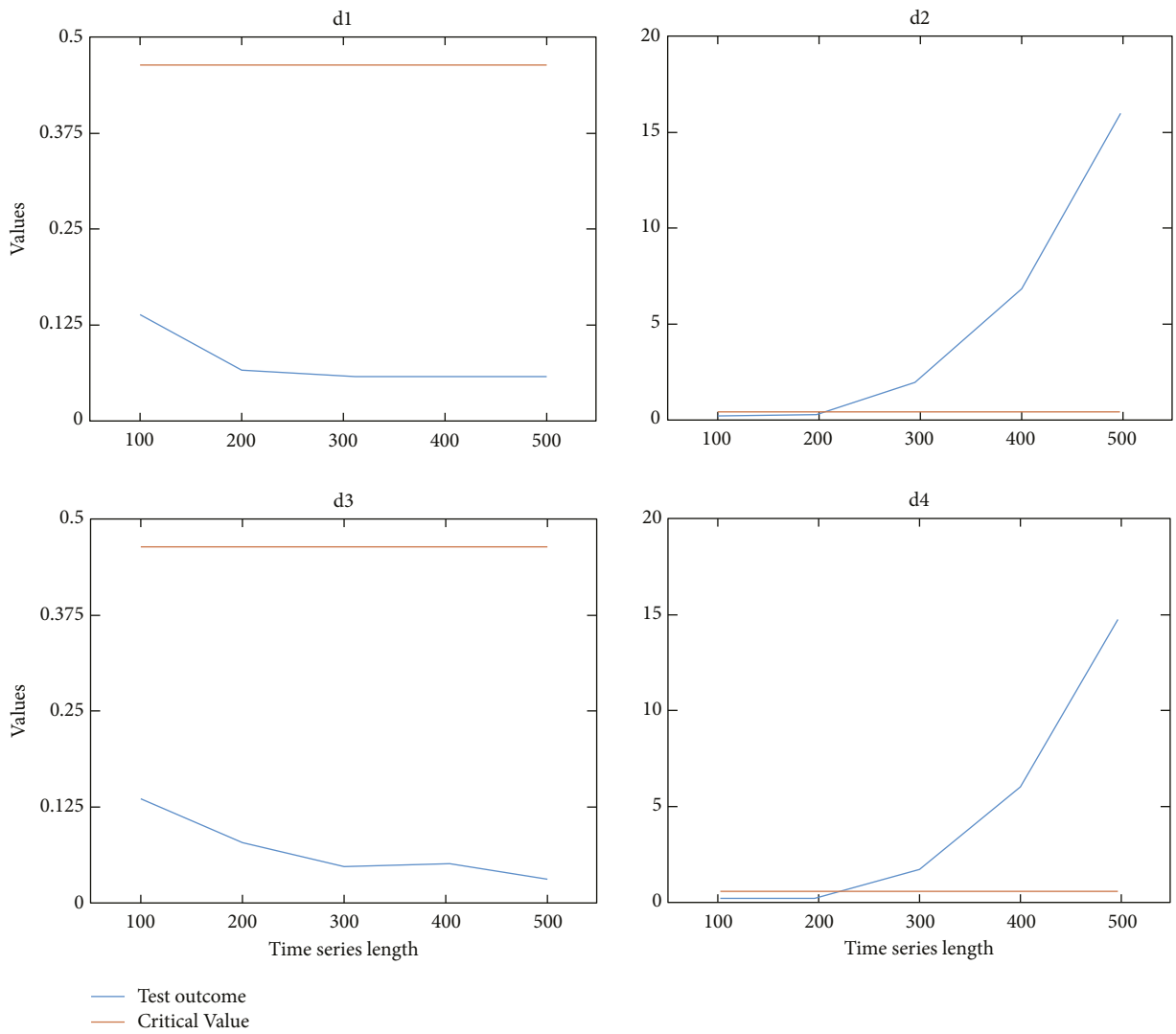


Figure 7: Impact of time series length and time series effects on test results and critical values of KPSS test.

Figure 2. The tests that can detect trend components can yield unbiased results without being concerned about the presence of seasonality or volatility in the time series. Tests useful for seasonality assessment provide biased results in the presence of trend components as they fail to detect differences between seasonal and trend variations. In contrast, the presence of volatility is not a prime concern. Lastly, as volatility is generally detected using normality tests or similar tests, the time series should be free from any trend or seasonality component for true evaluation of the time series volatility. In the presence of trend or seasonality, these tests will provide results pertaining to these facets, which means that these components overshadow the results as volatility remains unassessed.

The variations in the above-discussed tests necessitate a suitable classification for duly understanding the similarities and differences between all these tests. Thus, a novel standard classification tree for stationarity tests is proposed in Figure 3. A test's working principle being the main reason how a test is different from others is generally ignored in the existing stationarity test classifications; hence, after classifying the tests as unit root and nonunit root tests, the proposed classification is mainly prioritized on the working principles of the nonunit root tests. A detailed elaboration is provided underneath. The tests are classified as parametric, semiparametric, and nonparametric. Parametric tests assume the data to be normally distributed, whereas nonparametric tests do not have such assumptions. PP test is classified as semiparametric as initially, it follows a parametric model, but while constructing test statistics, a nonparametric approach is adopted. The tests are classified as EDF-based, regression and correlation-based statistical moment-based, sign function-based, and one-way ANOVA-based tests. Here, EDF-based tests function by computing the EDF of data. In contrast, regression and correlation-based tests use the ratio of normally distributed least squares estimates and the variance of the data sample. Similarly, statistical moment-based and sign function-based tests involve calculating moments and using sign function, respectively. One-way ANOVA-based tests mostly employ the

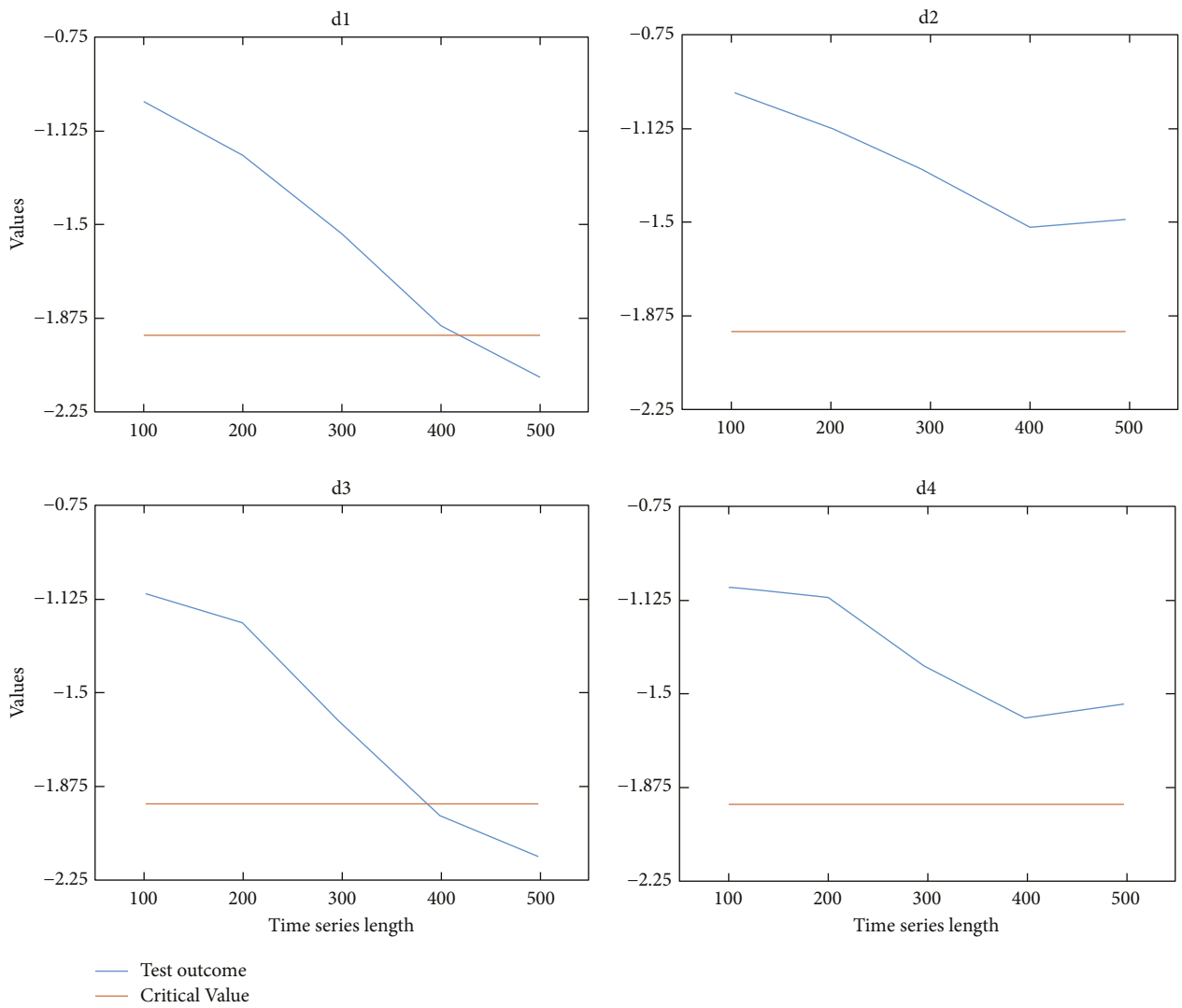


Figure 8: Impact of time series length and time series effects on test results and critical values of PP test.

same principles as classical one-way ANOVA tests; hence, they can also be called improvements of one-way ANOVA tests. Lastly, parametric EDF-based tests are further classified as supremum and quadratic based on the computation of test statistics by respective tests. Tests entailing supremum yield test results by calculating the largest difference between two quantities, whereas quadratic tests do the same by calculating the quadratic difference.

3. Stationarity Test Attributes

The results of the tests for time series stationarity are often biased by time series facets, further affected by the time series length [38, 41, 42] as well as time series clustering [14, 39, 40, 43]. Various other parameters associated with the tests can also bias the test results. A practical stationarity test is expected to correctly indicate whether a given time series is stationary or not. Further, the test outcome should remain unbiased irrespective of the above facets.

3.1. Impact of Time Series Length on Test Results. The length of the time series significantly impacts the test results and critical values [38, 41, 44]. This capability varies differently for different tests. The time series length must not influence a test in providing unbiased results. For any given length, the tests should ideally yield an unbiased outcome; therefore, it is necessary to study the impact of length on the test results. This would provide colossal information about the tests' efficacy while dealing with time series of different lengths. The effect of time series length on test results can be noted by comparing the test results and critical values for various lengths, which help notice the significance of a test declaring a time series as stationary or

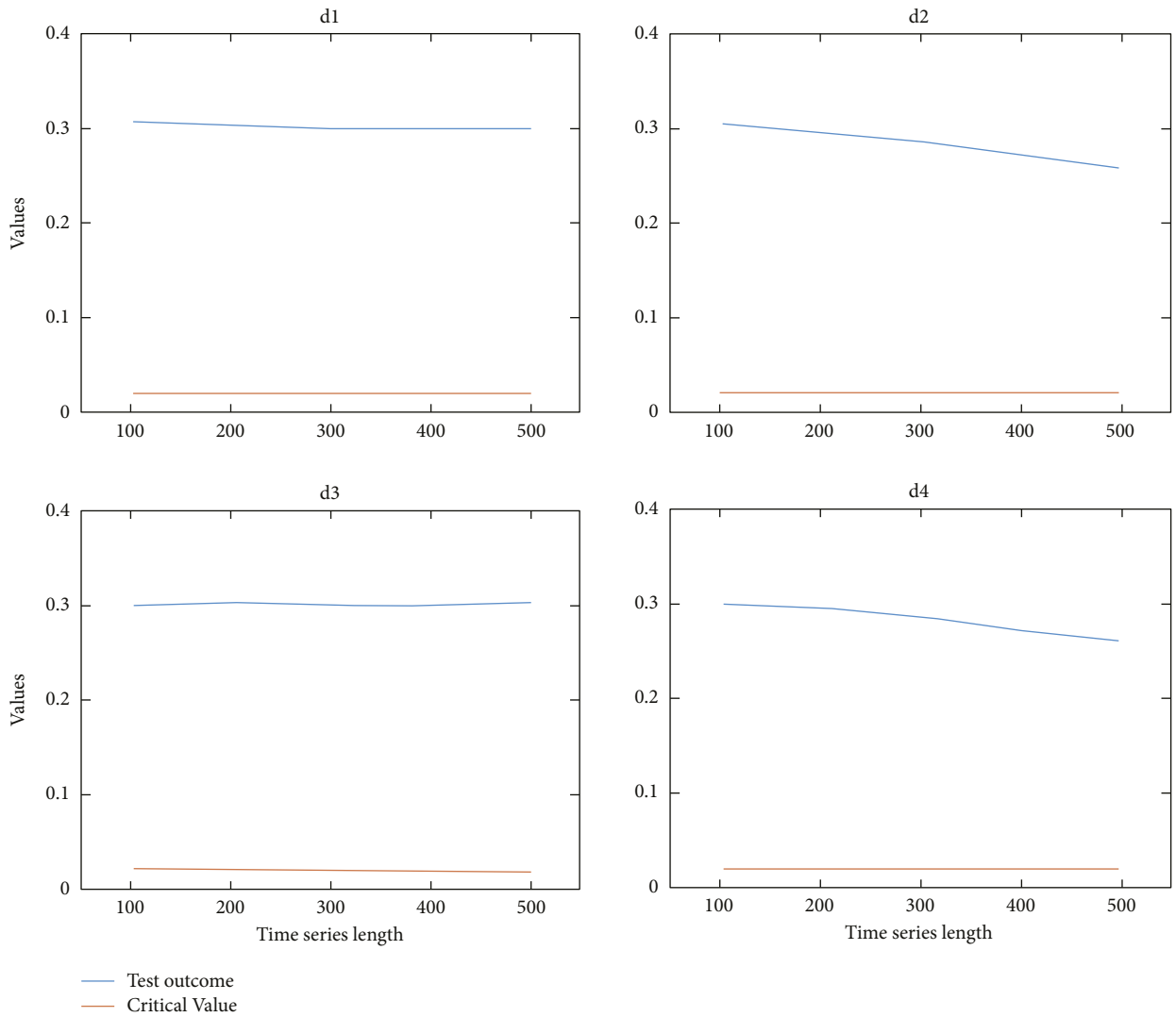


Figure 9: Impact of time series length and time series effects on test results and critical values of Breitung test.

nonstationary. Actually, the test result is compared with the critical value to accept or reject the null hypothesis. Thus, if the positive difference between the test result and the corresponding critical value is higher, the null hypothesis is accepted or rejected by a higher significance. Therefore, if a time series rejects the nonstationary hypothesis by a large factor, then the time series is more strongly stationary. A nonstationary component in the time series may be overlooked when the time series considered for testing has a high length. Perhaps for a shorter length, the presence of the component is highlighted through the results. The impact of length on test results can only be studied for ADF, PP, KPSS, Breitung, MK, and SW tests. The rest of the tests analyze time series by dividing them into groups or clusters; hence, the impact of length on these tests' outcomes cannot be characterized.

3.2. Impact of Time Series Clustering on Test Results. Some stationarity tests function by dividing the time series into various fragments [14, 39, 40, 43]. All these fragments are compared and analyzed, and test results are obtained. Among all the considered tests, Levene's, KW, and two-way KS tests are built to examine the time series by dividing them into various groups or clusters. These fragments, groups, or clusters are created by considering some parts of the time series without any intermixing of data. The size of the group can cause notable variations in test results. It is crucial to know the apt size of the groups for respective tests to obtain accurate and unbiased results. Considering a very small or very large group size may lead to significant discrepancies in test results. Further, cluster sizes can also be unequal for all these tests. KS test can only test two clusters at a time, while Levene's and KW tests can examine many groups at a time,

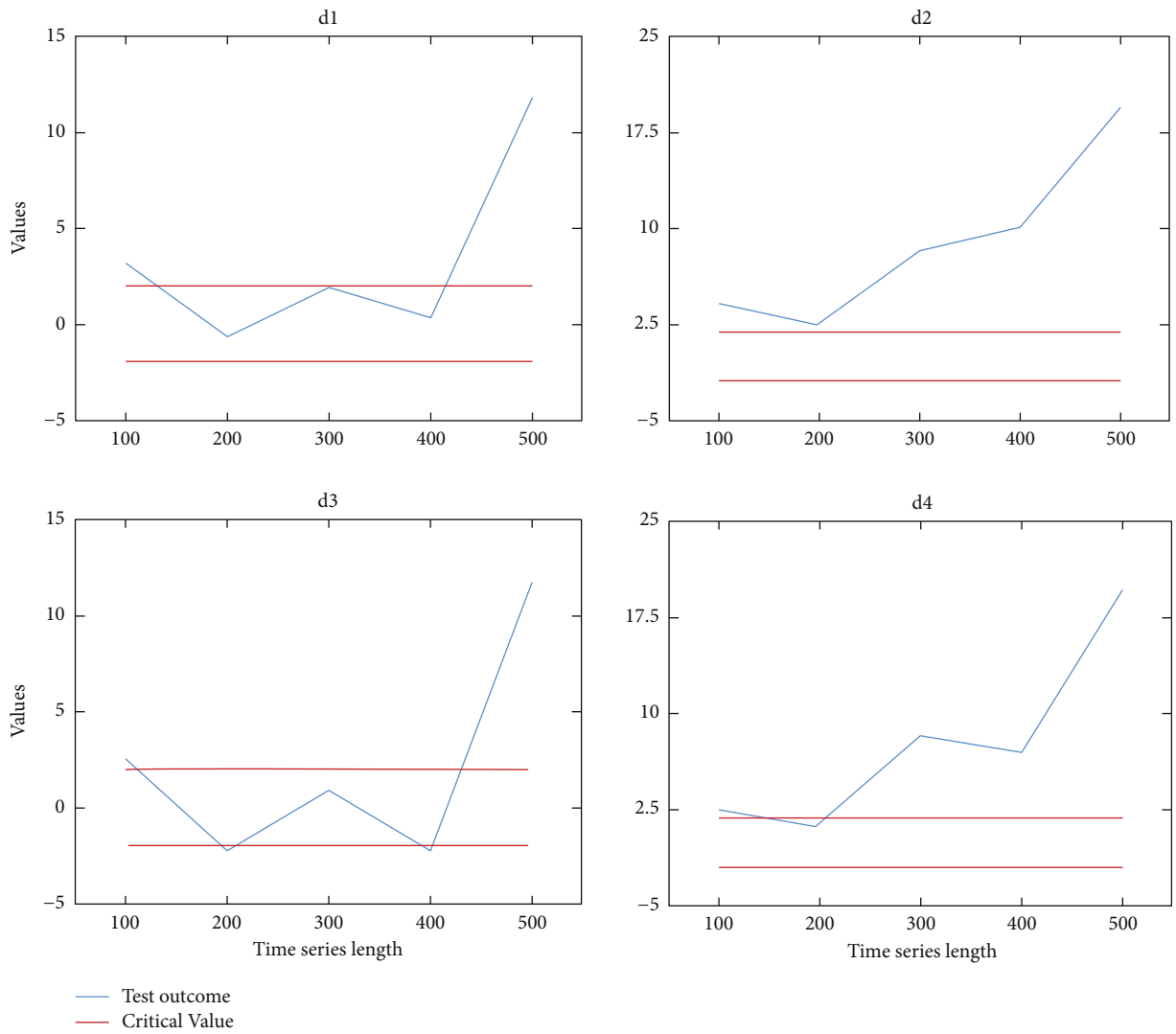


Figure 10: Impact of time series length and time series effects on test results and critical values of MK test.

and thus the number of groups also turns out to be an essential attribute. The size of clusters, the number of clusters, or the fact that they are equally or unequally sized, should not affect the outcome of an ideal test. But in reality, these elements have a substantial effect on the results, so it is necessary to note their behavior for all the variations in these elements.

3.3. Impact of Time Series Facets on Test Results. Trend and volatility effects in a seasonal time series account for its nonstationarity and might bias the test results [32]. It is also possible that the tests may overlook these effects and may fail to yield unbiased results. It is vital to notice the changes in test results and critical values due to the trend effect as the facet cause nonstationarity in time series [38, 41, 42, 44, 45]. Such an analysis can help understand how impactful a trend effect could be in making a test bias. Similar to the trend effect, it would also be interesting to notice the changes in test results due to volatility effects. The impact of volatility on test outcomes is often characterized based on the test results, as the volatility effect has less impact on changing critical values [29]. The presence of deterministic seasonality may inflate the size and reduce the power of the ADF test; as a result, the distribution of the test is shifted towards the left, and the dispersion is lower [44]. The presence of deterministic seasonality components can lead to discrepancies in the decision-making process of the KPSS test under the null hypothesis [44]. Possibly, other tests can also characterize seasonality through test results similarly.

3.4. Impact of Changes in Test Parameters on Test Results. Various test parameters characterize stationarity tests; they define the primary features of respective stationarity tests, e.g., the median is typically employed in Levene's test whenever the data distribution are not symmetrical, whereas the mean is appropriate when the data come from a symmetrical distribution [12]. Thus, significant impacts of these modifications could be observed through the test results. The parameters corresponding to various tests are listed as follows:

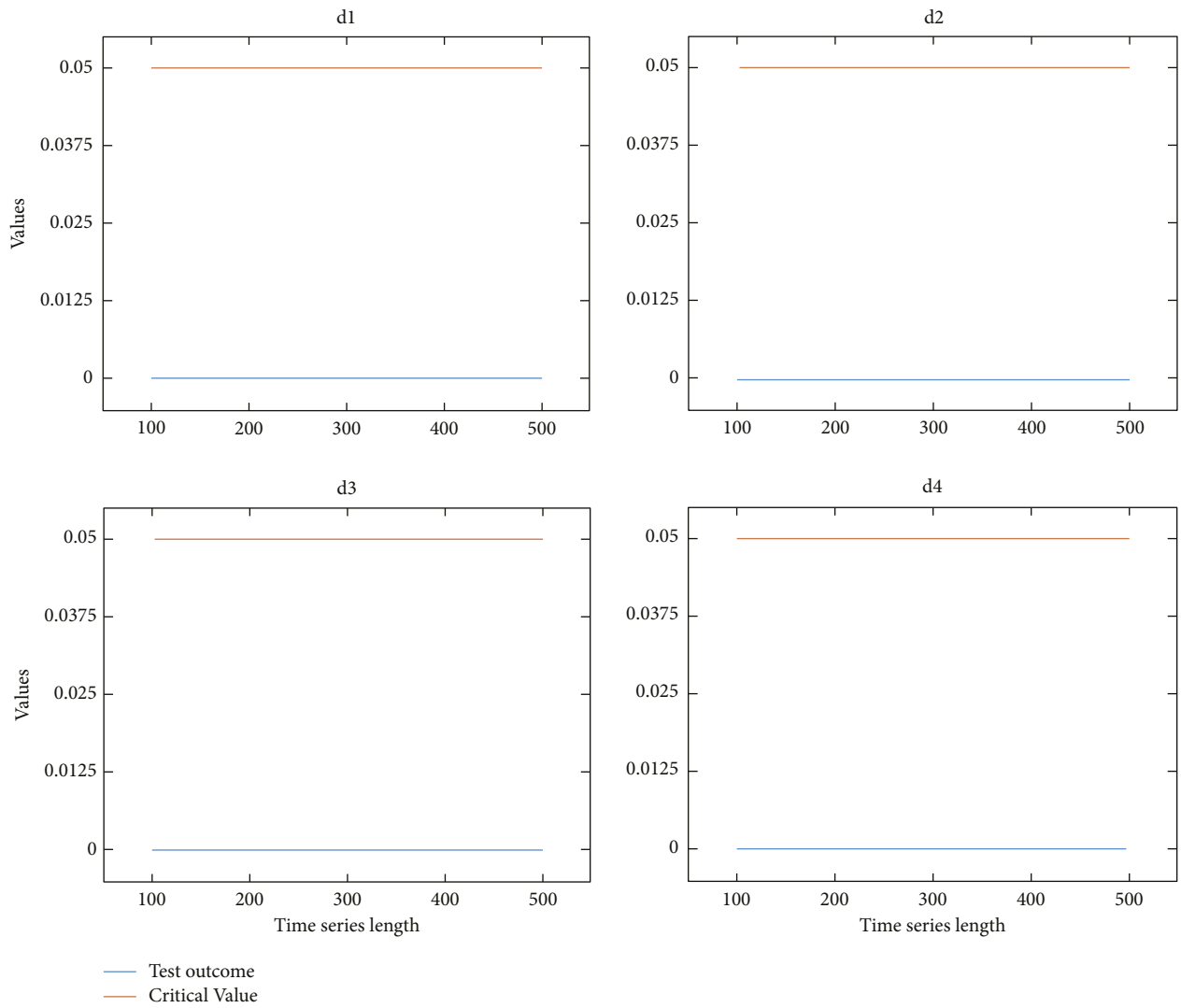


Figure 11: Impact of time series length and time series effects on test results and critical values of SW test.

(i) ADF – $k_1 = 0/k_1 \neq 0$, $k_1 = k_2 = 0/k_1 \neq 0, k_2 \neq 0$, n_1 [41].

(ii) KPSS – $k_2 = 0/k_2 \neq 0$ [44].

(iii) PP – $k_1 = 0/k_1 \neq 0$, $k_1 = k_2 = 0/k_1 \neq 0, k_2 \neq 0$ [45].

(iv) Breitung – $d_t = 0/d_t \neq 0$ [38].

(v) MK–Consideration of seasonality [14].

(vi) Levene's–Analysis using mean or median [12, 39].

4. Result Analysis

The selection of effective stationarity tests for a given application necessitates a thorough examination of well-established tests, investigating their efficacy when dealing with multifold seasonal datasets with trend and volatility effects. A comprehensive survey is usually helpful for exposing tests' weaknesses or failures when dealing with a particular time series facet. On this note, the performance of the nine well-established tests for four datasets (refer to Figure 4) is examined to check their ability to yield unbiased results. The analyses carried out in this section are fivefold. Firstly, the impact of test parameters on the test outcomes is studied considering four clean seasonal time series with synthetically embedded trend and volatility effects. Secondly, with the same datasets, the impact of time series length on test results is comprehensively studied. Furthermore, the sensitivity of stationarity tests' outcome to time series clustering considering cluster size, cluster ratio, and cluster numbers is examined. Lastly, deliberate discussions of the obtained results are carried out and provided with appropriate suggestions on the selection of tests. The non-stationary datasets used for the above analyses are summarized in Table 1. Considering the four WGN-assisted seasonal datasets with synthetically embedded time series facets helps make the study more generic and informative. The load power consumption of a restaurant in Anchorage,

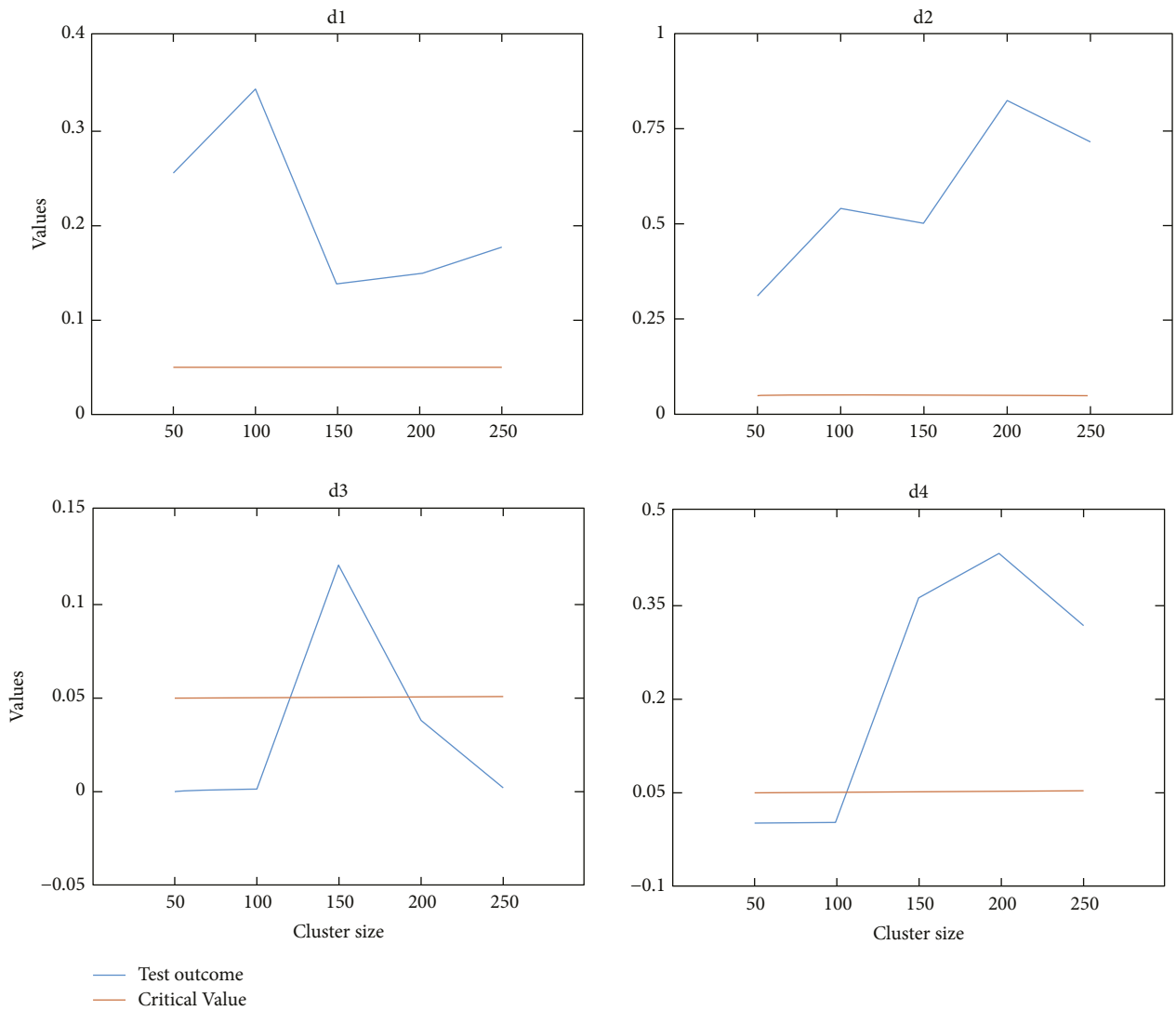


Figure 12: Impact of cluster size and time series effects on test results and critical values of Levene's test.

USA [46], denoted as d1, is the primary dataset for creating d2, d3, and d4.

4.1. Impact of Test Parameters on the Outcome of ADF, KPSS, PP, Breitung's, MK, and Levene's Tests

4.1.1. Impact of Test Parameters on the Outcome of Unit Root Tests, ADF, KPSS, PP, and Breitung. All unit root tests involve time series modeling (refer to Figure 1). While modeling time series, k_1 and k_2 values are essential parameters for tests such as ADF and PP. For the KPSS test, k_2 is only considered. The overall deterministic component is considered for the Breitung test, similar to ADF and PP tests. These tests model the time series with a drift and a time-varying trend component upon consideration of k_1 and k_2 values, respectively. These components are then removed from the time series using OLS detrending method in ADF, KPSS, and PP tests. In contrast, GLS detrending technique is used in the Breitung test. Finally, the residual is tested for stationarity. When time series is modeled under consideration of these parameters, provided these components exist in the time series, higher accuracy can be expected from results for conformance of time series stationarity. The observations from the obtained results are similar for all the unit root tests (refer to Table 2 through Table 5). For all these tests, consideration of k_2 or time-varying trend component in d_t does not have any impact on test outcome for data with no time-varying trend as in the case of d1 and d3. In contrast, due to the same trend component, notable changes are seen in results for d2 and d4 (refer to Table 2 through Table 5). Unit root tests effectively characterize trend components but cannot describe seasonality and volatility effects. Seasonality and volatility bias the test results of ADF, KPSS, and PP tests. Notably, the Breitung test did not render any biased results like all other unit root tests in the presence of the seasonality effect, thus showcasing its ability to detect the seasonality effect. Also, results without d_t indicate that it is better to consider the drift and trend components if they are present in time series for effectively detecting seasonal components. Volatility also has a minimal but positive impact on the test under all

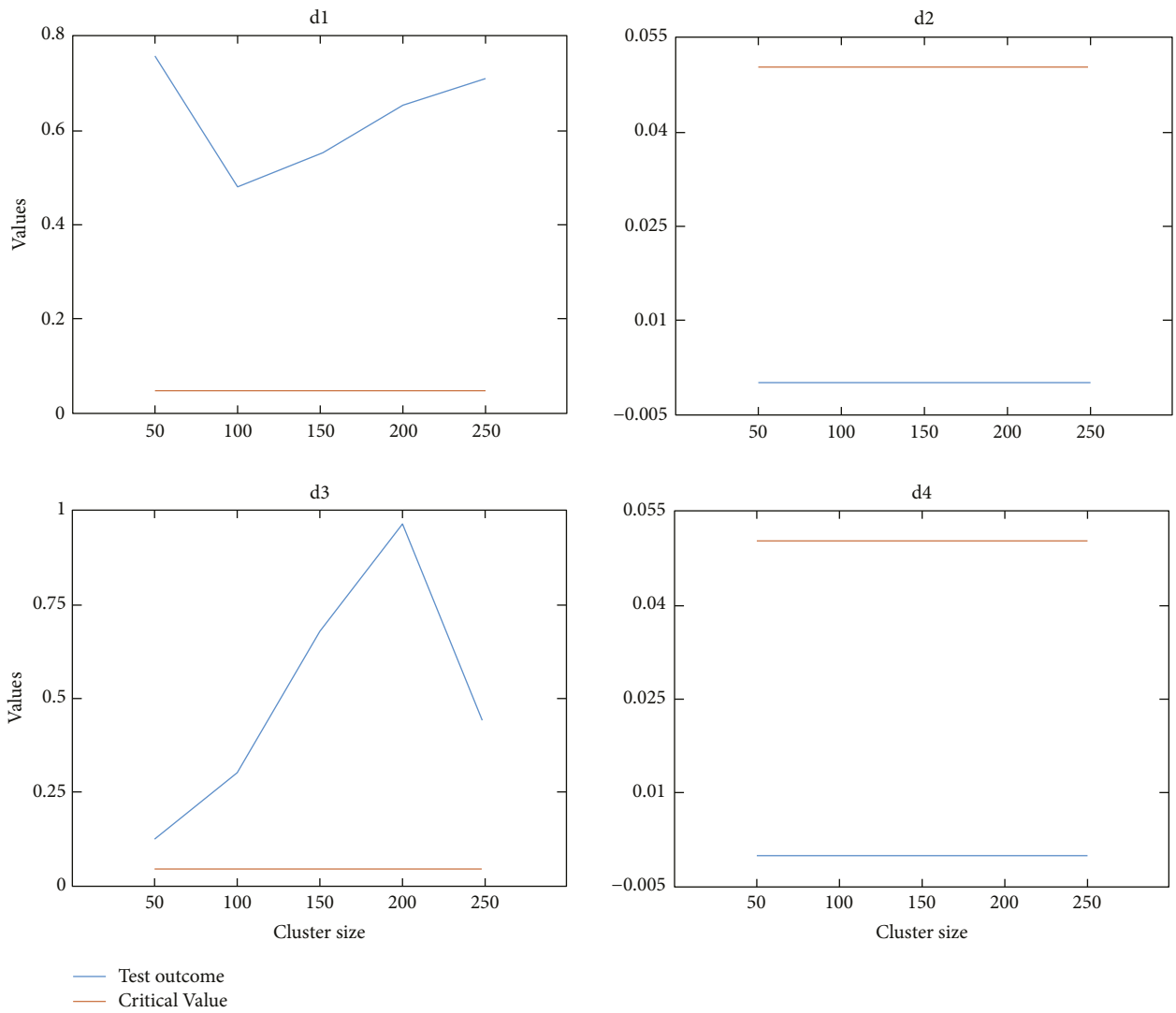


Figure 13: Impact of cluster size and time series effects on test results and critical values of KW test.

parameters. Further, critical values for all the tests are also reduced under consideration of k_1 and k_2 values or d_t. To check whether time series is not stationary according to the ADF and PP tests, $k_1 = k_2 = 0$ model is better for use. The results obtained under consideration of parameters $k_1 \neq 0$ with $k_2 = 0$ and $k_1 \neq 0$ with $k_2 \neq 0$ can be compared with of obtained with $k_1 = k_2 = 0$ as they indicate the necessary actions required for stationarizing the considered data. A similar approach can be used to check the presence of a trend with the KPSS test and the presence of drift and time-varying trends with the Breitung test.

4.1.2. Impact of Lag Number on ADF Test's Outcome. ADF test involves the determination of n_l values as it is an integral part of the model equation (refer to Figure 1). These lag numbers significantly impact ADF test results under various time-series effects. In all the cases, the test outcome decreases for $n_l = 1$, and then inconsistently rises as the n_l value increase to $n_{l,\max} = 17$. The critical values also decrease uniformly by a minimal factor with an increment in n_l value. Overall, it is seen that unbiased results are obtained for higher lag numbers. If the lag number is too small or too large, a loss in power of the ADF test is noticed [47]. This increases the chance of getting biased results. The trend component has notable effects on the test results as the entire plot shifts upward for d2 and d4, which can be visualized in Figure 5 as unit root tests can characterize the trend component effectively. The volatility effect is seen biasing the test outcome as the plots for cases d3 and d4 shift slightly down compared to d1 and d2 (refer to Figure 5).

4.1.3. Impact of Test Parameters on MK Test's Outcome. For the MK test, considering seasonality is an essential parameter, as seasonal variations can lead to biased test outcomes. Consideration of seasonality enables the test to detect trend components in time series in the presence of seasonality by ignoring the seasonal variations, as it is possible that these variations can bias the test results because

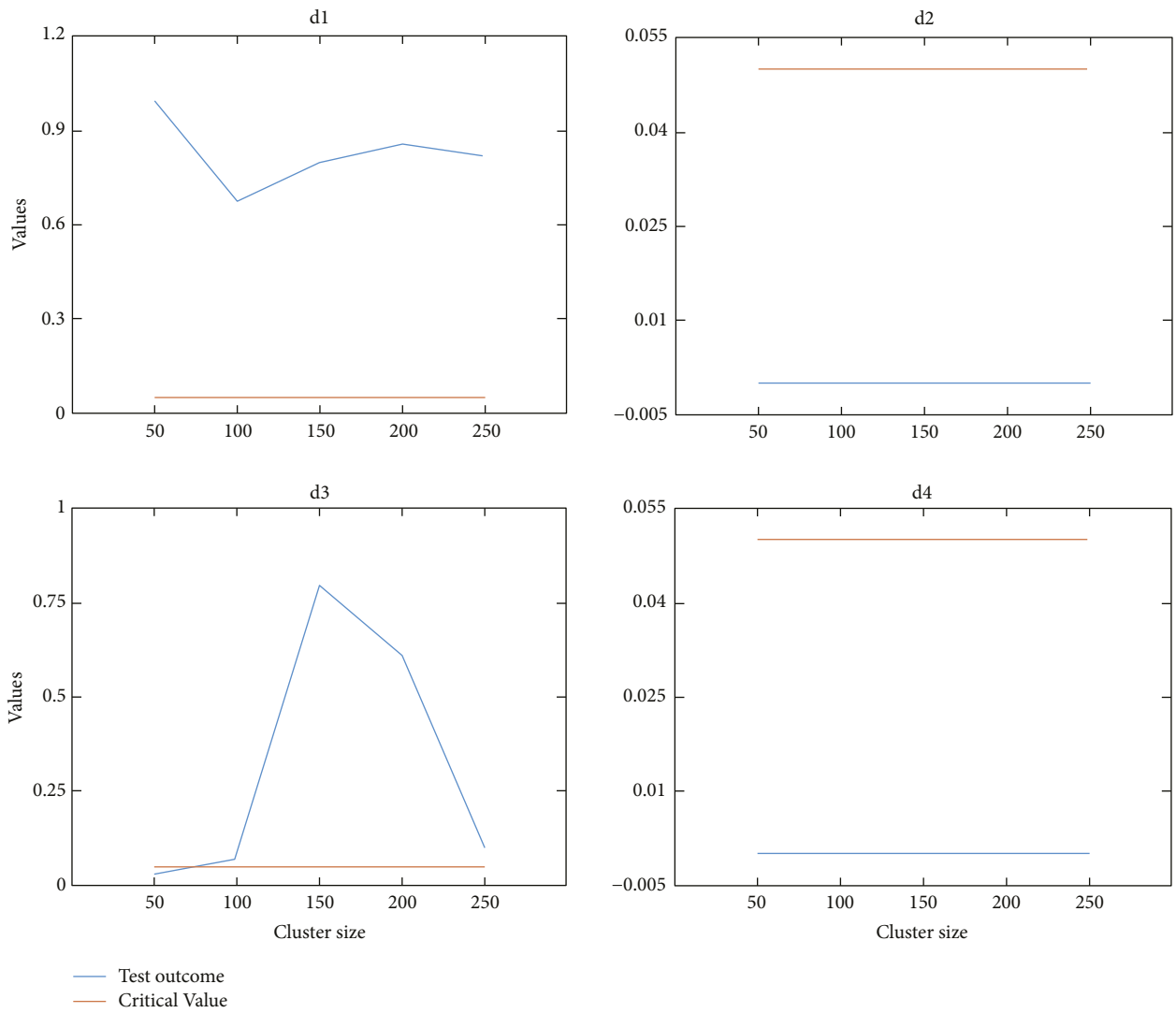


Figure 14: Impact of cluster size and time series effects on test results and critical values of KS test.

the test detects a monotonic trend by examining a gradual difference between data over time. MK test is designed to detect monotonic trends; hence, it effectively detects trend components in all the data (refer to Table 6). Biased results are obtained for the MK test, even considering seasonality due to the test's high frequency of committing type-II errors. Thus, the overall outcomes are better when no seasonality is considered. Also, volatility has minimal or no effects on test outcomes of the MK test.

4.1.4. Impact of Test Parameters on Levene's Test's Outcome. Levene's test is designed to test the stationarity of a time series by dividing it into clusters, and the test formulations require the calculation of the mean or median of the groups (refer to Figure 1). Therefore, consideration of the mean or median for time series analysis is a vital parameter for the test. In this analysis, the value of "G" is taken as 100. It is noticed that Levene's test using mean has an overall better performance than that using the median as the results obtained are more biased for analysis using median than that using mean (refer to Table 7). The volatility effect is evidently characterized by Levene's test, which can be seen in the results for d3 and d4. The test is seen as giving biased results for seasonality and trend components as the test is only capable of detecting a difference in variance. There is no significant difference in the variances of different groups, so biased results are obtained.

4.2. Impact of Time Series Effects on the Outcome of KW, KS, and SW Tests. KW test detects a difference in mean values between groups, whereas the two-way KS test checks the difference between the distributions of two groups. Thus, analyzing the time series is divided into five equal groups with $G = 100$ for the KW test and two groups with $G = 250$ each for the KW test. Both tests significantly characterized the trend's impact, as seen in Table 7. The test outcomes are returned as 0, indicating that the stationary null is being rejected with utmost confidence. Biased results are obtained with both the tests for data in the absence of trend components, e.g., d1 and d3, due to the incapability of these tests to detect seasonality. But overall, the KS test performed better than the KW test as the results for KW tests are very

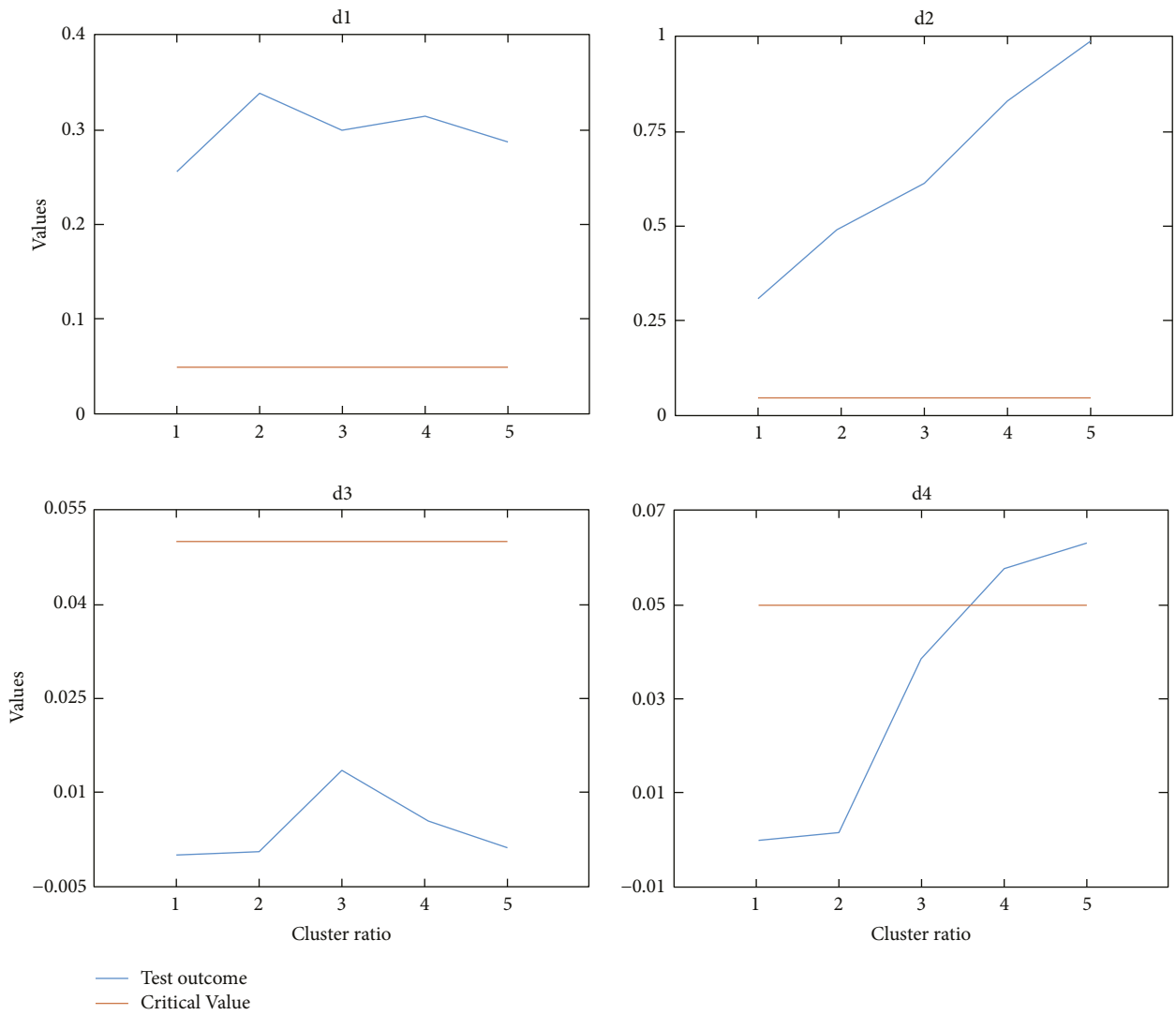


Figure 15: Impact of cluster ratio and time series effects on test results and critical values of Levene's test.

significantly biased compared to that of the KS test. Also, the volatility effect is quite effectively characterized by the KS test, as inferred through the results. Lastly, the SW test has given unbiased results. It is clear from the results that the test can detect seasonality (refer to Table 8). For d2 and d4, which have a trend component, an increase is seen in values of test results though insignificant. Similarly, the test result value rose for d3, having a volatility effect, but the difference is trivial. Thus, trend and volatility effects have minimal impacts on the SW test outcomes.

4.3. Impact of Time Series Lengths on Stationarity Tests' Outcomes. Time series length is an essential feature of a time series. A higher number of data points can assist in building an effective time series model. Therefore, for analysis of test results of ADF, KPSS, PP, Breitung, MK, and SW tests for different time series lengths, the same four sets of time series are considered, with lengths 100, 200, 300, 400, and 500. The variations in test results and critical values are visualized for all the given datasets and are thoroughly analyzed (refer to Figures 6 through Figure 11). For ADF and PP tests, the critical values rise by a small factor as the length increases. In contrast, the Breitung test's critical values decrease as length increases. There is no change in the critical values for KPSS, MK, and SW tests. ADF test result values show slight variations for length with d1, primarily due to seasonal variations. Further, if the same plot is seen for d2, test outcome values are increased as length increases due to the increasing impact of the trend effect. As the ADF test can characterize trend components effectively, the test results indicate higher nonstationarity. Considering the plots in d3 and d4, it is noticeable that they are similar to d1 and d3 but with more significant variations due to the volatility effect. KPSS test results are biased as the length increases for time series with no trend effect. The result values keep decreasing, implying that the null hypothesis is being retained with a higher significance indicating stronger stationarity. While d2 and d4 datasets have an inherent trend component, the outcome values increase with length. The outcome is stationary for

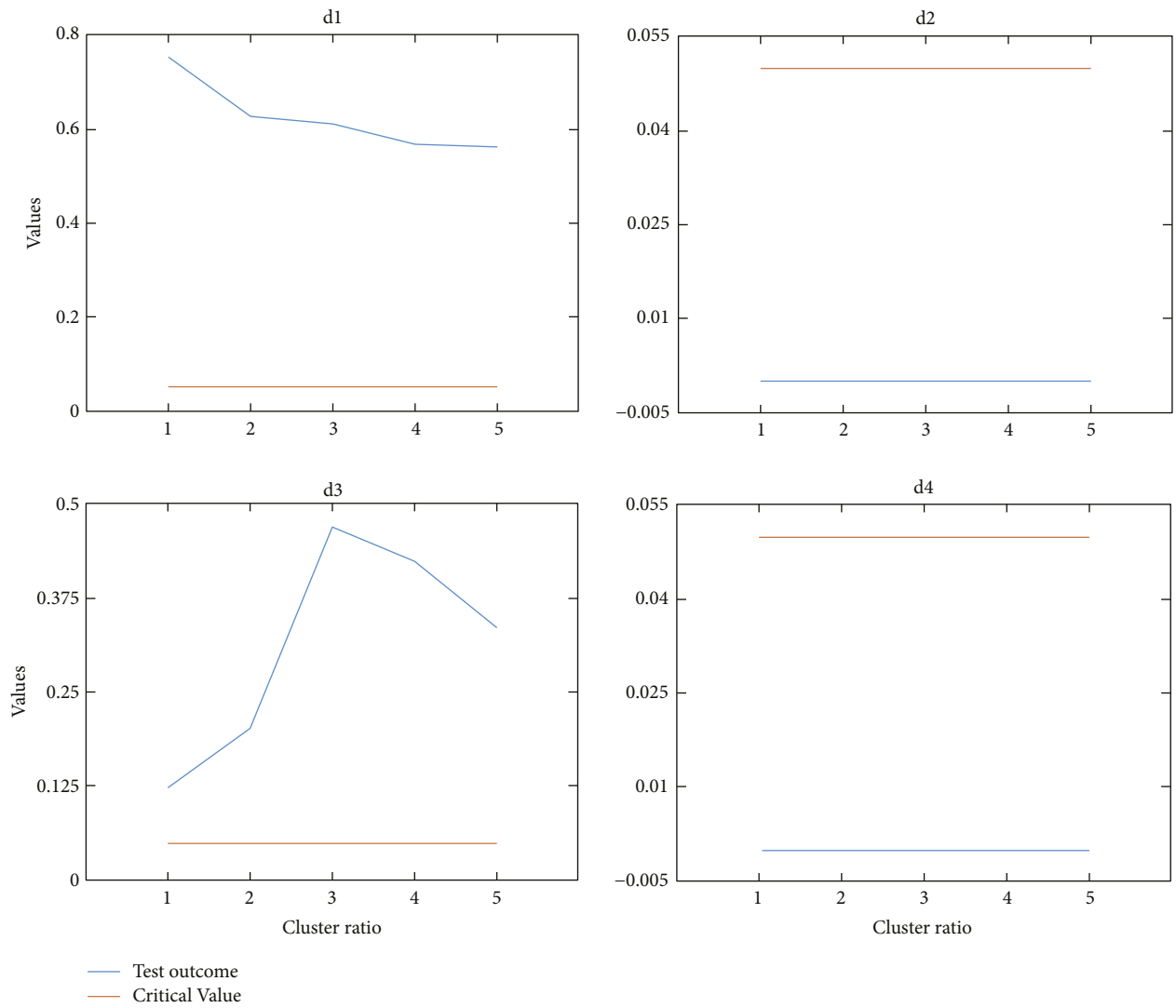


Figure 16: Impact of cluster ratio and time series effects on test results and critical values of KW test.

lengths 200 and 400, as there is no significant trend effect in these lengths. As the length increases, the stationarity null gets rejected with increasing significance due to the trend component. Volatility biases the test results overall as it can be noticed that for d3 and d4, the results indicate stronger stationarity or weaker nonstationarity than those of d1 and d2. Further, PP test result values decrease with an increase in length.

Considering that all the datasets under analysis are nonstationary, a decrease in outcome value with an increment in length indicates that the PP test performs better for time series with lower lengths. The test can also detect seasonal variations for lower lengths, possibly because the test characterizes the variation as a form of the stochastic trend for lower lengths. The test again effectively detects the trend component, as shown in Figure 8, as it is a unit root test. Similar to that of the KPSS test, the volatility effect biases the test results of the PP test too. In the case of the Breitung test, the test result values also decrease minimally as the length increases for d1 and d3 with no trend effect. The difference in the test results and critical values is more or less maintained. But, for d2 and d4, the test results decreased significantly with increased length. The trend effect becomes more prominent as the length increases, and it slightly biases the test results. Volatility has no significant impact on the test results, which can be noticed by comparing plots for d1 and d2 with d3 and d4, respectively. Further, the MK test has two lines of critical values as it is a two-tailed test, and seasonality is considered in this analysis. Trend component is detected for very low and very high lengths in all cases due to high Type-II error rates. For d2 and d4, as the length increases, the test rejects stationary null with higher significance as the trend effect becomes more and more prominent as the length increases. There are changes in plots of d3 and d4 to d1 and d2, respectively, indicating differences in results due to the volatility effect. The MK test possibly detects the volatility effect as a form of trend, and thus, nonstationary outcomes are yielded. Lastly, for the SW test, no significant changes are noticed for time series length for all the datasets. The test is seen behaving optimally against all

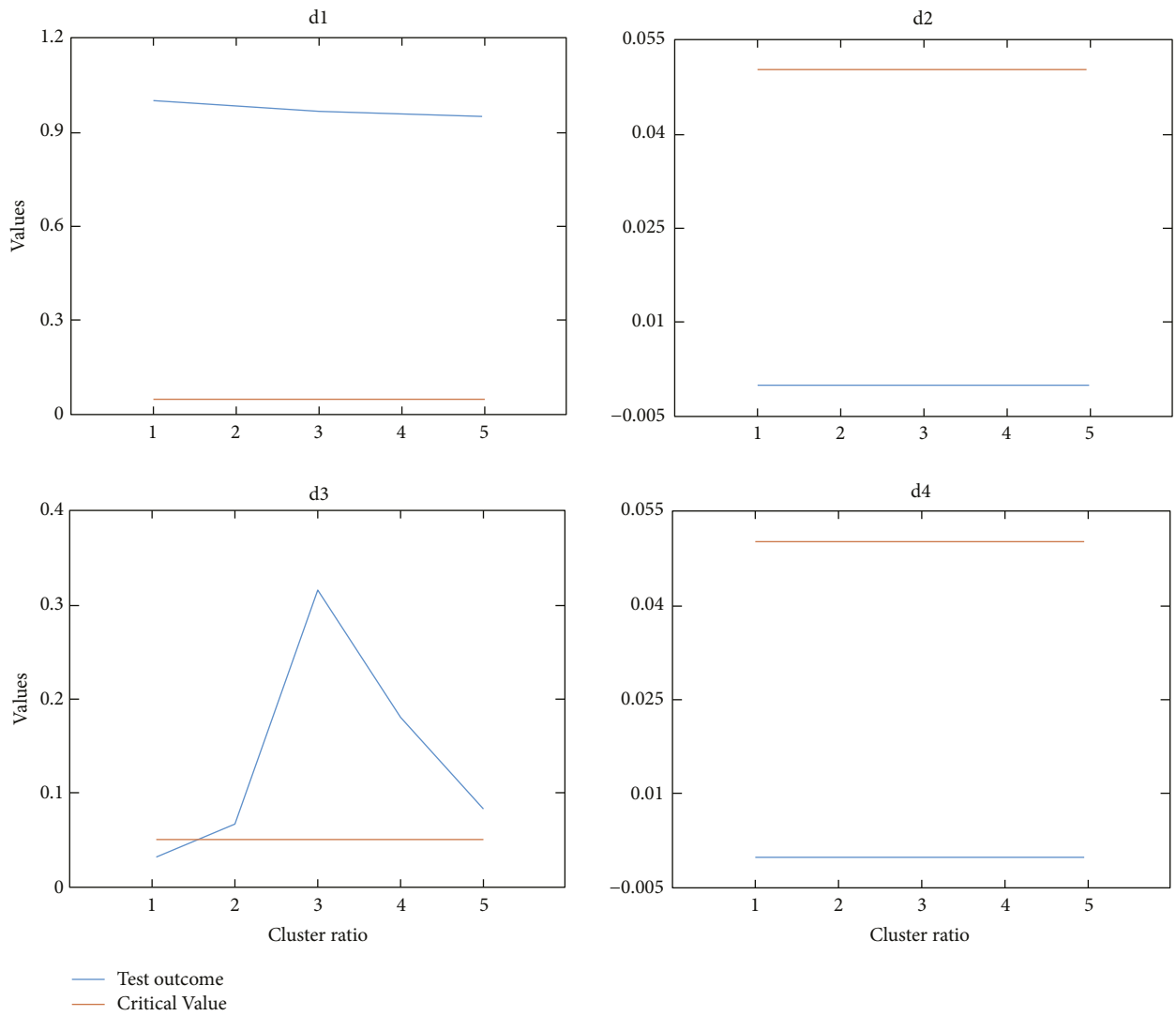


Figure 17: Impact of cluster ratio and time series effects on test results and critical values of KS test.

the considered lengths for all the time series effects, further stating that these components have no significant impact on the test.

4.4. Impact of Time Series Clustering on Stationarity Tests' Outcome. A threefold analysis of impacts of time series clustering considering various time series effects is executed. Levene's, KW, and KS tests can be used with equal and unequal-sized clusters. For equal-sized group analysis, two groups are taken with sizes varying from 50 to 250 with a difference of 50. Next, the term cluster ratio is employed to conduct a uniform analysis of these tests for unequal-sized groups. It is defined as the ratio of the size of group 2 to that of group 1. The cluster size of group 1 is taken as 50 throughout the cluster ratio analysis, whereas the size of group 2 increases to 50, 100, 150, 200, and 250. Therefore, to study the impacts of increasing inequality among groups, the test results are plotted against cluster ratio. Further, Levene's and KW tests can be employed with any clusters, while for the KS test; the number of clusters is always two. Thus, the variations in the test results for all three can be studied for varying cluster sizes (equal cluster sizes) and increasing cluster ratios (unequal cluster sizes), keeping the number of clusters as two. Hence, the impacts of the overall size of the clusters and the difference in sizes of the clusters on the test results can be visualized. For Levene's and KW's tests, the impacts of the number of clusters are also studied with a constant group size of 50. After analyzing the effects of changing sizes of two equal clusters, it can be noted that Levene's test performs better with a larger cluster size in the presence of only the seasonality effect. From all other plots in Figure 12, it can be marked that the test yields better results for a small sample size. Trend effect yet continues to bias results while the test effectually detects volatility. Actually, using a smaller cluster size would be beneficial for obtaining more precise results. Still, the value of "G" should not be minimal as it could call a stationary time series nonstationary. Further, from Figures 13 and 14, KW and KS tests have similar variations in test results. A drop is seen in result

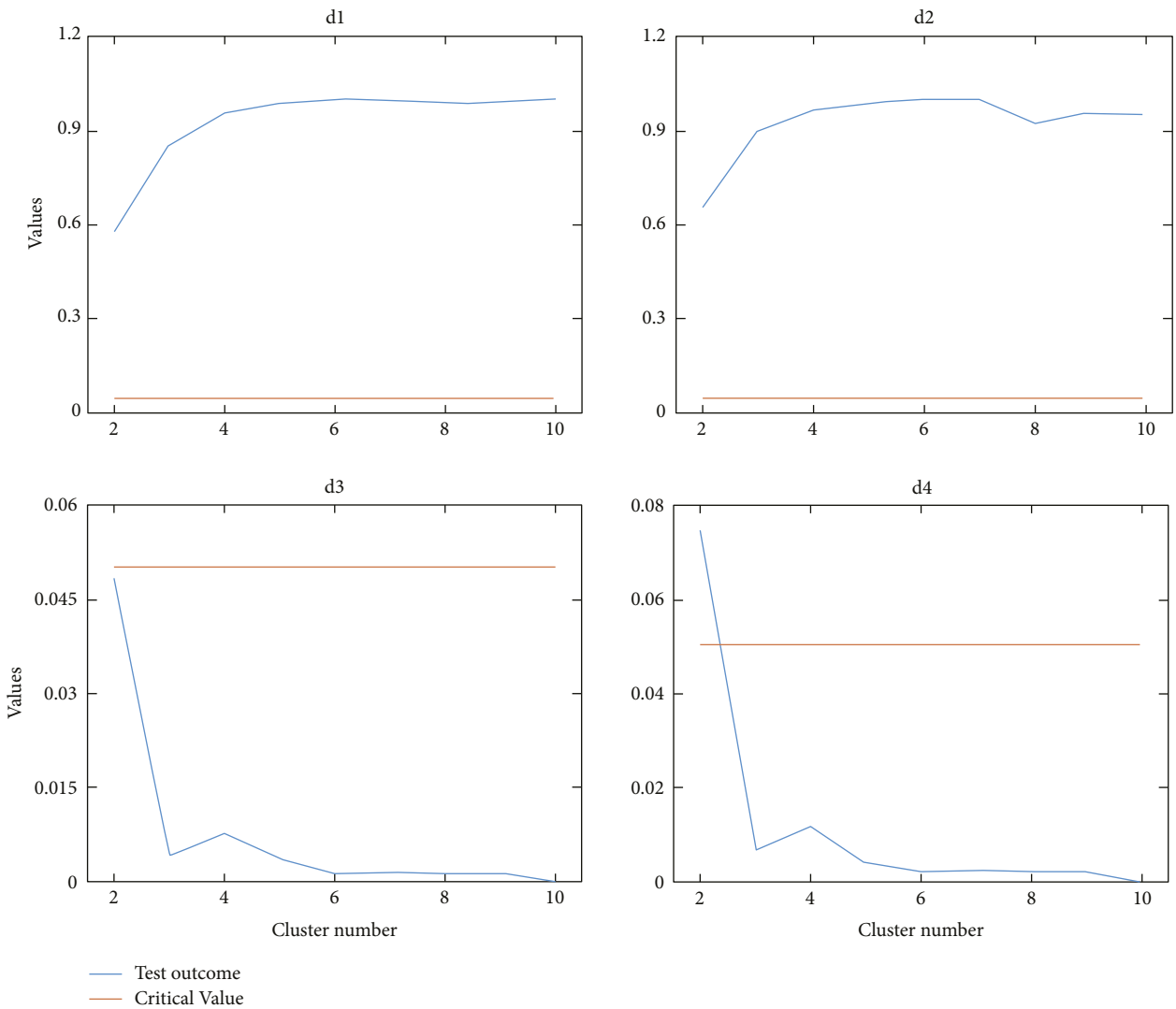


Figure 18: Impact of the number of clusters and time series effects on test results and critical values of Levene's test.

values for $G = 100$ and a further increase is seen as "G" values increments for d1. Likewise, for d3, the result values rise as "G" increases to 200 for the KW test and 150 for the KS test, and then a drop is seen. Overall, the performance of the tests is better for smaller cluster sizes for seasonal and volatile data. In addition to a trend component, as in d2 and d4, perfect results are obtained as these tests can effectively characterize the trend component.

Further, on studying the impacts of changes in cluster ratio on the test results, it can be seen that cluster ratio difference does not significantly affect Levene's test's results for detection of seasonality solely, which can be seen from results for d1 in Figure 15. For d3, overall performance is better for the lower cluster ratio, while an improvement is seen in the results again for a very high cluster ratio. For d2 and d4, having a trend effect, better results are obtained for a lower cluster ratio. Increasing cluster ratio also escalates the biased nature of the test's results for d4. KW test results for d1 indicate that a higher cluster ratio is preferable for detecting seasonality, whereas, for d3, overall performance is improved for a lower cluster ratio. The results show enhancement for a very high cluster ratio (refer to Figure 16). Further considering the KS test, somewhat similar observations are noticed. Variations in results are detected for d3 in Figure 17, considering all these tests due to the strong volatility effect. Lastly, considering d2 and d4, perfect results are obtained for KW and KS tests as both these tests are capable of noticing trend components.

Considering the number of clusters, Figure 18 shows that Levene's test yields better outcomes for lower cluster numbers for d1 and d2. In contrast, for d3 and d4, unbiased and more robust results are obtained for a higher number of clusters. The test is well-suited for analyzing volatility effects. Therefore, the test can notice more variations in data if the number of groups increases, as for d3 and d4. As the data are not volatile for d1 and d2, the variations noticed for the amount of data under testing dominate lower cluster numbers. The test's biased output becomes stronger. Lastly, the KW test fails to detect seasonality in d1 (refer to Figure 19), but comparatively better results are obtained when the cluster number is lower. Further, for d3, vicissitudes are seen in the plots due to the volatility effect. Considering the

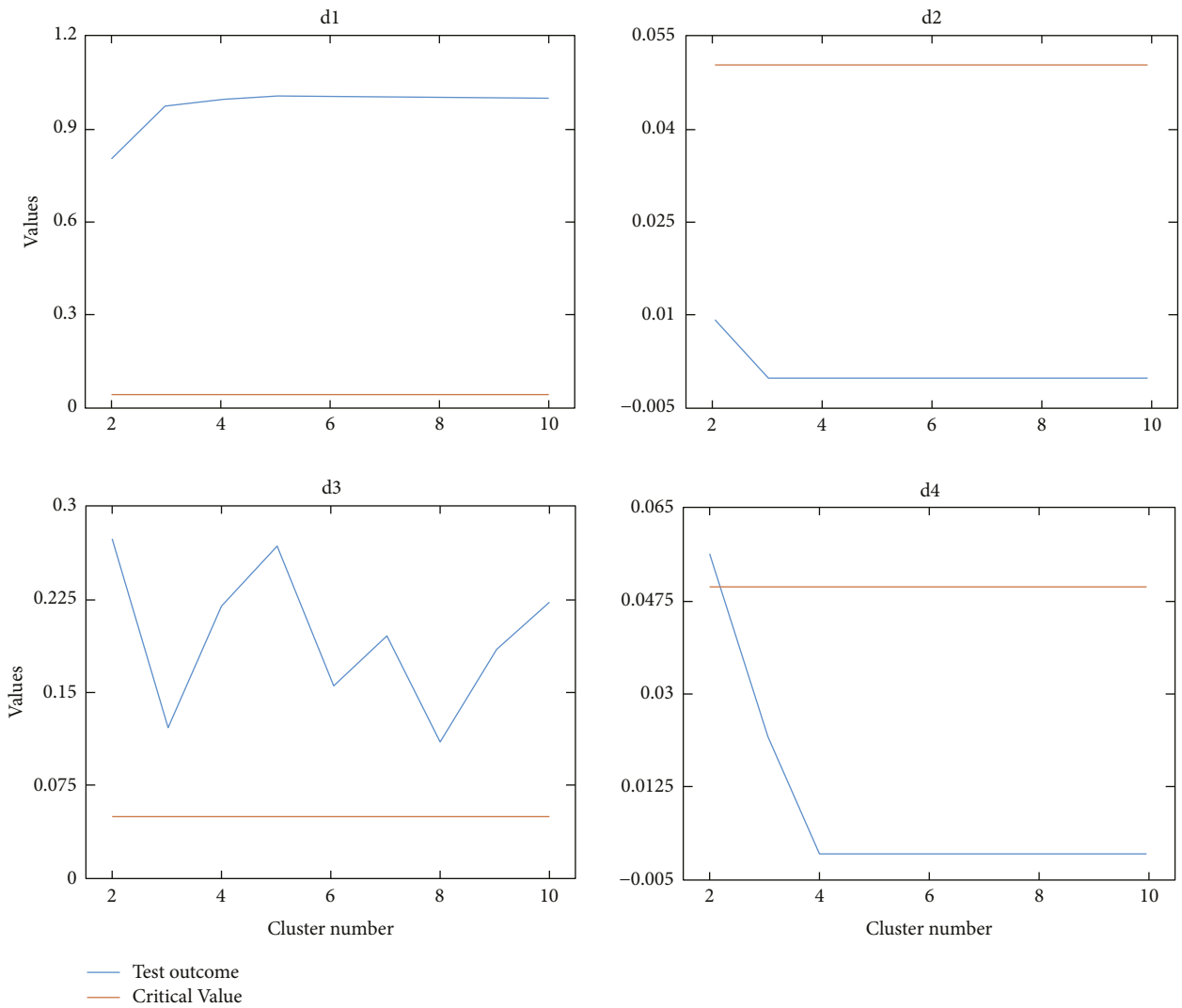


Figure 19: Impact of the number of clusters and time series effects on test results and critical values of Levene's test.

plots for d2 and d4, better and unbiased results are noted for a higher number of clusters as the trend effect is effectively highlighted by increasing the number of clusters.

4.5. Discussions. It can be noted from the obtained results that in some cases, the test outcome is contradictory, i.e., a nonstationary time series is represented as stationary or vice-versa. This can be due to the limitations in tests' working because they cannot characterize the nonstationary facets. If the test's working is entirely suitable for the desired application, calculating the power of a test would be beneficial to note how reliable the tests are in assisting the tests' selection [16]. Calculation of power of a test provides information on whether the test is completely efficient in distinguishing the given data as per its null and alternative hypotheses. A test cannot be considered reliable if it does not have good power. Also, using two tests capable of characterizing the same component in time series is not helpful. Further, it can be seen that all the tests cannot characterize all the nonstationary time series facets. Few of these tests also have biased results over changes in other attributes such as time series length or clustering and tests' parameters. For detecting trend components, all the tests except for Levene's test are capable of varying strength in the unbiased examination, but the best among these are KW and KS tests. After that, only Breitung and SW tests could detect seasonality in all datasets. Lastly, Levene's and SW's tests could only characterize the volatility effect. SW test can characterize all components effectively, and its results do not bias by changes in time series length. But, if the SW test generates a nonstationary outcome, it does not always mean that the time series is not stationary. The time series could have a stationary non-normal distribution. In that case, a combo of Breitung, Levene's, and KS/KW tests could be instrumental. The Breitung test can be considered with no deterministic components based on the results. Levene's test can be employed using small cluster sizes, more or less equal group

sizes, and higher cluster numbers. KW and KS tests can be used with any cluster size and cluster ratio, but a higher cluster number is preferable for the KW test.

Thus, the detailed result analysis and corresponding discussions would help a novice reader understand.

(1) Working models of various stationarity tests and the similarities and differences between them.

(2) Impacts of time series length and other time series facets such as trend, volatility, and seasonality on the test outcomes that assist in the selection of tests.

(3) The correct usage of clustering and various other test parameters helps achieve good efficiency in acquiring unbiased test outcomes.

(4) The best combination of tests that can be used to attain accurate information on the stationarity of a time series.

5. Conclusion

This research aimed to compare and examine time series stationarity tests, considering the effects of various time series facets to help the readers choose the best test for a given application. Nine well-established tests for stationarity were compared and analyzed thoroughly. The impacts of various time series effects on the test results with respect to various test parameters, time series lengths, and various types of time series clustering were taken into account. These impacts on the tests' performance were studied in detail by applying them to clean data and the same data with synthetically embedded trend and volatility effects. Test results and critical values of the tests were compared for all the above analyses that highlighted their capabilities in characterizing all these effects for various test parameters. The variations in capabilities of tests in portraying all these effects through their results for different test parameters helped in choosing these parameters for additional analysis. Based on the results and discussions, it is suggested to use the SW test first, and if a nonstationary outcome is generated by the test, then Breitung, Levene's, and KW/KS test is used. It is to mention here that a test's performance should be analyzed considering all possible parameters' settings to ascertain the optimal/appropriate parameter value to yield an unbiased test outcome. All the attributes considered in this paper are within certain limits. Possibly, more attributes exist that are not considered in this study. Further, the obtained test results may be biased. Critical research establishing the above future scope would add one more dimension to the judicious selection of stationarity tests for a particular application.

Nomenclature

ADF: Augmented Dickey–Fuller
ANOVA: Analysis of variance
AR: Autoregressive
ECDF: Empirical cumulative distribution function
GLS: Generalized least squares
KPSS: Kwiatkowski Phillips Schmidt Shin
KS: Kolmogorov Smirnov
KW: Kruskal Wallis
LM: Lagrange multiplier
OLS: Ordinary least squares
PP: Phillips Perron
PV: Photovoltaic
SW: Shapiro–Wilk
VR: Variance ratio
WGN: White Gaussian noise
$\widehat{\delta}_u$: Test statistic for Breitung test
ε: White noise
ζ , ϕ, ψ: Model parameters
$\theta(q)$: Variance term associated with Z-statistic for VR test
ρ: Parameter of lag 1 term σ
$\widehat{\rho}$: OLS estimator of parameter ρ
σ: Standard deviation
$\widehat{\sigma}_\rho$: OLS estimator of σ_ρ
σ^2: Variance
$\widehat{\sigma}_\varepsilon^2$: Long run estimate of variance σ_ε^2
a_i: Coefficient for i^{th} ordered data sample in SW test
c_α: Level of significance
d_t: Deterministic part of time series
i, j, l: Counting variable
k_1, k_2: Constant/intercept term and linear trend parameter respectively
r_i: i^{th} rank of data in KW test
n: Integer part of T/q in VR test; q is the period of VR test
n_l: Lag number for ADF test
$n_{l,\max}$: Maximum value of n_l
s_i: Seasonal interval or seasonal period
sgn(.): sign function used to extract sign of a given real number
$\widehat{u}$: Residual part of time series
$\overline{\overline{x}}$: Overall mean of all the groups in Levene's test
y: Time series data
$\overline{y}_i$: Mean of data in i^{th} group
$\widetilde{y}_i$: Median of data in i^{th} group
D_{n_1,n_2}: Test statistic for KS test
$E(.)$: Calculates expected value
G: Group size
H: Test statistic for KW test
H_0: Null hypothesis
H_A: Alternate hypothesis
$I(L)$: "L" order of integration (differencing of lag "L")
LM_s: LM statistic for KPSS test
P_D: Power consumption
T: Time series length
TS_ρ: Test statistic for ADF and PP tests
W: Test statistic for Levene's and SW tests
Z_{MK}: Test statistic for MK test

Data Availability

Data are openly available in a public repository [46] that does not issue DOIs.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] M. Fan, V. Vittal, G. T. Heydt, and R. Ayyanar, "Probabilistic power flow studies for transmission systems with photovoltaic generation using cumulants," *IEEE Transactions on Power Systems*, vol. 27, no. 4, pp. 2251–2261, 2012.

[2] B. R. Prusty and D. Jena, "A critical review on probabilistic load flow studies in uncertainty constrained power systems with photovoltaic generation and a new approach," *Renewable and Sustainable Energy Reviews*, vol. 69, pp. 1286–1302, 2017.

[3] D. D. Le, G. Gross, and A. Berizzi, "Probabilistic modeling of multisite wind farm production for scenario-based applications," *IEEE Transactions on Sustainable Energy*, vol. 6, no. 3, pp. 748–758, 2015.

[4] B. R. Prusty and D. Jena, "A sensitivity matrix-based temperature-augmented probabilistic load flow study," *IEEE Transactions on Industry Applications*, vol. 53, no. 3, pp. 2506–2516, 2017.

[5] B. R. Prusty and D. Jena, "A spatiotemporal probabilistic model-based temperature-augmented probabilistic load flow considering PV generations," *International Transactions on Electrical Energy Systems*, vol. 29, no. 5, pp. 28199–e2913, 2019.

[6] C. Chatfield, *Time-series Forecasting*, CRC Press, Boca Raton, Florida, 2000.

[7] S. Porter-Hudak, "An application of the seasonal fractionally differenced model to the monetary aggregates," *Journal of the American Statistical Association*, vol. 85, no. 410, pp. 338–344, 1990.

[8] B. G. Brown, R. W. Katz, and A. H. Murphy, "Time series models to simulate and forecast wind speed and wind power," *Journal of Climate and Applied Meteorology*, vol. 23, no. 8, pp. 1184–1195, 1984.

[9] D. Fedorová, "Selection of unit root test on the basis of length of the time series and value of AR (1) parameter," *Statistika*, vol. 96, p. 3, 2016.

[10] X. Zuo, "Several important unit root tests," in *Proceedings of the 2019 2nd International Conference on Information Communication and Signal Processing (ICICSP)*, pp. 10–14, IEEE, Weihai, China, September 2019.

[11] A. Lojowska, D. Kurowicka, G. Papaefthymiou, and L. van der Sluis, "Advantages of ARMA-GARCH wind speed time series modeling," in *Proceedings of the 2010 11th International Conference on Probabilistic Methods Applied to Power Systems*, pp. 83–88, IEEE, Singapore, June 2010.

[12] N. Masseran, A. M. Razali, K. Ibrahim, and W. Wan Zin, "Evaluating the wind speed persistence for several wind stations in Peninsular Malaysia," *Energy*, vol. 37, no. 1, pp. 649–656, 2012.

[13] C. Modin, "Short-term wind power forecasting," *Doctor, Economy and Society*, Dalarna University, Falun, Sweden, 2009.

[14] I. Ebtehaj, H. Bonakdari, M. Zeynoddin, B. Gharabaghi, and A. Azari, "Evaluation of preprocessing techniques for improving the accuracy of stochastic rainfall forecast models," *International journal of Environmental Science and Technology*, vol. 17, no. 1, pp. 505–524, 2020.

[15] D. Yang, "Choice of clear-sky model in solar forecasting," *Journal of Renewable and Sustainable Energy*, vol. 12, no. 2, Article ID 026101, 2020.

[16] A. A. Bawdekar and B. R. Prusty, "Selection of stationarity tests for time series forecasting using reliability analysis," *Mathematical Problems in Engineering*, vol. 20228 pages, Article ID 5687518, 2022.

[17] F. N. Melzi, T. Touati, A. Same, and L. Oukhellou, "Hourly solar irradiance forecasting based on machine learning models," in *Proceedings of the 2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pp. 441–446, IEEE, Anaheim, CA, USA, December 2016.

[18] M. Diagne, M. David, P. Lauret, J. Boland, and N. Schmutz, "Review of solar irradiance forecasting methods and a proposition for small-scale insular grids," *Renewable and Sustainable Energy Reviews*, vol. 27, pp. 65–76, 2013.

[19] S. Atique, S. Noureen, V. Roy, V. Subburaj, S. Bayne, and J. Macfie, "Forecasting of total daily solar energy generation using ARIMA: a case study," in *Proceedings of the 2019 9th annual computing and communication workshop and conference (CCWC)*, pp. 0114–0119, IEEE, Las Vegas, NV, USA, January 2019.

[20] R. P. Silva, B. B. Zarpelão, A. Cano, and S. B. Junior, "Time series segmentation based on stationarity analysis to improve new samples prediction," *Sensors*, vol. 21, p. 7333, 2021.

[21] L. Lijuan, L. Hongliang, W. Jun, and B. Hai, "A novel model for wind power forecasting based on Markov residual correction," in *Proceedings of the IREC2015 6th International Renewable Energy Congress*, pp. 1–5, IEEE, Sousse, Tunisia, March 2015.

[22] J. Wang, Q. Zhou, and X. Zhang, "Wind power forecasting based on time series ARMA model," in *Proceedings of the 2018 IOP Conference Series: Earth and Environmental Science*, Banda Aceh, Indonesia, September 2018.

[23] S. Gao, Y. He, and H. Chen, "Wind speed forecast for wind farms based on ARMA-ARCH model," in *Proceedings of the 2009 International Conference on Sustainable Power Generation and Supply*, pp. 1–4, IEEE, Nanjing, China, April 2009.

[24] H. Chen, Q. Wan, F. Li, and Y. Wang, "GARCH in mean type models for wind power forecasting," in *Proceedings of the 2013 IEEE Power & Energy Society General Meeting*, pp. 1–5, IEEE, Vancouver, BC, July 2013.

[25] X. Hu, J. Jaraite, and A. Kažukauskas, "The effects of wind power on electricity markets: an evaluation using the Swedish electricity market data," Department of Economics, Umeå University, Umeå, Sweden, 2020.

[26] N. M. Razali and Y. B. Wah, "Power comparisons of shapiro-wilk, Kolmogorov-smirnov, lilliefors and anderson-darling tests," *Journal of statistical modeling and analytics*, vol. 2, no. 1, pp. 21–33, 2011.

[27] J. Li, G. Lyu, and H. Zhang, "Characteristics and forecast of short-term wind speed series in the Donghai Bridge wind farm," *SCIENTIA SINICA Physica, Mechanica & Astronomica*, vol. 46, no. 12, Article ID 124713, 2016.

[28] N. Bokde, A. Feijóo, N. Al-Ansari, S. Tao, and Z. M. Yaseen, "The hybridization of ensemble empirical mode decomposition with forecasting models: application of short-term wind speed and power modeling," *Energies*, vol. 13, no. 7, p. 1666, 2020.

[29] N. H. Hussin, F. Yusof, R. Jamaludin, and S. M. Norrulashikin, "Forecasting wind speed in peninsular Malaysia: an application of ARIMA and ARIMA-GARCH

models," *Pertanika Journal of Science and Technology*, vol. 29, no. 1, 2021.

[30] H. Wilms, M. Cupelli, and A. Monti, "On the necessity of exogenous variables for load, pv and wind day-ahead forecasts using recurrent neural networks," in *Proceedings of the 2018 Electrical Power and Energy Conference (EPEC)*, pp. 1–6, IEEE, Toronto, ON, Canada, October 2018.

[31] N. Odam and F. P. de Vries, "Innovation modelling and multi-factor learning in wind energy technology," *Energy Economics*, vol. 85, Article ID 104594, 2020.

[32] J. L. Carrion-i-Silvestre and A. Sansó, "A guide to the computation of stationarity tests," *Empirical Economics*, vol. 31, no. 2, pp. 433–448, 2006.

[33] R. Gimeno, B. Manchado, and R. Minguez, "Stationarity tests for financial time series," *Physica A: Statistical Mechanics and Its Applications*, vol. 269, no. 1, pp. 72–78, 1999.

[34] J. H. Cochrane, "A critique of the application of unit root tests," *Journal of Economic Dynamics and Control*, vol. 15, no. 2, pp. 275–284, 1991.

[35] T. K. Kim, "T test as a parametric statistic," *Korean Journal of Anesthesiology*, vol. 68, no. 6, p. 540, 2015.

[36] R. C. Sprinthall, *Basic Statistical Analysis*, Pearson Education, London, UK, 9th. Edition, 2012.

[37] Y. W. Cheung and K. S. Lai, "Lag order and critical values of the augmented Dickey–Fuller test," *Journal of Business & Economic Statistics*, vol. 13, no. 3, pp. 277–280, 1995.

[38] J. Breitung, "Nonparametric tests for unit roots and cointegration," *Journal of Econometrics*, vol. 108, no. 2, pp. 343–363, 2002.

[39] H. Levene, "Robust tests for equality of variances." contributions to probability and statistics," *Essays in honor of Harold Hotelling*, pp. 278–292, Stanford University Press, Palo Alto, 1960.

[40] W. H. Kruskal and W. A. Wallis, "Use of ranks in one-criterion variance analysis," *Journal of the American Statistical Association*, vol. 47, pp. 583–621, 1952.

[41] D. A. Dickey and W. A. Fuller, "Distribution of the estimators for autoregressive time series with a unit root," *Journal of the American Statistical Association*, vol. 74, no. 366, pp. 427–431, 1979.

[42] G. Elliott, T. J. Rothenberg, and J. H. Stock, "Efficient tests for an autoregressive unit root," *NBER technical working papers series*, 1992.

[43] R. Simard and P. L'Ecuyer, "Computing the two-sided Kolmogorov-Smirnov distribution," *Journal of Statistical Software*, vol. 39, no. 11, pp. 1–18, 2011.

[44] D. Kwiatkowski, P. C. Phillips, P. Schmidt, and Y. Shin, "Testing the null hypothesis of stationarity against the alternative of a unit root: how sure are we that economic time series have a unit root?" *Journal of Econometrics*, vol. 54, no. 1-3, pp. 159–178, 1992.

[45] P. C. B. Phillips, "Time series regression with a unit root," *Econometrica*, vol. 55, no. 2, pp. 277–301, 1987.

[46] "Hourly load consumption," 2022, https://openei.org/datasets/files/961/pub.

[47] G. W. Schwert, "Tests for unit roots: a Monte Carlo investigation," *Journal of Business & Economic Statistics*, vol. 20, no. 1, pp. 5–17, 2002.

Hindawi
Mathematical Problems in Engineering
Volume 2022, Article ID 6583809, 12 pages
https://doi.org/10.1155/2022/6583809

Research Article

The Impact of Digital Economy on Industrial Carbon Emission Efficiency: Evidence from Chinese Provincial Data

Ning-Yu Xie and Yang Zhang

School of Economics and Management, Dalian Minzu University, China

Correspondence should be addressed to Yang Zhang; zhangyang@dlnu.edu.cn

Received 22 July 2022; Accepted 30 August 2022; Published 29 September 2022

Academic Editor: Chao Huang

Digital economy has become an important driving force for green economic growth in China. Based on the province-level data of China from 2003 to 2018, this paper constructed the Total-factor Nonradial Directional Distance Function (TNDDF) model to measure the carbon emission efficiency of industrial sector and discussed the impact of digital economy on carbon emission efficiency. Empirical analysis shows that the carbon emission efficiency of China's industrial sector is low, and there is obvious regional heterogeneity where the carbon emission efficiency of eastern China is higher than that of central and western China. Areas with high level of digital economy development have higher carbon emission efficiency, and digital economy is conducive to promoting energy conservation and pollution reduction in China's industrial sector. The optimal threshold interval of digital economy for promoting carbon emission efficiency is explored by means of threshold model. In view of this, the Chinese government should vigorously develop the digital economy, promote industrial enterprises to networking and digital evolution, and improve the efficiency of carbon emission as well.

1. Introduction

The continuous increasement of industrial carbon emissions has led to extreme climate problems such as global sea level rise, and it has become a consensus all over the world to improve carbon emission efficiency of industrial sector and reduce industrial greenhouse gas emissions [1, 2]. Since reform and opening-up, China's economy has maintained a medium-high growth rate, with its GDP reaching more than 10 trillion yuan in 2020. China's industrialization process and its rapid development of industrial economy have provided important impetus for the miracle of China's sustained economic growth. At the same time, resource exhaustion and environmental problems caused by excessive energy consumption have become a major problem facing China's economic development. In 2016, the Chinese government put forward its voluntary emission reduction commitment in the Paris Agreement, striving to achieve peak carbon dioxide emissions in 2030, of which the carbon emission intensity is reduced by 60% to 65% compared with 2005. Furthermore, the 14th Five-Year Plan promulgated by China proposed the idea of 'spurring green development and promoting harmonious coexistence between man and nature, which stresses that industrial transformation can promote green economic growth. Following the plan, by 2035, green production lifestyles will be widely formed, and carbon emissions will be steadily dropped after the peak, thereby striving to achieve the goal of building a beautiful China and a deep integration of digital and real economy. At the same time, relying on the rapid development of digital technologies such as big data and the Internet, China's digital economy has penetrated into all aspects of enterprise production and residents living. It also shows great potential in improving the production efficiency of industrial enterprises, optimizing the industrial structure, and improving the misallocation of resources [3]. Based on the total factor efficiency theory, carbon emission efficiency of industrial sector reflects the maximum economic output and minimum undesirable output like carbon emissions that can be obtained by industrial enterprises in terms of given capital,

labor, and energy input. The higher the carbon emission efficiency of industrial sector is, the more conducive it is to promoting the green development of the industry. Therefore, it is of great practical significance to study the impact of digital economy on the carbon emission efficiency of China's industrial sector.

In the present context of global low-carbon and ecological civilization construction, digital economy has become an important force to promote sustainable economic and social development, making it attract great attention from governments around the world [4]. And as an economic power, China has also entered the era of digital economy. According to the "Digital Economy Report 2021" released by the United Nations Conference on Trade and Development (UNCTAD), China and the United States are the best positioned to participate in and benefit from the digital economy. In accordance with the "White Paper on China's Digital Development (2021)" released by China Academy of Information and Communications Technology, the scale of China's digital economy has increased from 11 trillion yuan at the beginning of the 13th Five-Year Plan to 39.2 trillion yuan in 2020, accounting for 38.6% of GDP. Digital economy has become the core growth engine of the national economy. The global order is entering a new phase of adjustment, with digital technology deeply integrated with production and life, and the COVID-19 pandemic promoting the full penetration of the digital economy. China's 14th Five-Year Plan emphasizes the needs to continue to develop digital economy and promote digital industrialization and industrial digitalization, so as to realize the deep integration of digital economy and real economy. It can be said that digital economy has become an important driving force for China's green economic growth.

As a big energy consumer, Is a high level of digital economy able to promote carbon efficiency in the industrial sector? At what level of development can the digital economy be most conducive to promoting the carbon efficiency in the industrial sector? This paper is mainly focusing on the above problems and contributing to the existing research mainly from the following aspects. Firstly, the TNDDF model is used to calculate the carbon emission efficiency. Secondly, it discusses the green transformation of China's industrial sector from the perspective of digital economy. The development of digital economy has become an important channel for energy conservation and emission reduction, which is conducive to opening the black box of realizing green economic growth. Finally, in order to investigate the dynamic impact of digital economy on the carbon emission efficiency of industrial sector in a more comprehensive and in-depth way, the optimal intensity range of digital economy to stimulate the carbon emission efficiency of industrial sector is sought through nonlinear test with digital economy as the threshold variable.

The remaining parts of this paper are arranged as follows: the second part reviews the existing literature from the perspectives of the impact of digital economy on economic growth and green development. The third part calculates the carbon emission efficiency by selecting the carbon emission index and using the Total-factor Nonradial Directional Distance Function (TNDDF) model. The fourth part is the construction of threshold effect model and introduction to the relevant variables and data sources. The fifth part analyzes the empirical results, mainly from the basic, regional heterogeneity and threshold. The sixth part is to draw conclusions based on empirical analysis and put forward corresponding countermeasures and suggestions.

2. Literature Review

There are few literatures that directly analyze digital economy and carbon emission efficiency. This paper sorts out the existing literature from the perspectives of the impact of digital economy on economic growth and green development.

In the study of the influence of digital economy on economic growth, some scholars point out that digital economy has a positive role in promoting economic growth. Wu et al. [5] believe that digital economy is becoming a new driving force for global economic growth, and enterprises are increasingly relying on the Internet for production, operation, and sales. The Internet combines internal business processes with external business activities, reducing the cost of internal communication and information exchange through external communication [4, 6]. Perez and Lacalle [7] argue that the emergence of the Internet has changed the ability of global information sharing, while the knowledge and innovation it is releasing have a significant positive impact on economic growth, especially in developing countries. Jorgenson and Khuuong [8] pointed out that the notable feature of digital economy is the rapid development of e-commerce, which further promotes the construction of infrastructure such as information and communication, thereby accelerating the economic growth of a country. Wu et al. [9] analyzed the provincial panel data of China from 2006 to 2017 by using the dynamic spatial Dubin model, mediation effect model, and dynamic threshold panel model and pointed out that information and communication technology supported by the Internet has become an important driving force to promote the intelligent development of environmental governance in China. Deng and Zhang [10] used nonradial distance function (NDDF) to measure energy and carbon emission performance based on Chinese provincial data from 2006 to 2017. The bidirectional fixed effect model is constructed, and it is concluded that the development of Internet mainly improves energy and carbon emission efficiency by promoting industrial structure upgrading and technology diffusion. Wang et al. [11], based on OECD data, KPWW method, and multipanel regression, discusses the impact of digital technology innovation and technology spillover on domestic carbon emission intensity and its mechanism and points out that the information industry has become a 'new engine' to drive world economic growth. However, there is little research on the negative impact of the Internet on economic growth. Steffen et al. [12] believe that the influence of digitalization on energy consumption mainly includes inhibition effect and growth effect. The inhibition effect is mainly reflected in the fact that digitalization can reduce energy intensity by improving

energy use efficiency and optimizing industrial structure [13]. The mechanism of growth effect is mainly manifested in two aspects: the energy consumption directly increased by the production, use, and disposal of digital technology; and the energy demand indirectly caused by the economic development empowered by digital technology [5, 14].

The research on the relationship between digital economy and green development is also one of the hottest issues in academic circles in recent years [3, 15, 16]. Whether digital economy can promote green development is controversial at present, with most scholars believing that digital economy promotes green development. Wu and Zhang [17] found that China's technological Internet and platform Internet have significantly improved the green total factor productivity of forestry in the short term. As the Chinese government speeds up the 'Internet+' development plan, the integrative development of energy and Internet has become a new way of green urban development. Li et al. [18] evaluated the environmental effects of the Energy Internet demonstration project implemented by the Chinese government and found that the measure was improving the air quality of the pilot city and surrounding cities. Shahnazi and Shabani [19] found that there is a spatial spillover effect between the progress of Internet technology and environmental pollution. From the perspective of environmental detection, the development of Internet technology plays a significant role in reducing environmental pollution. Koomey et al. [20] pointed out that when digital technology spills over to the agricultural sector, manufacturing sector, and housing construction sector, it is also conducive to eliminating carbon emission levels. Ismagilova et al. [21] believe that the information system (IS) method promotes the intelligent use of information and communication technology (ICT), thereby providing optimized and advanced services for individuals as well as affecting their quality of life and the sustainable management of natural resources. Singh and Sahu [22] argue that green information technology and related innovations mainly make use of the contribution of technology to emission reductions, while ICT replaces the physical world. Ishida [23] estimated the long-term and short-term relationship between ICT and energy consumption in Japan from 1980 to 2010. The study found that ICT investments contribute to modest reductions in energy consumption in the long and short term. Schulte et al. [24], with the industrial panel data of OECD countries, and Khuntia et al. [25], with the cross-sectional data of Internet investment and energy consumption in India's manufacturing industry, believe that the development of internet can reduce energy consumption. Saidi et al. [26] studied the panel data of 67 countries in the world and found that the higher the ICT development level represented by the amount of Internet access and the number of mobile phone users, the higher the power consumption level. Salahuddin and Gow [27] studied the relationship between Internet development and energy consumption in Australia and found that there was no significant correlation between Internet development and carbon emissions in the short and long term. Lange and Pohl [12] found that the hope of digitalization to reduce energy consumption had not been proved, and digitalization did not save energy, while it increased energy consumption instead. Existing studies did not directly point out that the digital economy will hinder environmental pollution. On the contrary, some scholars believe that the Internet has not played a positive role in energy efficiency and energy consumption. Salahuddin and Alam [28] studied the long-term and short-term impacts of Internet information technology on electricity consumption by using OECD panel data from 1985 to 2012, but it shows that OECD countries have not yet achieved improvements in energy efficiency through the adoption of Internet information technology.

Above all, the review of the literature indicates that many scholars have made use of all kinds of analysis method to examine the impact of the digital economy on economic growth and economic development from different perspectives, and China's industrial sector currently has problems such as low-carbon emission efficiency, high pollution, and high energy consumption in industrial development [29–31]. These provide an important reference for the research of this paper. But there are still some deficiencies that need to be further improved. First of all, most scholars have conducted detailed studies on the relationship between digital economy and economic growth but have not incorporated digital economy and carbon emission efficiency into a unified research framework. Secondly, digital economy is an important driving factor of technological innovation, but the existing research has not considered the spatial spillover and nonlinear impact of digital economy on the efficiency of energy saving and emission reduction. Finally, few studies have taken China's industrial sector as the research object to analyze the impact of digital economy on the carbon emission efficiency of China's industrial sector. Therefore, our study helps clarify the relationship between digital economy and carbon emission efficiency and then provides pertinent suggestions for further promoting the development strategy of digital economy and helping China achieve green economic transformation.

3. Measuring Carbon Emission Efficiency of China's Industrial Sector

3.1. TNDDF Model. Since Chambers et al. [32] proposed a Directional Distance Function (DDF) that maximizes the expected output while minimizing the undesirable output based on the Data Envelopment Analysis (DEA) method, a large number of studies apply the directional distance function to seek the best combination of economic development and energy conservation and emission reduction. Based on the practice of China's green economy development, Zhang et al. [33] proposed the Total-factor Nonradial Directional Distance Function (TNDDF), which has the advantage of not only avoiding the problem of no solution in linear programming, but also achieving the comparability of intertemporal efficiency. Therefore, this paper will use the TNDDF function proposed by Zhang et al. [33] to measure the carbon emission efficiency of China's industrial sector.

Suppose that there are j decision-making units (DMUs), and each DMU uses the input factor $x \in R_+^m$ to obtain the desirable output $y \in R_+^n$ and undesirable output $b \in R_+^k$.

Then, according to the definition of Färe et al. [34], the environmental technology function $P(x)$ is

$$P(x)=\left\{\left(x^t,y^t,b^t\right):\ \sum_{j=1}^{J} z_j^t y_{jn}^t \le y_n^t, \sum_{j=1}^{J} z_j^t b_{jk}^t = b_k^t, z_j^t \ge 0, \quad \forall j,m,n,k, \right. \tag{1}$$

where z_j^t is the intensity variable, which assigns weights to each observation unit. The equal sign in the constraints means that the environmental technology function has weak disposability of output as well as zero binding characteristics. Then, according to Zhang et al. [33], the TNDDF function can be defined as

$$\overrightarrow{ND}\left(x,y,b;g_x,g_y,g_b\right)=\sup\left\{\omega^T\beta:\ \left(x-\beta_x g_x, y+\beta_y g_y, b-\beta_b g_b\right)\in P(x),\right. \tag{2}$$

where $\omega^T = (\omega^x, \omega^y, \omega^b)^T$ is the weight vector. According to the setting of Zhang et al. [33], the weights of inputs and desirable and undesirable outputs are all 1/3. In this study, the input factors include capital (K), labor (L), and energy (E), while the desirable output is the industrial added value (Y) of each province, and the undesirable output is the total carbon dioxide emissions of each city (S). Therefore, the weight vector in this study can be set as $\omega^T = (1/9, 1/9, 1/9, 1/3, 1/3)^T$, the slack variable as $\beta = (\beta_K, \beta_L, \beta_E, \beta_Y, \beta_S)^T \ge 0$, and the direction vector as $g = (-K, -L, -E, Y, -S)$.

Then, the TNDDF function can be solved as follows:

$$\begin{aligned}
\overrightarrow{ND}^G\left(x^t,y^t,b^t;g_x,g_y,g_b\right) &= \mathrm{Max}\,\omega_K\beta_K^{Gt}+\omega_L\beta_L^{Gt}+\omega_E\beta_E^{Gt}+\omega_Y\beta_Y^{Gt}+\omega_S\beta_S^{Gt},\\
s.t\ \sum_{j=1}^{J}\sum_{t=1}^{T} z_j^t Y_j^t &\ge \left(1+\beta_Y^{Gt}\right)Y_j^t, \sum_{j=1}^{J}\sum_{t=1}^{T} z_j^t K_j^t \le \left(1-\beta_K^{Gt}\right)K_j^t, \sum_{j=1}^{J}\sum_{t=1}^{T} z_j^t L_j^t \ge \left(1+\beta_L^{Gt}\right)L_j^t,\\
\sum_{j=1}^{J}\sum_{t=1}^{T} z_j^t E_j^t &\le \left(1-\beta_E^{Gt}\right)E_j^t, \sum_{j=1}^{J}\sum_{t=1}^{T} z_j^t S_j^t \ge \left(1+\beta_S^{Gt}\right)S_j^t,\\
z_j &\ge 0, j=1,2,\ldots J; \beta_K^{Gt}\ge 0, \beta_L^{Gt}\ge 0, \beta_E^{Gt}\ge 0, \beta_Y^{Gt}\ge 0, \beta_S^{Gt}\ge 0.
\end{aligned} \tag{3}$$

According to Lin and Zhu [35], carbon emission efficiency (CEE) can be defined as

$$CCE_{i,t}=\frac{1/2\left(1-\beta_{E,\mathrm{it}}^*\right)+1/2\left(1-\beta_{S,\mathrm{it}}^*\right)}{1-\beta_{Y,\mathrm{it}}^*}, \tag{4}$$

where $\beta^* = (\beta_L^*, \beta_K^*, \beta_E^*, \beta_Y^*, \beta_S^*)$ is the optimal solution for the slack variable in equation (4).

3.2. Indicators' Selection. This study takes 30 provinces in China as the research object and uses Matlab software to measure the carbon emission efficiency of the industrial sector from 2004 to 2017. The input factors in the TNDDF function are capital, labor, and energy. Drawing on the research of Jin and Shen [36], this paper uses the perpetual inventory method to estimate the capital stock of each province in China. The relevant data required for capital stock calculation are all from the China Statistical Yearbook (2004–2019) and transformed into constant prices with a base period of year 2000. The energy input and labor input data of the industrial sector in each province are from the China Industrial Statistical Yearbook (2004–2019).

Desirable output is measured by using the value added of the industrial sector in each province. The original data are from the China Industrial Statistical Yearbook (2004–2019) and converted to constant prices by taking year 2000 as the base period. According to the research of Guo and Yuan [37], the undesirable output is measured by using the total emissions of carbon dioxide, and the original data come from the China Environmental Statistics Yearbook (2004–2019).

3.3. Results of Carbon Emission Efficiency. Figures 1 and 2 show the CEE distribution in China in 2008 and 2017, respectively. It can be seen that the spatial distribution of CEE in China has strong regional regularity. The provinces with the highest CEE values are mainly concentrated in the eastern part of China and gradually distributed to the periphery. Further, Figure 3 depicts the change of CEE in terms of time trend.

As shown in Figure 3, during the whole observation period, the average carbon emission efficiency index in China was greater than 0.50, except in 2003 (SARS broke out in China in 2003, the secondary industry with industry as the main body increased significantly; especially the high-carbon manufacturing industry led to the decline of carbon emission efficiency), the average over the sample period was 0.603. And carbon emission efficiency is showing a growing trend, with an average annual growth rate of 2.6%. This is consistent with the general background of China's recent mandatory restrictions on reducing carbon emission intensity, increasing intensity of energy conservation and emission reduction, and economic development gradually entering a new normal. The Chinese government has issued a series of environmental policies focusing on carbon emission reduction. In 2011, the year 2011 carbon intensity target was officially included in China's 12th Five-Year Plan. In 2014, the Chinese government officially proposed that carbon emission could peak by around 2030. And in 2015, China set a target of reducing carbon emission intensity by 60% to 65% compared with 2005 by 2030. In the report of the 19th National Congress of the Communist Party of China, it is proposed to promote efficiency reform, improve total factor productivity, reduce carbon emissions, and improve carbon emission efficiency to achieve sustainable development.

From the regional perspective, there are great differences in the values of carbon emission efficiency index among the three regions of eastern, central, and western China in different periods. As can be seen from Figure 1, the eastern coastal area has the highest carbon emission efficiency value with a high of 0.980, mainly due to the rapid economic growth, energy structure, and technological progress in the

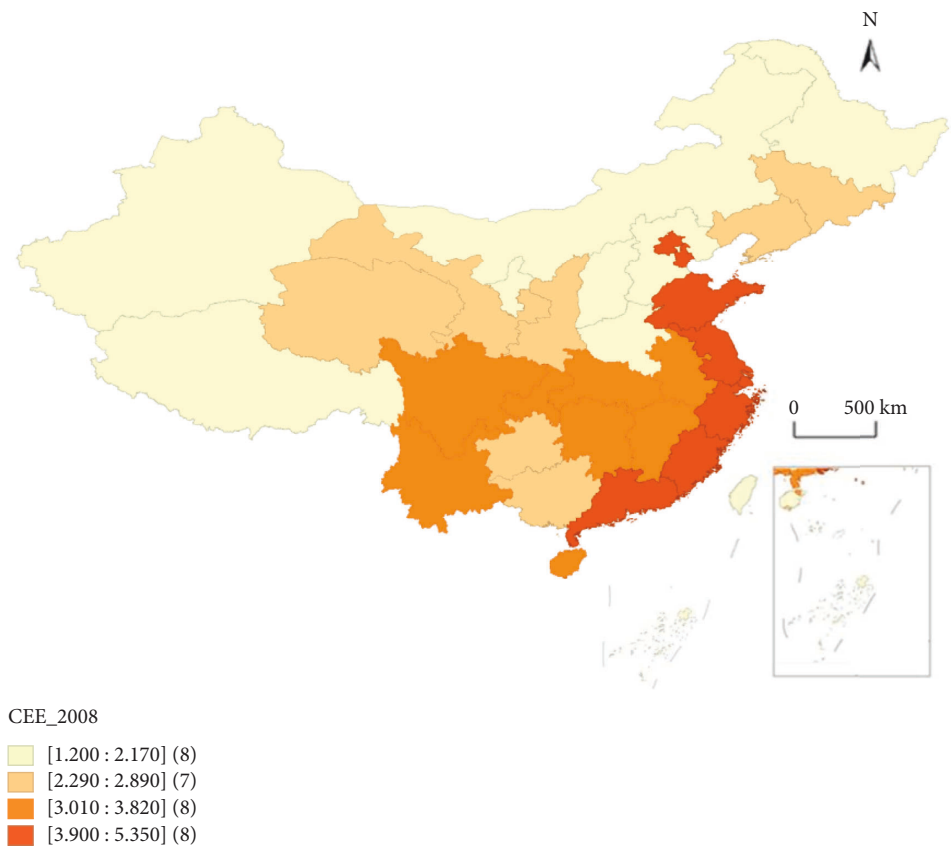


Figure 1: CEE of China in 2008.

eastern region. In particular, technological progress can reduce energy intensity and improve carbon emission efficiency to a certain extent. Meanwhile, part of the industrial gradient in the eastern region is transferred to the central and western regions. After 2010, the gap between the carbon emission efficiency of the central and western regions and that of the eastern regions has gradually increased; especially the carbon emission efficiency in western region began to decrease in 2011. This is largely related to the transfer of 'high-carbon' industries from the eastern regions to the central and western regions since 2010. The central and western regions mainly undertake the industrial transfer of high-polluting and high-energy-consuming industries from the eastern region. Although it is beneficial to improve the level of industrial development in the short term, such rapid economic development in the short term comes at the cost of environmental pollution, so the carbon emission efficiency gradually decreased in the long term. Therefore, the government agencies of central and western provinces should establish and improve the long-term mechanism of industrial carbon emission efficiency growth in practice and maintain the steady growth of carbon emission efficiency by means of continuously improving technological level.

In addition, this paper remeasures the carbon emission efficiency of China's industrial sector using the Data Envelopment Analysis (DEA) model that does not consider undesired outputs. According to Figure 4, it is not difficult to observe that the variation trend of industrial carbon emission efficiency in Eastern, Central, and Western China is consistent with the TNDDF model. However, it is worth noting that, before 2013, the carbon emission efficiency of western China was significantly higher than that of the central region, which is obviously contrary to the fact that western China has undertaken the transfer of a large number of pollution-intensive industries. Furthermore, the carbon emission efficiency results calculated by the DEA model are obviously high, which indicates that the results of the DEA model calculation without considering the undesirable output are deviation biased. This is also in line with the research conclusions by some scholars that the traditional carbon emission efficiency is higher than that considering the undesirable output. It shows that choosing the TNDDF model to measure the carbon emission efficiency of China's industrial sector is more scientific and robust.

4. Empirical Model and Variables

4.1. A Dynamic Panel Model. One of the core issues to be explored in this paper is the impact of digital economy on carbon emission efficiency of industrial sectors. Based on the theoretical analysis above, this paper constructs the following dynamic econometric model:

$$\ln \mathrm{CEE}_{i,t} = \alpha_0 + \beta_1 \ln \mathrm{CEE}_{i,t-1} + \beta_2 \ln \mathrm{Digital}_{i,t} + \sum \delta \ln X_{i,t} + \varepsilon_{i,t}, \tag{5}$$

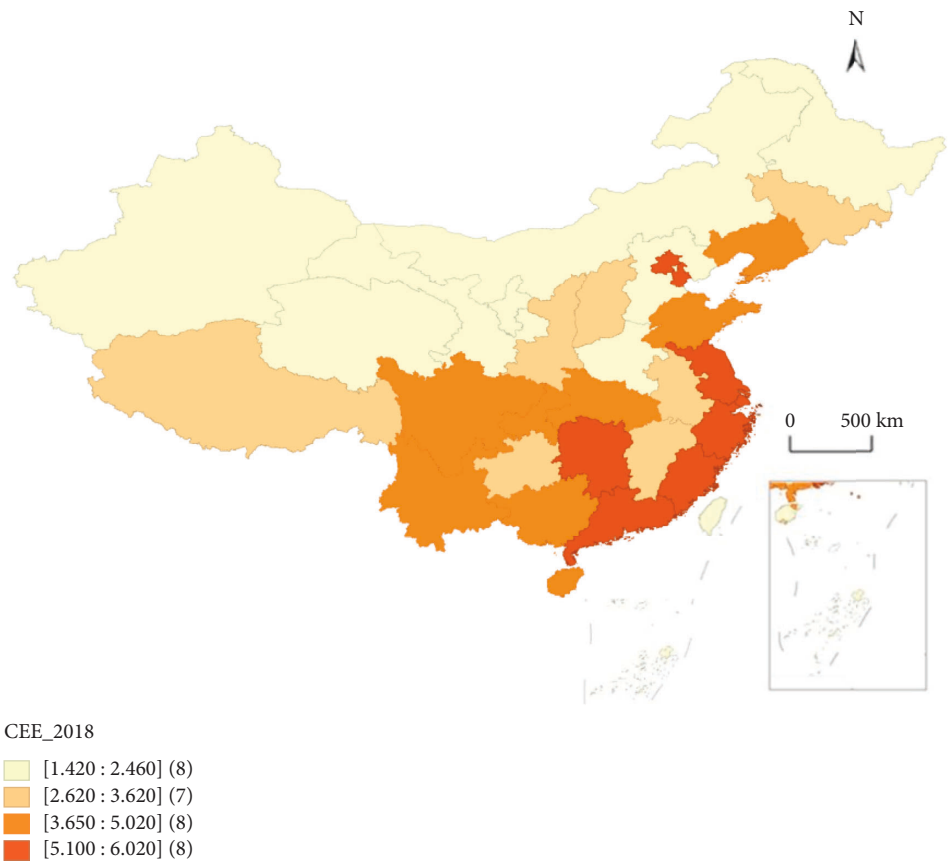


Figure 2: CEE of China in 2018.

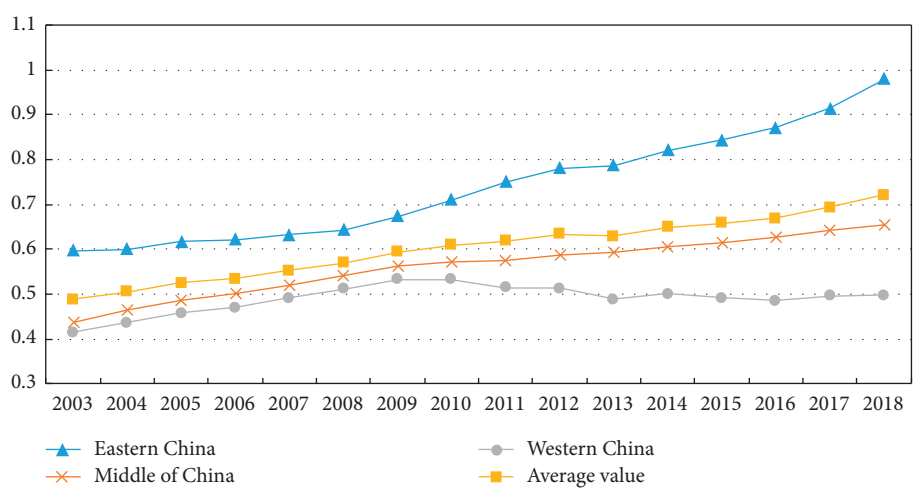


Figure 3: Trend of CEE of the three economic zones in China.

where i denotes the region, and t denotes the year. CEE represents carbon emission efficiency of China's industrial sector. Digital represents the Digital economy development index. X represents the control variable selected based on existing literature: specifically, foreign direct investment (FDI), capital deepening (Inv), export trade (Ex), gross domestic product (GDP), and government size (Gov). ε represents the error term, and ln represents taking the natural log.

The existing studies have not reached a consensus on the view that China's digital economy can improve the carbon emission efficiency of some industries, indicating that digital economy may have a nonlinear impact on the carbon emission efficiency of China's industrial sector. Therefore, in order to investigate the dynamic impact of digital economy on carbon emission efficiency of industrial sector in a more comprehensive and in-depth manner, especially to examine the optimal intensity range of the carbon emission efficiency

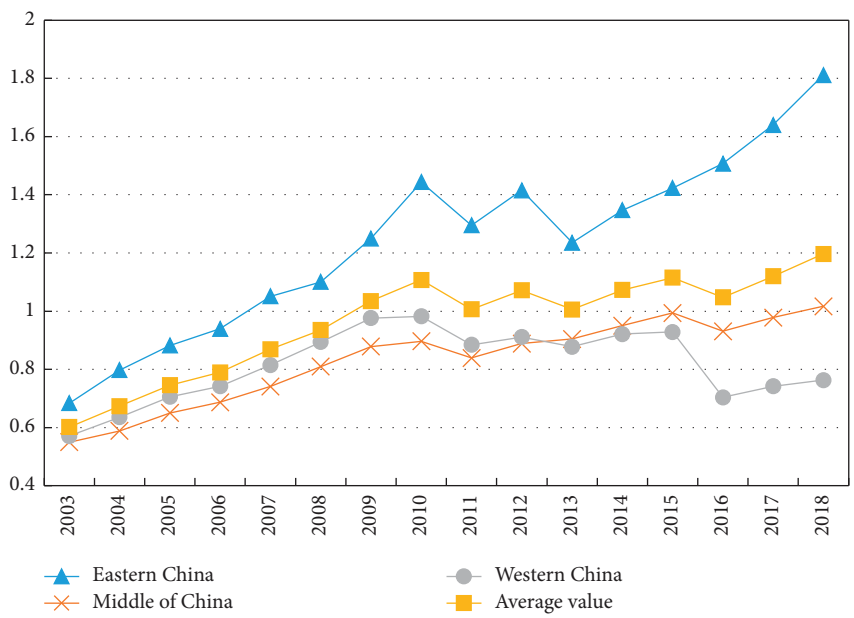


Figure 4: CEE trend of China's three economic zones measured by DEA model.

of the industrial sector where the digital economy stimulates environmental regulation and high-quality economic effects, this paper further takes the digital economy as the threshold variable and adopts the method of Hansen [38] to test the threshold effect. The specific model is set as follows:

$$\begin{aligned}\ln CEE_{i,t} &= \alpha_0 + \beta_1 \ln CEE_{i,t-1} + \theta_1 \ln \text{Digital}_{i,t} \times I\left(\ln \text{Digital}_{i,t} \le s_1\right) + \\ \theta_2 \ln \text{Digital}_{i,t} &\times I\left(s_1 \le \ln \text{Digital}_{i,t} \le s_2\right) + \ldots + \theta_n \ln \text{Digital}_{i,t} \times I\left(\ln \text{Digital}_{i,t} \ge s_n\right) + \\ &+ \sum \delta \ln X_{i,t} + \varepsilon_{i,t}, \\ \text{In}CEE_{i,t} &= a_0 + \beta_1 \text{In}CEE_{i,t-1} + \theta_1 \text{InDigital}_{i,t} \times I\left(\text{InDigital}_{i,t} \le s_1\right) + \theta_2 \text{InDigital}_{i,t} \times I\left(s_1 < \text{InDigital}_{i,t} \le s_2\right) + \\ &\ldots + \theta_n \text{InDigital}_{i,t} \times I\left(\text{InDigital}_{i,t} \ge s_n\right) + \sum .\delta \text{In}X_{i,t} + \varepsilon_{i,t},\end{aligned} \tag{6}$$

where $I(\cdot)$ is indicator function, and X is the set of control variables in the above baseline regression.

4.2. Variable and Data Source. At present, there are two main categories for measuring the development level of digital economy: single index measurement method and comprehensive index measurement method. The disadvantage of single index measurement method is that the research results may be biased due to the subjectivity of index selection or fail to fully reflect the development level of Internet in various regions due to the one-sidedness of index selection. In view of this, this paper draws on the research methods of Han et al. [39] to construct a comprehensive index of digital economy development by using the number of fixed broadband Internet users, mobile phone users, and general Internet users per 100 people. The specific measurement method of the comprehensive index of digital economy development is as follows:

$$\text{Digital}_{\text{it}} = \left(\text{net}_{\text{it},1} \times \text{net}_{\text{it},2} \times \text{net}_{\text{it},3}\right)^{1/3}, \tag{7}$$

where $\text{net}_{it,\ 1}$, $\text{net}_{it,\ 2}$, $\text{net}_{it,\ 3}$ successively represent the number of fixed broadband Internet users per 100 people, the number of mobile phone users per 100 people, and the number of general Internet users per 100 people in region i during period t. After geometric weighted average of the three indicators, the comprehensive index Digital of Digital economy development can be obtained.

There are 5 control variables in this study. First, regarding the foreign direct investment (FDI), this paper chooses the ratio of foreign direct investment to regional GDP to measure it; second, regarding the capital deepening (Inv), it is measured by the ratio of the net fixed assets of each region to its GDP; third, regarding the export trade (Ex), it is measured by the proportion of the total import and export volume of each region to the region's GDP; fourth, regarding the government size (Gov), it is measured by the ratio of government consumption to GDP of each region; finally,

Table 1: Summary statistics of main variables.

Variables	Mean	Sd	Min	Max
EQ	0.5681	0.0927	0.4933	0.8694
Digital	0.4323	0.0708	0.3990	0.7202
Tech	6.1222	0.0633	5.2037	7.0166
Allocation	0.1653	0.7112	0.0033	0.2191
FDI	69.7877	2.6604	45.5537	89.1336
Inv	2.8352	0.0273	0.7899	4.5385
Ex	9.1754	0.1644	2.1082	14.7367
Gov	0.1081	0.0097	0.0042	0.1691
GDP	11.8736	0.8566	2.1083	14.8766

Table 2: Dynamic panel data regression results.

	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)	Model (6)
Cons_	1.051^{***} (0.035)	1.052^{***} (0.034)	1.051^{***} (0.035)	1.054^{**} (0.032)	1.057^{*} (0.031)	1.150^{***} (0.039)
LnDigital_{it}	0.114^{***} (0.005)	0.145^{***} (0.004)	0.145^{***} (0.004)	0.117^{***} (0.003)	0.120^{***} (0.001)	0.205^{***} (0.002)
Lnfdi_{it}		0.012^{**} (0.000)	0.014 (0.001)	0.012^{***} (0.003)	0.015^{*} (0.000)	0.011^{***} (0.004)
LnInv_{it}			0.052^{***} (0.115)	0.054^{***} (0.107)	0.055 (0.110)	-0.056^{**} (0.111)
LnEx_{it}				-0.094^{**} (0.051)	-0.095^{*} (0.051)	-0.101^{***} (0.050)
Lngdp_{it}					0.135^{**} (0.005)	0.139^{*} (0.005)
LnGov_{it}						0.022^{***} (0.054)
AR(1)	0.0155	0.0117	0.0154	0.0015	0.0015	0.0025
AR(2)	0.1501	0.5255	0.4045	0.5714	0.4555	0.1512
Sargan	0.2021	0.4255	0.1750	0.4015	0.2557	0.2251

Notes. Standard errors are shown in brackets; ***, ** and *, respectively, represent significance at 1%, 5%, and 10% levels.

regarding the gross domestic product (GDP) is measured by the per capita GDP of each region.

In this study, panel data of 30 provinces in China from 2003 to 2018 were obtained from China Urban Statistical Yearbook, China Demographic Statistical Yearbook, China Industrial Statistical Yearbook, China Science and Technology Statistical Yearbook, etc. Based on the scientific principle of research samples, this paper uses household consumption index, industrial producer price index, GDP deflator, etc. to reduce all monetary quantities to comparable prices in the base period of 2000. At the same time, in order to avoid the impact of dimensional and order of magnitude differences of variables on test results, the original values of panel data in this paper are standardized. The summary statistics of main variables are shown in Table 1.

5. Empirical Results and Discussions

5.1. Baseline Regression Results. Table 2 reports the basic model regression results of the impact of digital economy on carbon emission efficiency of China's industrial sectors. According to model (1) in Table 2, without the inclusion of control variables, the estimated coefficient of the digital economy variable (Digital) is 0.114, which is significant at the 1% level, indicating that every 1% increase in the development level of Digital economy, the Carbon emission efficiency will increase by 11.4%. Meanwhile, models (2)–(6) further incorporate control variables based on Model (1). We can find that the estimated coefficient values and significance of digital economy variables have not fundamentally changed, and the digital economy is still positively promoting the carbon emission efficiency of China's industrial sector. This also verifies the first hypothesis where "digital economy can effectively stimulate the carbon emission efficiency of the industrial sector." In terms of control variables, GDP, Inv, Gov, and FDI are positively correlated with carbon emission efficiency, indicating that rising incomes, accelerating capital accumulation, government support, and foreign-invested factories are important drivers of carbon emission efficiency, which can enhance the infrastructure construction of the digital economy. In addition, there is a significant negative correlation between Ex and carbon emission efficiency. It is because that the existing international trading system is mainly biased in favor of the economically advanced countries, and the negative attitude of developed countries to foreign trade is not conducive to improving the carbon emission efficiency of China's industrial sector.

5.2. Regional Heterogeneity Analysis. Table 3 reports the regional heterogeneity test results of the digital economy on the carbon emission efficiency of the industrial sector. The results from Model (1) to Model (4) show that digital economy has a significant promoting effect on carbon emission efficiency in eastern and central China, where the influence coefficient of eastern region is the highest, which is 0.2528, and is much higher than 0.0767 of central province. Model (5) and Model (6) show that digital economy has a significant negative impact on carbon emission efficiency in western China, while export trade and income levels have a positive impact on it, which is opposite to that in eastern China. The results confirm that digital economy has significant regional differences in carbon emission efficiency.

Table 3: Regional heterogeneity analysis results.

	Eastern region		Central region		Central region	
	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)	Model (6)
Cons_	1.2476*** (0.042)	1.2642*** (0.046)	1.7224*** (0.001)	1.7788*** (0.002)	1.0524*** (0.052)	1.0257*** (0.065)
$LnDigital_{it}$	0.2528** (0.161)	0.2512** (0.174)	0.0767** (0.102)	0.0572* (0.105)	−0.0048*** (0.180)	−0.0022* (0.181)
$Lnfdi_{it}$		0.0642*** (0.066)		0.0812** (0.115)		−0.0762 (0.017)
$LnInv_{it}$		0.0177*** (0.007)		0.0058*** (0.077)		0.0056*** (0.006)
$LnEx_{it}$		−0.1250*** (0.027)		−0.0761** (0.051)		0.0682* (0.052)
$LnGov_{it}$		0.0002 (0.002)		−0.0464 (0.001)		−0.0524 (0.002)
$Lngdp_{it}$		−0.7146*** (0.164)		−0.0814** (0.247)		0.0685 (0.264)
AR(1)	0.0445	0.0411	0.0218	0.0041	0.0082	0.0091
AR(2)	0.1409	0.2001	0.2126	0.6802	0.1986	0.5984
Sargan	0.6626	0.4466	0.2541	0.4259	0.1109	0.2066

Notes. Standard errors are shown in brackets; ***, ** and * respectively represent significant at 1%, 5% and 10% levels.

Table 4: Threshold estimators from sampling test results.

Threshold variable	The threshold number	Estimated value	F	10%	5%	1%
Digital	Single	10.753***	27.51	23.578	26.716	34.692
	Double	11.541**	25.50	20.061	25.365	35.601
	Triple	9.707	9.13	19.502	24.163	37.714

The authors believe that the main reasons for the above results are as follows: China's regional industrial development gap is relatively large, and the industrial layout has also undergone fundamental changes in recent years, showing a trend of gradient transfer of industries from the east to the central and western. The Internet + initiative was launched earlier in the eastern region, and the development level of the digital economy is relatively high. Energy-consuming industries have been moving to the central and western regions. During the implementation and promotion of the strategy of the rise of central China, the central region has gradually increased the construction of network infrastructure and improved the allocation level of regional network resources. Overall, the eastern and central regions have achieved an effective connection for the development of the digital economy, which makes the digital economy in the eastern and central regions promote carbon emission efficiency. Due to the relatively lagging development of the Internet and the low level of digital economy in western China, its economic development mainly depends on government policies and investment. At the same time, the promotion process of digital economy is accompanied by the undertaking of energy consumption industry, which makes the development of digital economy in western China show the result of restraining carbon emission efficiency.

Above all, carbon emissions are closely related to economic development, and economic development requires energy consumption. China proposes two-stage carbon reduction targets (dual carbon strategic goal), to peak carbon dioxide emissions by 2030 and become carbon neutral by 2060. This goal will not be achieved without the help of digital development.

5.3. Threshold Effects of Digital Economy. According to the above theoretical analysis, there may be a nonlinear relationship between digital economy and carbon emission efficiency. In order to further verify the correlation between the two, this paper adopts the threshold panel technology proposed by Hansen to set digital economy as the threshold variable, sampling 500 times and estimating the critical value for F statistic and threshold variable. Table 4 reports the test results of digital economy as the threshold variable, where the digital economy variable presents its characteristics of significant under single threshold and double threshold while being not significant under triple threshold. The estimated value of single threshold is 10.753, and the estimated value of double threshold is 11.541. The dynamic threshold effect under different digital economy levels can be analyzed according to the different intervals divided by the double threshold.

As can be seen from the threshold regression results reported in Table 5, when LnDigital <10.753, the estimated coefficient is −0.021, which is significantly negative. In the sample grouping data of this paper, Anhui, Heilongjiang, Jiangxi, Gansu, and other western provinces with relatively

Table 5: Threshold model regression.

Variable	Estimated value	Standard error	P
LnDigital (LnDigital < 10.753)	-0.021^{*}	0.0045	0.000
LnDigital (10.753 ≤ LnDigital ≤ 11.541)	0.098^{***}	0.0895	0.082
LnDigital (LnDigital > 11.541)	0.152^{*}	0.0043	0.226
Lnfdi_{it}	0.042^{***}	0.0042	0.000
LnInv_{it}	0.026	0.0586	0.225
LnEx_{it}	-0.071^{**}	0.0056	0.043
LnGov_{it}	0.088^{***}	0.0402	0.007

Notes. ***, ** and *, respectively, represent significance at 1%, 5%, and 10% levels.

Table 6: Robustness checks.

	Model (1)	Model (2)	Mode l(3)	Model (4)	Model (5)	Model (6)
Cons_	2.142^{***} (0.134)	2.140^{***} (0.133)	2.143^{***} (1.134)	2.144^{**} (1.131)	2.147^{*} (1.132)	2.241^{***} (1.136)
LnDigital_{it}	1.114^{***} (0.002)	1.142^{***} (0.004)	1.142^{***} (0.004)	1.117^{***} (0.003)	1.110^{***} (0.001)	1.102^{***} (0.001)
Lnfdi_{it}		0.030^{**} (0.001)	0.040 (0.003)	0.043^{***} (0.001)	0.042^{*} (0.000)	0.038^{***} (0.002)
LnInv_{it}			-0.072^{**} (0.022)	-0.074^{*} (0.019)	−0.075 (0.020)	-0.077^{*} (0.024)
LnEx_{it}				0.022 (1.001)	-0.065^{**} (1.012)	-0.063^{**} (1.005)
Lngdp_{it}					0.651^{***} (0.001)	0.659^{***} (0.002)
LnGov_{it}						0.191^{**} (0.012)
AR(1)	0.0412	0.0447	0.0413	0.0241	0.0341	0.0141
AR(2)	0.4204	0.1413	0.4042	0.1743	0.4131	0.4146
Sargan	0.3043	0.4211	0.3710	0.2041	0.4127	0.3514

Notes. Standard errors are shown in brackets; ***, ** and *, respectively, represent significance at 1%, 5%, and 10% levels.

backward economic development level are in this range, indicating that, with the improvement of digital economy, the carbon emission efficiency of these regions decreases. In recent years, the development level of digital economy in western China lags far behind the central and eastern regions, and there is a huge "digital gap" between western China and developed regions. Although the region vigorously promotes the integrated development of 'Internet +' and traditional manufacturing, it mainly combines industrial enterprises with heavy pollution. Therefore, with the continuous development of digital economy, carbon emission efficiency is low.

When 10.753 ≤ LnDigital ≤ 11.541, the estimated coefficient is 0.098, indicating that, with the improvement of digital economy, carbon emission efficiency increases significantly. Fujian, Guizhou, Sichuan, and other provinces in central China are in this range. Although the low-carbon industry in central China started later than that in eastern China, its low-carbon industry develops at a faster speed, with higher average annual growth rate of relevant indicators of low-carbon industry development and greater potential for industrial low-carbon development. Taking Guizhou as an example, according to the 2019 China Big Data Industry Development Report released by China Industrial Information Security Development Research Center, the big data industry development Index of Guizhou is 76, ranking third in the country after Beijing and Guangdong.

When LnDigital >11.541, the estimated coefficient is 0.152, which is significantly positive. The possible reason is that the six samples in this region, that is, Beijing, Tianjin, Shanghai, Jiangsu, Zhejiang, and Guangdong, are all relatively developed eastern provinces and cities. These provinces have already played a leading role in the development of digital economy. In the development process of digital economy, it has successfully seized the initiative in e-commerce model, digital supply chain, etc. and pioneered the integration of the digital economy with the physical industry. The industries initially applied the integration with high pollution and high energy consumption have been gradually transferred to the central and western regions. At present, the development of digital economy is mainly integrated with high-end manufacturing such as artificial intelligence. The higher the development level of digital economy, the higher the carbon emission efficiency, and this is significantly correlated.

5.4. Robustness Checks. In order to ensure the robustness of the research results, the green total factor productivity of the industrial sector, which emphasizes the "win-win situation between the environment and the economy," is selected as a proxy indicator to measure the explained variables, and the model is reestimated. As shown in Table 6, the results of reestimating the model show that the digital economy has a positive and significant impact on the green total factor productivity of China's industrial sector, supporting the research conclusion that developing digital economy is conducive to energy conservation and emission reduction in China's industrial sector. Thus, the empirical results of substitution variables have verified the robustness of the previous research conclusion.

6. Conclusion

With the improvement of the development level of digital economy, its development level has affected the carbon emission efficiency of China's industrial sector.

Based on existing research, the digital economy is seen as increasingly important for improving the efficiency of carbon emissions. However, there are few studies on the impact of the digital economy on the carbon emission efficiency of China's industrial sector. Based on the data of 30 Chinese provinces from 2003 to 2018, this paper examines the impact of digital economy on carbon emission efficiency from both linear and nonlinear perspectives and conducts a threshold test on the heterogeneous impact of digital economy in different regions on carbon emission efficiency of China's industrial sector.

The development of digital economy inevitably produces unnecessary output, which means that although it is expected that its development can make GDP continue to grow, it is accompanied by a certain carbon emission. Therefore, the TNDDF model can be used to measure the carbon emission efficiency more effectively. From the perspective of regional carbon emission efficiency, China's carbon emission efficiency is currently at a low level, and there are significant regional differences. The development level of digital economy in eastern China is higher than that in central and western China, and its carbon emission efficiency is also significantly higher. It can be said that digital economy plays a significant role in promoting carbon emission efficiency. In order to find the optimal interval for its development, this paper conducts threshold effect analysis and concludes that although double threshold has passed the significance test, only in the appropriate interval, the development of digital economy will contribute to the growth of carbon emission efficiency. At present, the most significant interval is the eastern coastal region.

According to the research conclusions of this paper, the following policy implications can be obtained. First, steadily improve the development level of the digital economy, give full play to its incentive effect on carbon emission efficiency, and strive to narrow the information gap and "digital divide" between regions. Secondly, the Chinese government should fully consider the regional and industrial heterogeneity of real industry development and formulate targeted policies for digital economy development. The eastern region should maintain a reasonable and scientific speed of digital economic development, while the central and western regions should promote the application of new digital technologies such as big data and cloud computing in real industry and strive to narrow the "digital divide" between regions and industries. Third, the deep integration of digital economy and real industry is an effective way to improve China's carbon emission efficiency. On the one hand, China should promote collaborative innovation platforms with the digital economy as the carrier and promote the public sharing and rational flow of innovative elements such as knowledge, information, technology, and talent among high and small schools. On the other hand, through the comprehensive promotion and application of new technologies and new models, industrial enterprises are promoted to network and digital evolution, thus improving the carbon emission efficiency.

Data Availability

As the paper used unofficial government data such as carbon emissions in its analysis, it is requested not to disclose the original data and processing process.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors appreciate the financial support by Social Science Fund of Liaoning Province of China (L20BJY043) and Social Science Research Fund of Dalian Minzu University (202002003).

References

[1] S. Hosan, S. C. Karmaker, M. M. Rahman, A. J. Chapman, and B. B. Saha, "Dynamic links among the demographic dividend, digitalization, energy intensity and sustainable economic growth: empirical evidence from emerging economies," *Journal of Cleaner Production*, vol. 330, Article ID 129858, 2022.

[2] A. Zakari and I. Khan, "Energy efficiency and sustainable development goals (SDGs)," *Energy*, vol. 239, Article ID 122365, 2022.

[3] J. Guo, L. Wang, W. J. Zhou, and C. Wei, "Powering green digitalization: evidence from 5G network infrastructure in China," *Resources, Conservation and Recycling*, vol. 182, no. 182, Article ID 106286, 2022.

[4] M. Kovacikova, P. Janoskova, and K. Kovacikova, "The impact of emissions on the environment within the digital economy," *Transportation Research Procedia*, vol. 55, no. 55, pp. 1090–1097, 2021.

[5] H. T. Wu, Y. Hao, S. Y. Ren, X. Yang, and G. Xie, "Does internet development improve green total factor energy efficiency? Evidence from China," *Energy Policy*, vol. 153, pp. 112247–112313, 2021.

[6] T. Peter and Č. Matej, "Humanizing digital life: reducing emissions while enhancing value-adding human processes," *International Journal of Information Management*, vol. 63, Article ID 102443, 2022.

[7] Y. M. Perez and C. M. Lacalle, "The impact of knowledge diffusion on economic growth across countries," *World Development*, vol. 132, Article ID 104995, 2020.

[8] D. W. Jorgenson and K. M. Vu, "The ICT revolution, world economic growth, and policy issues," *Telecommunications Policy*, vol. 40, no. 5, pp. 383–397, 2016.

[9] H. Wu, Y. Xue, Y. Hao, and R. Siyu, "How Does Internet Development Affect Energy-Saving and Emission Reduction? Evidence from China," *Energy Economics*, vol. 103, Article ID 105577, 2021.

[10] R. R. Deng and A. X. Zhang, "The impact of urban digital finance development on carbon emission performance in China and mechanism," *Resources Science*, vol. 43, no. 11, pp. 2316–2330, 2021.

[11] L. Wang, Y. Y. Chen, T. S. Ramsey, and G. J. Hewings, "Will researching digital technology really empower green

development?" *Technology in Society*, vol. 66, pp. 101638–101713, 2021.

[12] S. Lange, J. Pohl, and T. Santarius, "Digitalization and energy consumption. Does ICT reduce energy demand?" *Ecological Economics*, vol. 176, Article ID 106760, 2020.

[13] B. Lin and Y. Zhou, "Does the Internet development affect energy and carbon emission performance?" *Sustainable Production and Consumption*, vol. 28, no. 28, pp. 1–10, 2021.

[14] Z. G. Li and J. Wang, "The dynamic impact of digital economy on carbon emission reduction: evidence city-level empirical data in China," *Journal of Cleaner Production*, vol. 351, no. 351, Article ID 131570, 2022.

[15] D. D'Amato and J. Korhonen, "Integrating the green economy, circular economy and bioeconomy in a strategic sustainability framework," *Ecological Economics*, vol. 188, Article ID 107143, 2021.

[16] D. Ma and Q. Zhu, "Innovation in emerging economies: research on the digital economy driving high-quality green development," *Journal of Business Research*, vol. 145, pp. 801–813, 2022.

[17] L. Wu and Z. Zhang, "Impact and threshold effect of internet technology upgrade on forestry green total factor productivity: evidence from China," *Journal of Cleaner Production*, vol. 271, Article ID 122657, 2020.

[18] L. Li, Y. Zheng, S. Zheng, and H. Ke, "The new smart city programme: evaluating the effect of the internet of energy on air quality in China," *Science of the Total Environment*, vol. 714, Article ID 136380, 2020.

[19] Z. Dehghan Shabani and R. Shahnazi, "Energy consumption, carbon dioxide emissions, information and communications technology, and gross domestic product in Iranian economic sectors: a panel causality analysis," *Energy*, vol. 169, pp. 1064–1078, 2019.

[20] J. G. Koomey, H. Scott Matthews, and E. Williams, "Smart everything: will intelligent systems reduce resource use?" *Annual Review of Environment and Resources*, vol. 38, no. 1, pp. 311–343, 2013.

[21] E. Ismagilova, L. Hughes, Y. K. Dwivedi, and K. R. Raman, "Smart cities: advances in research-An information systems perspective," *International Journal of Information Management*, vol. 47, pp. 88–100, 2019.

[22] M. Singh and G. P. Sahu, "Towards adoption of Green IS: a literature review using classification methodology," *International Journal of Information Management*, vol. 54, Article ID 102147, 2020.

[23] H. Ishida, "The effect of ICT development on economic growth and energy consumption in Japan," *Telematics and Informatics*, vol. 32, no. 1, pp. 79–88, 2015.

[24] P. Schulte, H. Welsch, and S. Rexhauser, "ICT and the demand for energy: evidence from OECD countries," *Environmental and Resource Economics*, vol. 63, no. 1, pp. 119–146, 2016.

[25] J. Khuntia, T. J. V. Saldanha, S. Mithas, and V. Sambamurthy, "Information technology and sustainability: evidence from an emerging economy," *Production and Operations Management*, vol. 27, no. 4, pp. 756–773, 2018.

[26] K. Saidi, H. Toumi, and S. Zaidi, "Impact of information communication technology and economic growth on the electricity consumption: empirical evidence from 67 countries," *Journal of the Knowledge Economy*, vol. 8, no. 3, pp. 789–803, 2017.

[27] M. Salahuddin and J. Gow, "The effects of Internet usage, financial development and trade openness on economic growth in South Africa: a time series analysis," *Telematics and Informatics*, vol. 33, no. 4, pp. 1141–1154, 2016.

[28] M. Salahuddin and K. Alam, "Information and Communication Technology, electricity consumption and economic growth in OECD countries: a panel data analysis," *International Journal of Electrical Power & Energy Systems*, vol. 76, no. 3, pp. 185–193, 2016.

[29] U. Ahmed, O. Ilhan, S. Ullah, and A. Hassan, "Does ICT Have Symmetric or Asymmetric Effects on CO2 Emissions? Evidence from Selected Asian Economies," *Technology in Society*, vol. 67, Article ID 101692, 2021.

[30] W. Zhang, X. M. Liu, D Wang, and J Zhou, "Digital economy and carbon emission performance: Evidence at China's city level," *Energy Policy*, vol. 165, Article ID 112927, 2022.

[31] X. Y. Zhou, D. Q. Zhou, Z. Zhao, and Q. Wang, "A framework to analyze carbon impacts of digital economy: the case of China," *Sustainable Production and Consumption*, vol. 31, pp. 357–369, 2022.

[32] R. G. Chambers, Y. Chung, and R. Färe, "Benefit and distance functions," *Journal of Economic Theory*, vol. 70, no. 2, pp. 407–419, 1996.

[33] N. Zhang, F. B. Kong, Y. Choi, and P. Zhou, "The effect of size-control policy on unified energy and carbon efficiency for Chinese fossil fuel power plantsed energy and carbon efciency for Chinese fossil fuel power plants," *Energy Policy*, vol. 70, pp. 193–200, 2014.

[34] R. Färe, S. Grosskopf, and C. Pasurkajr, "Environmental production functions and environmental directional distance functions," *Energy*, vol. 32, no. 7, pp. 1055–1066, 2007.

[35] B. Lin and J. Zhu, "Fiscal spending and green economic growth: evidence from China," *Energy Economics*, vol. 83, pp. 264–271, 2019.

[36] G. Jin and K. R. Shen, "Beggar-thy-neighbor or partner? Environmental regulation enforcement interaction and urban productivity growth," *Management World*, vol. 34, no. 12, pp. 43–55, 2018.

[37] R. Guo and Y. J. Yuan, "Different types of environmental regulations and heterogeneous influence on energy efficiency in the industrial sector: evidence from Chinese provincial data," *Energy Policy*, vol. 145, Article ID 111747, 2020.

[38] B. E. Hansen, "Threshold effects in non-dynamic panels: estimation, testing, and inference," *Journal of Econometrics*, vol. 93, no. 2, pp. 345–368, 1999.

[39] J. Han, F. Feng, and X. Y. Jiang, "Internet development and global value chain embedment: a Transnational empirical Study based on GVC Index," *Nankai Economic Research (Chinese)*, vol. 4, pp. 21–35+52, 2018.

Hindawi
Mathematical Problems in Engineering
Volume 2022, Article ID 1440323, 11 pages
https://doi.org/10.1155/2022/1440323

Research Article

Transformer-Based Data-Driven Video Coding Acceleration for Industrial Applications

Yixiao Li,[1,2] Lixiang Li,[1,2] Zirui Zhuang,[3] Yuan Fang,[1,2] Haipeng Peng,[1,2] and Nam Ling[4]

[1]*Information Security Center, State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China*
[2]*National Engineering Laboratory for Disaster Backup and Recovery, Beijing University of Posts and Telecommunications, Beijing 100876, China*
[3]*State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China*
[4]*Department of Computer Science and Engineering, Santa Clara University, 95053 Santa Clara, USA*

Correspondence should be addressed to Lixiang Li; lixiang@bupt.edu.cn

Received 14 March 2022; Revised 4 September 2022; Accepted 14 September 2022; Published 27 September 2022

Academic Editor: Ardashir Mohammadzadeh

With the exploding development of edge intelligence and smart industry, deep learning-based intelligent industrial solutions are promptly applied in the manufacturing process. Many intelligent industrial solutions such as automatic manufacturing inspection are computer vision based and require fast and efficient video encoding techniques so that video streams can be processed as quickly as possible either at the edge cluster or over the cloud. As one of the most popular video coding standards, the high efficiency video coding (HEVC) standard has been applied to various industrial scenes. However, HEVC brings not only a higher compression rate but also a significant increase in encoding complexity, which hinders its practical application in industrial scenarios. Fortunately, a large amount of video coding data makes it possible to accelerate the encoding process in the industry. To speed up the video coding process in some industrial scenes, this paper proposes a data-driven fast approach for coding tree unit (CTU) partitioning in HEVC intracoding. First, we propose a method to represent the partition result of a CTU as a column vector of length 21. Then, we employ lots of encoding data produced in normal industry scenes to train transformer models used to predict the partitioning vector of the CTU. Finally, the final partitioning structure of the CTU is generated from the partitioning vector after a postprocessing operation and used by an industrial encoder. Compared with the original HEVC encoder used by some industrial applications, experiment results show that our approach achieves 58.77% encoding time reduction with 3.9% bit rate loss, which indicates that our data-driven approach for video coding has great capacity working in industrial applications.

1. Introduction

With the development of the smart industry, image and video play a key role in many industrial scenarios [1–3]. As a result, video coding standards have been used more widely than ever [4]. Although advanced video coding (AVC) was introduced in 2003, it is still used by many applications today due to its fast coding speed [5]. However, with the increasing demand for high resolution and ultra-high resolution video and the improvement of hardware computing, HEVC is gradually replacing AVC because of its superior coding efficiency. Although high efficiency video coding (HEVC) was developed about one decade ago, the computational complexity is still not low at present due to the introduction of many advanced coding tools [6]. Some certain industrial scenarios have an urgent demand for real-time high definition display, and high encoding complexity hinds the applications of HEVC in these scenarios. Fortunately, it is becoming more mature for the use of big data and deep learning techniques in the industry [7–12]. Besides, mobile video applications have exploded in recent years and produced massive encoding data, which makes it possible for us

to employ big data and deep learning techniques to reduce the encoding time and speed up the encoding process in industrial applications [13].

The main computing burden of HEVC encoding comes from a complicated partition rule for intracoding called quadtree structure. To lower the computational complexity, many scholars have developed lots of fast algorithms for HEVC encoding by using various techniques, such as the learning-based technique and 3-type fuzzy logic system (FLS) [14]. Learning-based techniques are good at finding patterns in data. Dong et al. [15] proposed a learning-based fast algorithm for versatile video coding (VVC) from two aspects of mode selection and prediction terminating to reduce coding complexity, and 3-type fuzzy logic systems are good at solving equations or finding a mathematical model to represent the relationship between output and associated input variables [16]. With the advances in modeling problems, 3-type FLS is suitable for mode decision or rate control tasks in video coding and may bring potential improvement for encoding performance [14]. There are quite a number of algorithms focusing on fast coding unit (CU) partitioning. Furthermore, with the popularity of deep learning technology [17], researchers use neural networks (NNs) to boost CU partitioning and achieve satisfactory results [18]. Works using the deep learning method can be roughly classified into two main categories which are the multistage partition approach and the end-to-end structure decision approach.

In the category of multistage partition, approaches regard the structure determination of CTU partitioning as a combination of several binary classification problems. Xu et al. [19] proposed a three-level CTU splitting algorithm based on a convolution neural network (CNN). They trained three deep CNN models, of which each predicted the split flag for a CU in a certain depth. Kim and Ro [20] also proposed a CTU partition algorithm by using three CNN-based classifiers, they used different networks to predict the splitting decisions of CUs in different sizes. Shi et al. [21] proposed an AK-CNN for fast CTU partition prediction. Their AK-CNN classifiers are well-designed and can detect texture complexity of the CU quickly. Shen et al. [22] proposed early determination and a bypass strategy for CU size decisions by using the texture property of the current CU and coding information from neighboring CUs. Moreover, Shen et al. [23] proposed a fast intermode decision algorithm for HEVC by jointly using the interlevel correlation of quadtree structure and the spatiotemporal correlation. Based on visual perception and machine learning, Chen et al. [24] proposed a fast algorithm by using random forest models to quickly select the partition for VVC intracoding. However, methods in this category usually need as many as three NN models, which require much training work and cause implementation difficulties. Besides, the splitting of the CU is usually related to partition statuses of neighboring CUs, and methods in this category obviously ignore the splitting information from neighboring CUs by considering the splitting problem hierarchically.

In the category of end-to-end decision, one partition structure or several possible result candidates of the CTU can be generated through a single prediction. Liu et al. [18] proposed a VLSI friendly approach to partition a CTU by designing a shallow CNN. Feng et al. [25] proposed a fast block partitioning algorithm using CNN-based depth map prediction for HEVC intracoding. They used a depth map to represent the partition structure of the CTU so that quadtree structure can be predicted end to end. Tissier et al. [26] proposed an edge possibility prediction approach by using CNN. In their approach, most possible CTU partition results were generated, and the final partition was determined through postprocessing and rate-distortion optimization (RDO). Although methods in this category consider the influence of the entire CTU space information on the splitting of the current CU, they do not take into account the influence of the splitting of parent CUs. Furthermore, for some end-to-end methods, which take a strategy of reducing the CTU partition candidates for RDO, the RDO process is not fully skipped so that the complexity reduction is limited.

To make full use of splitting information from both neighboring and parent CUs, we first propose a new representation for CTU partitioning structure. Specifically, we use an array called split vector (SV) to represent a CTU partitioning structure. Then, we design and train a transformer model to predict the PV of the CTU. With the introduction of the transformer, the CTU partition is regarded as a sequence problem. Finally, the CTU partition structure is decided from SV through postprocessing, and the RDO process is no longer needed for CTU partition searching. The main contributions in this paper are described as follows:

(1) An array called SV is proposed to represent the partition structure of the CTU. With the use of SV, RDO progress searching for optimal CTU partitioning structures can be fully skipped.

(2) We introduce transformer models into the fast determination task for CTU partition structure and imaginatively model the partition problem as a sequence problem affecting each other.

(3) We not only design effective transformer models to predict the SV of the target CTU with high accuracy but also build several datasets for transformer model training.

This paper is organized as follows: Section 1 introduces the background of the proposed approach. Section 2 introduces the fundamental knowledge used in this paper. Section 3 describes the proposed fast algorithm for CTU partitioning. Section 4 shows the experiment results, and Section 5 concludes this paper.

2. Fundamental Knowledge

In this section, we introduce the quadtree structure of HEVC intracoding, and a brief description of the transformer is also given.

2.1. Quadtree Structure of HEVC Intrapartitioning. First, each intraframe is divided into nonoverlapped square blocks called coding tree units (CTUs) which are usually 64×64 pixels. To cope with the texture characteristic of the CTU,

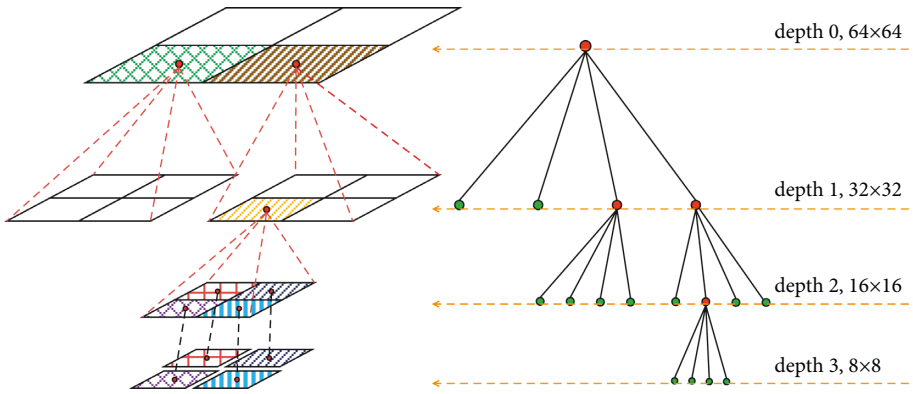


Figure 1: A partition sample of the CTU and the corresponding quadtree structure.

the CTU can be further split into four equal-sized square blocks, and each block can be iteratively further split into four squared sub-blocks according to the quadtree partition structure. In the quadtree of the final CTU partitioning, the CTU serves as the root, and each leaf node represents a coding unit (CU). The CU size of HEVC intracoding varies from 64×64 pixels to 8×8 pixels since the max depth of a quadtree is 3. Figure 1 shows the partitioning result of the CTU and the corresponding quadtree structure. The deeper the quadtree, the smaller the CU.

CTU can be adaptively partitioned into CUs of different sizes to achieve optimal coding efficiency, and the final partition structure is decided according to a brute-force method called rate-distortion optimization (RDO) [27]. RDO evaluates the cost of every possible partition structure in terms of bit rate and visual quality first. As a result, the partition structure with minimal cost is selected as the final partition result for the CTU. Obviously, the encoding time spent on the RDO process increases exponentially with the increase of quadtree depth. Thus, it is essential to reduce the computational burden by replacing the RDO process with an end-to-end approach for CTU partition structure determination.

2.2. Transformer Network. Attention mechanism has been widely used in neural network models such as recurrent neural networks (RNNs) and convolutional neural networks (CNNs) [28]. Since the attention mechanism was proposed, sequence-to-sequence models with attention have shown performance improvement in various tasks. In 2017, Ashish et al. [29] first proposed an attention fully based model named transformer. In order to integrate the advantages of CNNs and RNNs, they creatively used the full attention mechanism to build the network. They applied the transformer to machine translation tasks and achieved state-of-the-art effects at that time.

Like most sequence-to-sequence models, the transformer can be divided into two main parts, i.e., encoder and decoder. The encoder is responsible for mapping the input sequence into a hidden layer, which is the mathematical expression of the input sequence. The decoder then maps the hidden layer back to the target sequence. Using a transformer, we can solve various problems, such as image classification, summary generation, and machine translation. Figure 2 shows the model structure diagram of the transformer. In Figure 2, the encoder consists of N subencoders, and each subencoder includes two layers, which are a multihead attention mechanism and a fully connected feed-forward network, respectively. Besides, residual connection and normalization are also added to each layer. As we can see from Figure 2, transformer decoders are also composed of N subdecoders, and each subdecoder has one more masked multihead attention layer than the subencoder.

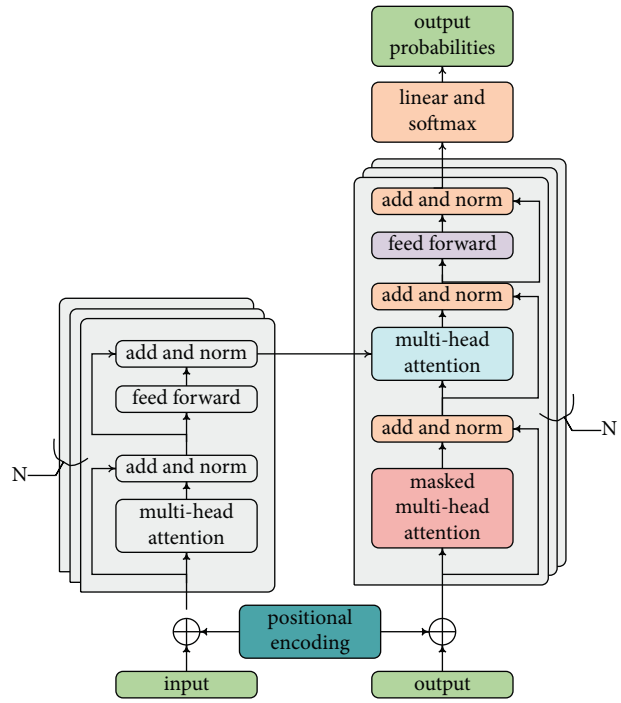


Figure 2: Network structure of the transformer employed by the proposed approach.

The transformer is a new network architecture designed to replace RNN and CNN. It can be stacked to very deep depth so that it can fully explore the characteristics of a deep neural network and achieve high accuracy. Unlike CNN, which is only able to obtain local information, it can directly obtain global information and capture sequence

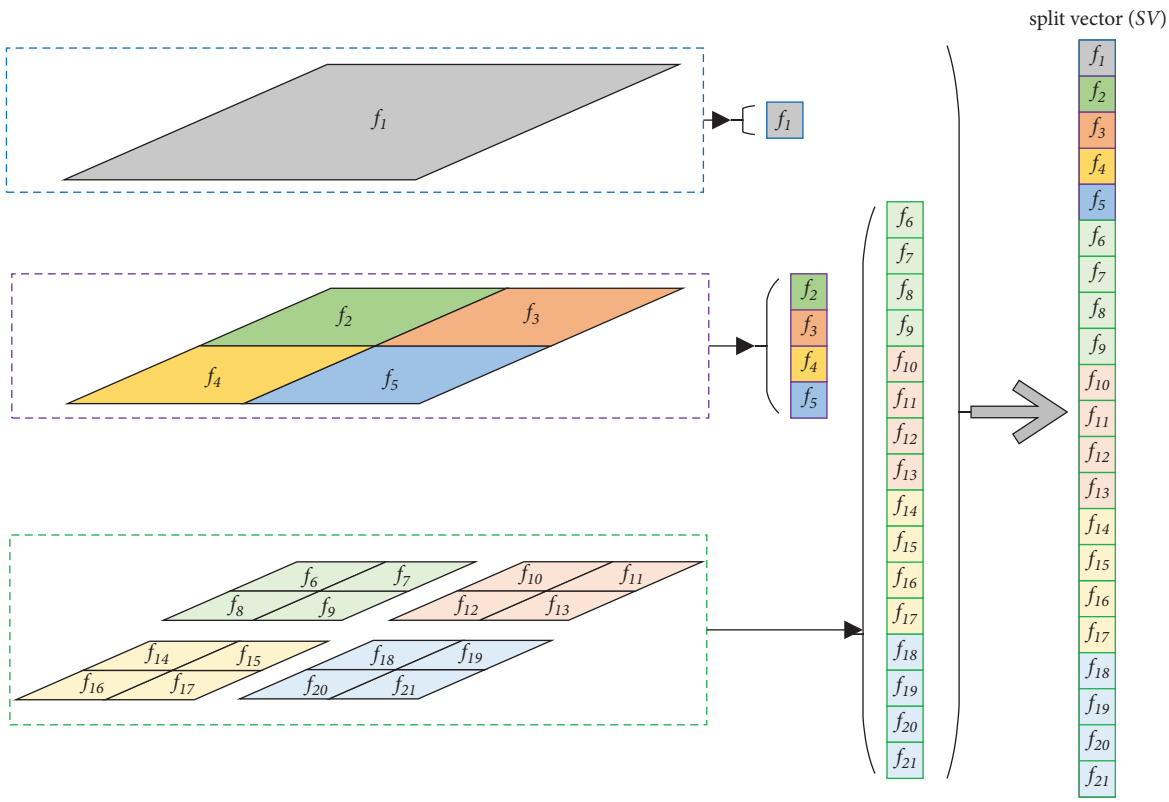


Figure 3: Representation of the CTU partition structure.

dependence. Compared with RNN, the transformer can realize fast parallelism using the attention mechanism, and the training time is greatly reduced.

3. Proposed Method

This section is divided into parts. Firstly, we introduce how the CTU partition structure is represented in the proposed approach. Then, we introduce the training process of transformer models, and loss function designing and train dataset preparation are both included. Finally, we present the postprocessing procedure for the outputs of transformer models.

3.1. Representation of the CTU Partition Structure. In this paper, we propose a novel representation way for the CTU partition structure. Specifically, each of the CTU partition structures can be represented by a vector called split vector (denoted as SV) which contains 21 Boolean split flags. Figure 3 illustrates how the CTU partition structure is reflected in an SV.

According to the quadtree partition rule used in HEVC intracoding, the CTU is split recursively until the max CU depth is reached. Thus, a CTU partition structure can be represented depth wisely. In depth 0, f_1 is the split flag of a CTU. In depth 1, four Boolean values (f_2, f_3, f_4, and f_5 in Figure 3) are used to represent the split status of the four sub-CUs of the current CTU. Furthermore, f_6–f_{21} represent the split flag of each 16×16 CU in depth 2, respectively. Finally, these 21 CU split flags form the SV of the current CTU.

It is convenient and effective to put these flags together into SV. The SV provides us with a new way of reviewing the CTU partitioning jointly, and it avoids treating the CTU partition problem as a level-wise binary classification task by using several cascade prediction models. In other words, SV makes it possible for our method to predict the split flags of blocks in the same CTU jointly due to the high correlation among them.

3.2. Overview of the Proposed Approach. Most existing methods, either statistic-based or learning-based, partition a CTU by deciding whether to split CUs in particular depths. These methods usually employ as many as 3 modes to complete the binary prediction for CUs in different depths. However, with the proposed end-to-end method called TBFA, the partition structure of CTUs is decided through a single prediction by using SV.

Besides, the splitting of the CU is not only highly related to the entire CTU it locates but is also affected by other CUs locating the same CTU. However, many previous methods only focus on the current block without considering information from neighboring or father CUs. BTFA is designed to be able to consider them jointly by using SV, which consists of 21 split flags of neighboring and father CUs in the current CTU.

Also, we specifically design a transformer to predict the SV of the CTU. In BTFA, the transformer is used to explore

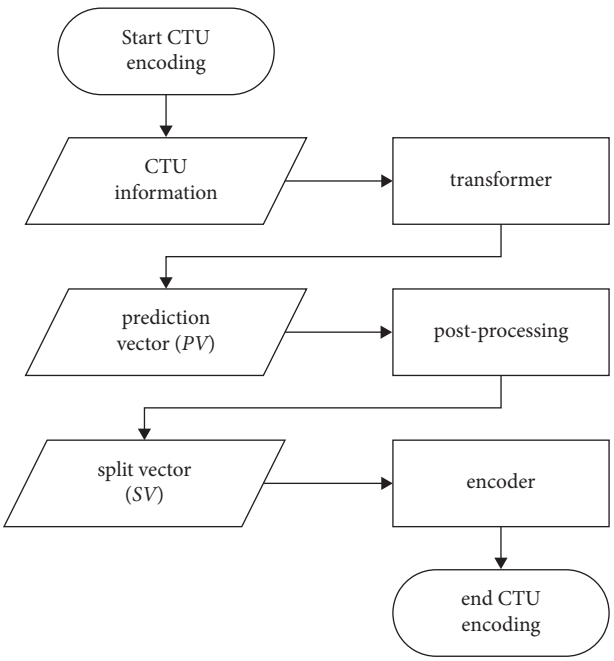


Figure 4: Flowchart of the proposed approach.

the relations among those 21 split flags. The transformer takes the luminance values of the CTU as input and outputs a prediction vector (denoted as PV) of the CTU. Through training the transformer models, PV equals the target SV of the current CTU with high accuracy. Thus, SV can be obtained from PV after the postprocessing procedure. Once the SV is obtained, the entire CTU partition is determined. Then, the encoder can encode each CTU directly, completely avoiding the RDO process for partitioning. Figure 4 shows the whole flow of the proposed TBFA.

3.3. Transformer Training. The transformer predictor plays a key role in BTFA, and a classic transformer structure is employed by our approach. In this part, the design of the loss function is described, and the way in which data are prepared for training is also given.

3.3.1. Loss Function Design. The loss function is used to measure the distance or bias between the ground truth and the predicted result outputted by a model for the current sample. In the case of this paper, the loss function is employed to measure the error between PV and SV.

Actually, PV output by the transformer contains the partition probability of each part corresponding to the current CTU. Conversion from PV output by the transformer to SV is required for encoding, and the conversion process is quite simple that only a comparison between each element and 0.5 is needed. Specifically, if one element in PV is smaller than 0.5, the corresponding element located at the same position in SV is set to 0. On the contrary, if one element in PV is greater than 0.5, the colocated element in SV is set to 1. Contradictory values may exist in SV transformed from PV since the judging of each element of PV is independent during the conversion process. When a CTU is not split, related sub-CUs in depth 1 and 2 are not split either. Thus, if f_1 in Figure 3 equals 0, the values of rest elements in SV are supposed to be 0. To address the inconsistency among values of SV, we carefully designed a loss function for training.

When optimizing the weight parameters of the model iteratively, there is no point in considering the prediction error of the corresponding four sub-CUs if the real label of the current CU is nonsplit. Based on this, we designed the loss function, and the final loss of a sample can be calculated as

$$L = g_1 + l_1 \sum_{i=2}^{5} \left(g_i + l_i \sum_{j=4i-2}^{4i+1} g_j \right), \tag{1}$$

$$g_i = \left((1 - l_i) \log \frac{1}{1 - p_i} + l_i \log \frac{1}{p_i} \right), \tag{2}$$

where l_i denotes the ith element of SV, p_i denotes the ith element of PV, i is an integer from 1 to 21, and g_i is the cross entropy function for each of 21 element pairs.

The loss function used in our model training is designed and improved from cross entropy. It removes the effect brought by contradictory values during the training procedure, improves prediction accuracy, and reduces computational burden as well.

3.3.2. Train Data Preparation. The dataset for training is so important that it can influence model prediction performance greatly. To make sure transformer models achieve their best performance, we construct a reasonable dataset for each model. In this paper, the transformer model is used to make a one-shot prediction of the partition structure for the CTU. Thus, only one transformer model is needed when a video sequence is being encoded. However, the quantization parameter (QP) is one of the key factors affecting CU splitting result, and the partitioning structure of the CTU differs from different QP values. To make our algorithm more specific and achieve higher prediction accuracy and better encoding performance, we train one dedicated transformer model for each QP. To validate the proposed method, we take 4 classic QP values in this paper, and there are a total of 4 datasets constructed for QP 22, 27, 32, and 37, respectively.

We select five videos from standard test sequences and employ HEVC reference software HM16.7 to encode their entire frames under all-intramode and four QP values (22, 27, 32, and 37). These five sequences are *Traffic* (2560 × 1600), *ParkScene* (1920 × 1080), *BasketballDrill* (832 × 480), *BQSquare* (416 × 240), and *FourPeople* (1280 × 720). Sequences used for model training are with different scenes and resolutions, which make trained models more robust and generalized.

Once encoding is completed on these training sequences, we can obtain encoding results of each training sequence under each certain QP. In particular, partitioning structures of all CTUs in training sequences are generated and are further converted to ground truth SVs served as target labels in the training dataset for corresponding QP. As a result,

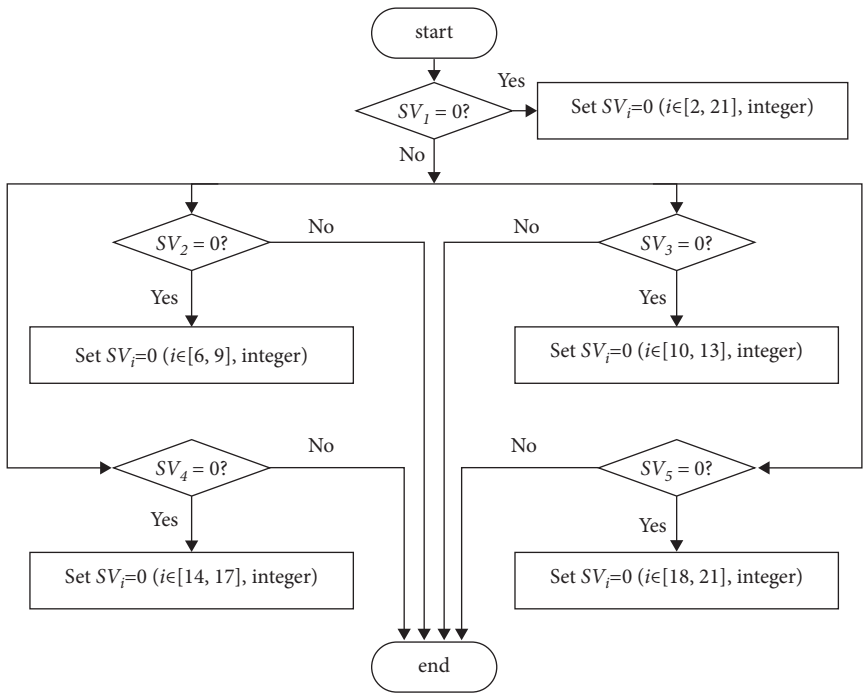


Figure 5: Postprocessing procedure on SV converted from PV.

four training datasets are assembled, and each sample of a dataset consists of luminance pixels and an SV of each CTU. It is worth noting that the only difference among these four datasets is SVs. Samples of different datasets share the same luminance pixel values since they are from the same training sequences.

3.4. Postprocessing of the Split Vector. PV is first output by our transformer model once the derivation is completed. Then, the predicted SV of the current CTU is generated from PV through a conversion process. SV consists of 21 binary elements, and each corresponds to a potential sub-CU of the current CTU. Although the contradiction among SV elements has been mitigated slightly by a well-designed loss function, the influence of loss function designing is not great enough to remove all contradictions in SV. So SV cannot be used directly as the split flags of CUs in the CTU. It is necessary to carry out a postprocessing procedure on SV converted from PV, and the detailed postprocessing procedure is described in Figure 5.

As we can see from Figure 5, the postprocessing procedure is easy enough to be carried out quickly. To be specific, we set SV_2 all the way to SV_{21} to be 0 if SV_1 is 0. If SV_2 is 0, we set SV_6, SV_7, SV_8, and SV_9 to be 0. Similarly, we set SV_{10}, SV_{11}, SV_{12}, and SV_{13} to be 0 if SV_3 is 0. We set SV_{14}, SV_{15}, SV_{16}, and SV_{17} to be 0 if SV_4 is 0. We set SV_{18}, SV_{19}, SV_{20}, and SV_{21} to be 0 if SV_5 is 0. After the fine-tuning described above, SV is able to be used in video encoding finally.

4. Experiment Result

In this section, we first analyze the prediction accuracy of transformer models employed by the proposed approach. Deep learning framework PyTorch is used to complete training and prediction. Simulations were executed on a Windows 7 64 bit operating system workstation with NVIDIA RTX 2080s GPU and Intel(R) Xeon(R) CPU E5-2623 v3 @ 3.00 GHz and 3.00 GHz (2 processors), 64.0 GB.

Then, to verify the effectiveness of the proposed one-shot approach for CTU partitioning, we implemented it on HEVC reference platform HM16.7. Coding parameters, such as additional coding tools, were set as default. Besides, all-intra main configuration was adopted to encode all the frames of video sequences in the standard test set. The BD-rate was employed to evaluate the coding performance of the proposed method, and the time saving ratio denoted by TS was used to measure the complexity reduction of encoding algorithms. It is defined as

$$TS = \frac{\text{time}_o - \text{time}_p}{\text{time}_o}, \tag{3}$$

where time_o denotes time spent by the original HM16.7 encoder and time_p is the time spent by the encoder on which the proposed algorithm is implemented.

At last, partition results of original HM16.7 and the proposed approach are compared on one frame randomly selected. The comparison results visually demonstrate the partitioning effect of the proposed one-shot algorithm for CTU partitioning structure determination.

Table 1: Splitting accuracy of four transformer models at each CU depth.

Depth level	QP 22 (%)	QP 27 (%)	QP 32 (%)	QP 37 (%)
Depth 0	95.58	91.08	91.04	89.38
Depth 1	78.12	81.81	82.44	82.33
Depth 2	74.54	80.60	82.55	84.87

Table 2: Percentage of inferring time to encoding time under different QPs.

	QP 22 (%)	QP 27 (%)	QP 32 (%)	QP 37 (%)
Percentage	0.015	0.021	0.025	0.031

Table 3: Bit rate loss and time saving comparison between the proposed algorithm and state of the art

Class	Sequence	CNNFA		FICUSA		BTFA		Proposed	
		BD-rate (%)	TS (%)	BD-rate (%)	TS (%)	BD-rate (%)	TS (%)	BD-rate (%)	TS (%)
A	*PeopleOnStreet*	3.97	55.59	2.63	45.00	1.16	38.06	4.03	59.04
	Average	**3.97**	**55.59**	**2.63**	**45.00**	**15.41**	**38.06**	**4.03**	**59.04**
B	*Kimono*	2.38	72.72	0.42	16.27	3.32	59.23	6.50	75.35
	Cactus	6.02	62.98	2.44	37.61	1.11	44.95	4.70	66.02
	BasketballDrive	6.02	69.51	1.46	27.96	2.46	49.09	5.43	71.68
	BQTerrace	4.82	57.89	2.41	41.10	1.09	46.88	2.66	58.07
	Average	**4.81**	**65.78**	**1.68**	**30.74**	**2.00**	**50.04**	**4.82**	**67.78**
C	*BQMall*	8.08	52.14	4.33	44.72	0.97	41.52	4.44	56.98
	PartyScene	9.45	58.75	4.66	46.06	1.36	37.68	1.49	44.07
	RaceHorses	4.42	58.19	—	—	0.60	47.49	2.70	58.22
	Average	**7.32**	**56.36**	**4.50**	**45.39**	**0.98**	**42.23**	**2.88**	**53.09**
D	*BasketballPass*	8.40	64.02	3.34	41.48	0.67	36.34	2.32	48.14
	BlowingBubbles	8.33	60.78	3.09	43.45	0.14	25.95	0.92	33.65
	RaceHorses	4.95	57.29	—	—	0.73	33.76	1.88	42.56
	Average	**7.23**	**60.70**	**3.22**	**42.47**	**0.51**	**32.02**	**1.71**	**41.45**
E	*Johnny*	7.96	66.55	2.23	38.56	2.38	63.13	7.45	73.12
	KristenAndSara	5.48	64.72	3.07	43.19	2.09	58.57	6.19	71.87
	Average	**6.72**	**65.64**	**2.65**	**40.88**	**2.24**	**60.85**	**6.82**	**72.49**
Overall average		**6.01**	**60.81**	**2.93**	**40.89**	**4.23**	**44.64**	**3.90**	**58.77**

4.1. Splitting Accuracy. Accuracy of splitting decisions predicted by transformer models for CUs of each depth is shown in Table 1. Depth levels in Table 1 represent different sizes of the CU, accuracy of depth level 0 means the percentage of right splitting decisions predicted by transformer model for CTUs, and accuracy of depth levels 1 and 2 means the percentage of right splitting decisions for CUs of size 32×32 and 16×16, respectively.

As we can see from Table 1, every transformer model achieves its own highest accuracy on the splitting prediction on depth level 0, and accuracies on levels 1 and 2 do not have much difference and are lower than that on level 0 by around 10%. The reason is that splitting prediction of large CUs is easier than that of small CUs since large CU provides more information for prediction, and texture is more obvious. Another phenomenon is that prediction accuracy of transformer models at depth level 0 decreases as QP increases, while the accuracy at depth level 1 and 2 increases as QP increases. It means that the transformer employed by the proposed approach is good at predicting the splitting results of large CUs when QP is small, while it predicts the splitting results of small CUs easily under large QP values. These properties provide guidance for the industry application of our method.

4.2. Inferring Time Overhead. The proposed approach aims at reducing the encoding time of HEVC by employing transformer models. However, inferring time of transformers must be considered due to their parameter scale. During the encoding process, the inferring time is included for a fair performance evaluation. Besides, we calculate the percentage of inferring time to encoding time under different QPs, and the results are shown in Table 2. As we can see from Table 2, the inferring time overhead of transformers takes a quite low percentage of the encoding time while the proposed method is used, which is because we use GPUs to complete the inferring calculation of transformers.

4.3. Encoding Performance. The last two columns in Table 3 show the encoding performance of the proposed method. Compared to original HM16.7, 58.77% encoding time on

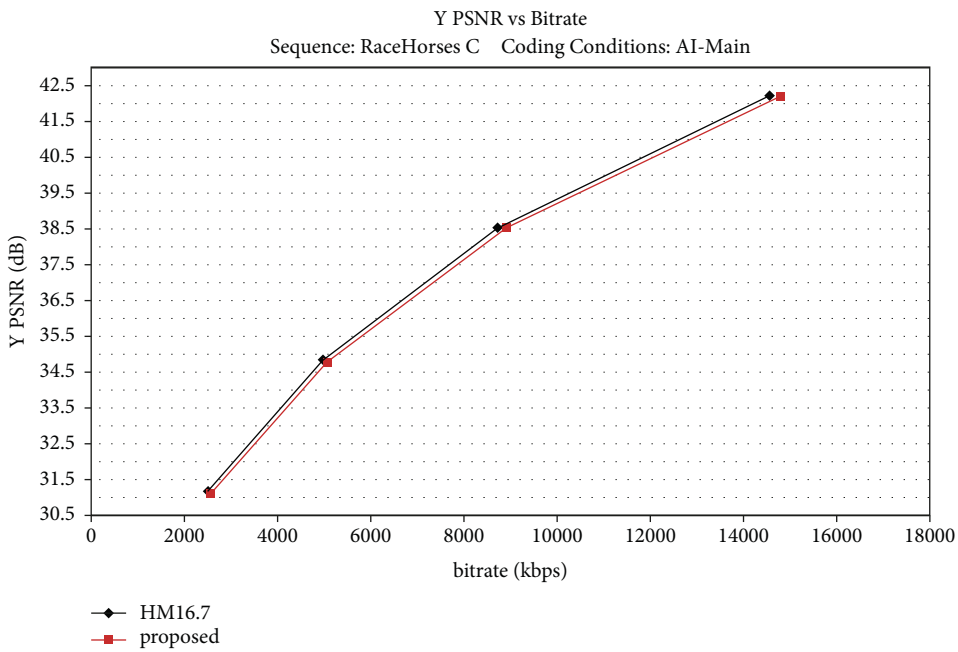


Figure 6: Rate-distortion curve comparison of the proposed approach and original HM16.7 on sequence *RaceHorses* (832 × 480).

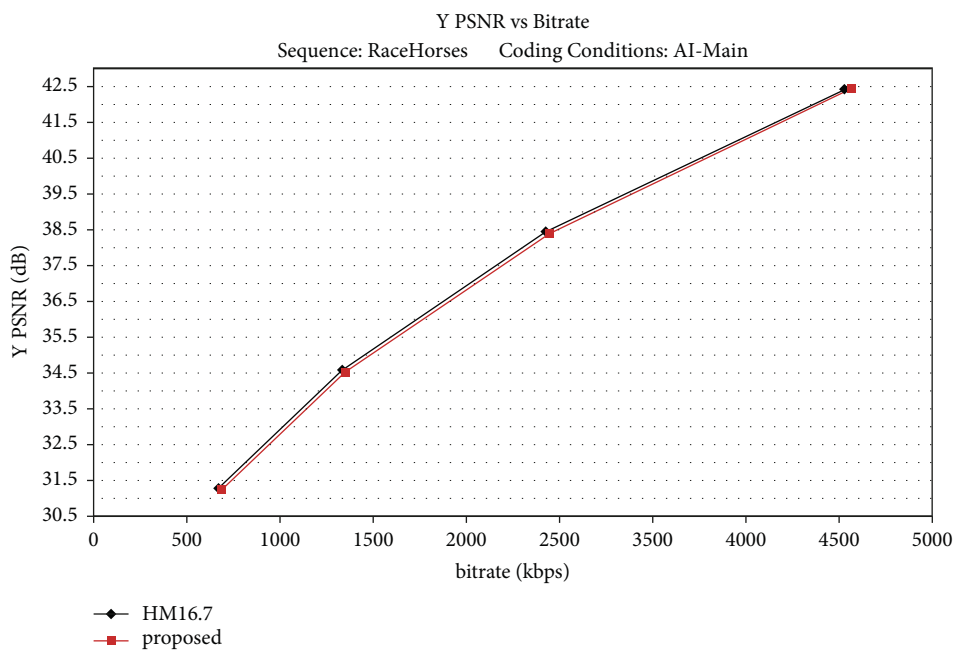


Figure 7: Rate-distortion curve comparison of the proposed approach and original HM16.7 on sequence *RaceHorses* (416 × 240).

average is saved while sacrificing about 3.90% bit-distortion rate. Besides, we observe that our fast approach achieves better results among sequences of resolution class *C* and *D* in terms of BD-rate loss, and time saving is acceptable. Though the proposed algorithm is able to save as much as 59.04%, 67.78%, and 72.49% encoding time on resolution class *A*, *B*, and *E*, respectively, BD-rate loss is not quite low while acceptable. It means that our transformer-based one-shot approach works better on the sequence of low resolution. Especially, we find encoding results to be good and balanced on sequences *RaceHorses* (832 × 480) and *RaceHorses* (416 × 240), and their rate-distortion (RD) curves are shown in Figures 6 and 7, respectively. As we can see from Figures 6 and 7, though the proposed approach saves much encoding time on these two sequences, their RD curves are both quite close to that of original HM16.7, which means the visual difference is negligible for the proposed approach. Observing contents of these two sequences, we find that they

Table 4: Encoding performance of the proposed approach on different resolution classes.

Class	ESDA		BTRNFA		Proposed	
	BD-rate (%)	TS (%)	BD-rate (%)	TS (%)	BD-rate (%)	TS (%)
A	1.00	48.50	2.28	62.21	4.03	59.04
B	1.16	45.60	2.69	70.66	4.82	67.78
C	0.55	33.75	1.32	52.82	2.88	53.09
D	0.30	33.00	0.88	41.86	1.71	41.45
E	2.73	61.33	3.30	74.19	6.82	72.49
Average	**1.15**	**44.44**	**2.09**	**60.35**	**3.90**	**58.77**

Figure 8: Partition result comparison between the proposed approach and original HM16.7.

both contain objects moving fast, which indicates our fast approach is ideal for industrial applications working in scenes involving a lot of movement.

To further analyze the coding performance of the proposed algorithm, we compare it with three recent studies. Respectively, they are convolution neural network-based fast algorithm (CNNFA) proposed by Liu et al. [18], fast intra-CU splitting algorithm (FICUSA) proposed by Zhong et al. [30], and bagged tree-based fast algorithm (BTFA) proposed by Li et al. [31]. CNNFA, FICUSA, and BTFA are all effective schemes for intra-CU size decision of HEVC, and their performance on standard test sequences is listed in Table 3. As we can see, FICUSA provides 0.97% less BD loss compared to ours, but the time saving of our method outperforms about 18%. Considering CNNFA, though its time saving is slightly higher than ours by 2.04%, their BD loss is also higher by 2.11%. Moreover, compared with BTFA, our approach outperforms in terms of both BD-rate and time saving.

Moreover, we also compared the proposed approach with other two fast partition approaches, which are the effective CU size decision approach (ESDA) [22], bagged tree, and ResNet joint fast approach (BTRNFA) [32]. Table 4 shows the encoding performance of ESDA and BTRNFA on different sequence resolution classes. As we can see from Table 4, the proposed approach outperforms ESDA in terms of time saving, while it causes more BD-rate loss. Compared with BTRNFA, the proposed approach makes more BD-rate loss by 2.85% and almost the same time saving. Though the proposed approach does not defeat BTRNFA, it does not require feature extraction and finishes partition prediction using as few as one transformer model.

Tables 3 and 4 prove that the performance of the proposed transformer-based fast approach for CTU partitioning is satisfactory and competitive and has good capacity in practical industrial applications according to the comprehensive performance of various coding scenarios.

4.4. Partition Comparison. To visualize the encoding performance of the proposed fast approach for HEVC intra-coding, we give the partition results predicted by our algorithm on the 200th frame of sequence *Basketball Pass* (416×240) under QP 22. Black, red, and green lines in Figure 8 represent borders of CUs for final encoding. To verify the correctness of CU splitting, we compare the partitioning results of our approach and original HM16.7, and differences are shown with red and green lines in Figure 8. The black line represents that our algorithm and original HM16.7 share the same partition results. The green line represents the boundaries of CUs, which are split by original HM16.7 but are not split by our approach.

Boundaries of CUs decided nonsplit by original HM16.7 but are split by our approach are shown with red lines in Figure 8.

As we can see from Figure 8, the partition results of the proposed algorithm are consistent with those of original HM16.7 in most CU cases. Splitting results of CUs located on main subjects in a frame are almost the same, while differences mainly exist in the background region of a frame. Compared with original HM16.7, the proposed algorithm is more likely to split the CU. Red lines outnumber green lines, which indicates that the splitting errors are far more than nonsplitting errors of the proposed approach.

5. Conclusion

In this paper, we focus on speeding up the video encoding process of industry applications. As a result, we propose a transformer-based fast CTU partitioning algorithm for HEVC intracoding. We convert the CTU partitioning structure to a split vector and employ transformer models to predict that of the target CTU while encoding. Compared with the original HM 16.7 encoder, our approach reduces encoding time by 58.77% on average while sacrificing negligible rate-distortion performance on selected video sequences. Intensive analysis and experiments show that the proposed solution has great capacity for working in industry applications, especially for scenes involving a lot of movement.

Data Availability

Data used and analyzed during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest

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

Acknowledgments

This work was supported by the National Key Research and Development Program of China (Grant no. 2020YFB1805403), the National Natural Science Foundation of China (Grant no. 62032002), the Natural Science Foundation of Beijing Municipality (Grant no. M21034), and the 111 Project (Grant no. B21049).

References

[1] P. Hořejší, K. Novikov, and M. Šimon, "A smart factory in a smart city: virtual and augmented reality in a smart assembly line," *IEEE Access*, vol. 8, pp. 94330–94340, 2020.

[2] L. Wang and Z. Zhang, "Automatic detection of wind turbine blade surface cracks based on UAV-taken images," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 9, pp. 7293–7303, 2017.

[3] H. Yang, C. Huang, L. Wang, and X. Luo, "An improved encoder–decoder network for ore image segmentation," *IEEE Sensors Journal*, vol. 21, no. 10, pp. 11469–11475, 2021.

[4] M. Wang, W. Xie, J. Zhang, and J. Qin, "Industrial applications of UHD video coding with an optimized super-SAO framework," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 12, pp. 7613–7623, 2020.

[5] N. Dolati, A. Beheshti, and H. Azadegan, "A selective encryption for H.264/AVC videos based on scrambling," *Multimedia Tools and Applications*, vol. 80, no. 2, pp. 2319–2338, 2021.

[6] G. J. Sullivan, J. R. Ohm, W. J. Han, and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1649–1668, 2012.

[7] W. Wang, J. Le, Z. Wang et al., "Event-triggered consensus control for high-speed train with time-varying actuator fault," *IEEE Access*, vol. 8, pp. 50553–50564, 2020.

[8] C. Huang, L. Wang, X. Luo, H. Zhang, and Y. Song, "Evolutionary computing assisted deep reinforcement learning for multi-objective integrated energy system management," in *Proceedings of the IEEE 33rd International Conference on Tools with Artificial Intelligence (ICTAI)*, pp. 506–511, IEEE, Washington, DC, USA, November 2021.

[9] Z. Wang, L. Wang, C. Huang, Z. Zhang, and X. Luo, "Soil-moisture-sensor-based automated soil water content cycle classification with a hybrid symbolic aggregate approximation algorithm," *IEEE Internet of Things Journal*, vol. 8, no. 18, pp. 14003–14012, 2021.

[10] C. Huang, H. Zhang, Y. Song, L. Wang, T. Ahmad, and X. Luo, "Demand response for industrial micro-grid considering photovoltaic power uncertainty and battery operational cost," *IEEE Transactions on Smart Grid*, vol. 12, no. 4, pp. 3043–3055, 2021.

[11] L. Li, K. Ota, and M. Dong, "Deep learning for smart industry: efficient manufacture inspection system with fog computing," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 10, pp. 4665–4673, 2018.

[12] T. Cerquitelli, D. J. Pagliari, A. Calimera et al., "Manufacturing as a data-driven practice: methodologies, technologies, and tools," *Proceedings of the IEEE*, vol. 109, no. 4, pp. 399–422, 2021.

[13] F. Al-Turjman and B. D. Deebak, "Seamless authentication: for IoT-big data technologies in smart industrial application systems," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 4, pp. 2919–2927, 2020.

[14] M.-W. Tian, A. Mohammadzadeh, J. Tavoosi et al., "A deep-learned type-3 fuzzy system and its application in modeling problems," *ActaPolytechnicaHungarica*, vol. 19, no. 2, pp. 151–172, 2022.

[15] X. Dong, L. Shen, M. Yu, and H. Yang, "Fast intra mode decision algorithm for versatile video coding," *IEEE Transactions on Multimedia*, vol. 24, no. 2021, pp. 400–414, 2022.

[16] C. Ma, A. Mohammadzadeh, H. Turabieh, M. Mafarja, S. S. Band, and A. Mosavi, "Optimal type-3 fuzzy system for solving singular multi-pantograph equations," *IEEE Access*, vol. 8, pp. 225692–225702, 2020.

[17] F. Zaki, A. E. Mohamed, and S. G. Sayed, "CtuNet: a deep learning-based framework for fast CTU partitioning of H265/HEVC intra- coding," *Ain Shams Engineering Journal*, vol. 2021, no. 1, 2021.

[18] Z. Liu, X. Yu, Y. Gao, S. Chen, X. Ji, and D. Wang, "CU partition mode decision for HEVC hardwired intra encoder using convolution neural network," *IEEE Transactions on Image Processing*, vol. 25, no. 11, pp. 5088–5103, 2016.

[19] M. Xu, T. Li, Z. Wang, X. Deng, R. Yang, and Z. Guan, "Reducing complexity of HEVC: a deep learning approach,"

IEEE Transactions on Image Processing, vol. 27, no. 10, pp. 5044–5059, 2018.

[20] K. Kim and W. W. Ro, "Fast CU depth decision for HEVC using neural networks," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 29, no. 5, pp. 1462–1473, 2019.

[21] J. Shi, C. Gao, and Z. Chen, "Asymmetric-kernel CNN based fast CTU partition for HEVC intra coding," in *Proceedings of the 2019 IEEE International Symposium on Circuits and Systems (ISCAS)*, May 2019.

[22] L. Shen, Z. Zhang, and Z. Liu, "Effective CU size decision for HEVC intracoding," *IEEE Transactions on Image Processing*, vol. 23, no. 10, pp. 4232–4241, 2014.

[23] L. Shen, Z. Zhang, and Z. Liu, "Adaptive inter-mode decision for HEVC jointly utilizing inter-level and spatiotemporal correlations," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 24, no. 10, pp. 1709–1722, 2014.

[24] M. J. Chen, C. A Lee, Y. H. Tsai et al., "Efficient partition decision based on visual perception and machine learning for H. 266/Versatile video coding," *IEEE Access*, vol. 10, no. 2022, pp. 42141–42150, 2022.

[25] A. Feng, C. Gao, L. Li, and L. Dong, "Cnn-based depth map prediction for fast block partitioning in HEVC intra coding," in *Proceedings of the 2021 IEEE International Conference on Multimedia and Expo (ICME)*, July 2021.

[26] A. Tissier, W. Hamidouche, J. Vanne, and F. Galpin, "CNN oriented complexity reduction of VVC intra encoder," in *Proceedings of the 2020 IEEE International Conference on Image Processing (ICIP)*, 25-28 October 2020.

[27] B. Huang, Z. Chen, and K. Su, "Low-complexity rate-distortion optimization for HEVC encoders," *IEEE Transactions on Broadcasting*, vol. 2021, no. 99, pp. 1–15, 2021.

[28] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," *Computer Science*, vol. 2014, 2014.

[29] V. Ashish, B. Samy, and B. Eugene, "Tensor2Tensor for neural machine translation," *Machine Learning*, 2018.

[30] W. Zhong, C. Dong, and X. Yao, "Fast intra-coding unit splitting algorithm based on spatial-temporal correlation in HEVC," *Journal of Image and Graphics*, 2018.

[31] Y. Li, L. Li, Y. Fang, H. Peng, and Y. Yang, "Bagged tree based frame-wise beforehand prediction approach for HEVC intra-coding unit partitioning," *Electronics*, vol. 9, no. 9, p. 1523, 2020.

[32] Y. Li, L. Li, Y. Fang, H. Peng, and N. Ling, "Bagged tree and ResNet-based joint end-to-end fast CTU partition decision algorithm for video intra coding," *Electronics*, vol. 11, no. 8, p. 1264, 2022.

Hindawi
Mathematical Problems in Engineering
Volume 2022, Article ID 7754356, 16 pages
https://doi.org/10.1155/2022/7754356

Research Article

Insulation Resistance Measurement of Airport Navigational Lighting System Based on Deep Learning and Transfer Learning

Z. B. Liu,[1] Q. Wang,[2] H. Li,[2] C. Y. Wang,[2] and J. Y. Fei[2]

[1]*School of Mechanical Engineering, Dalian Jiaotong University, Dalian 116028, China*
[2]*School of Locomotive and Rolling Stock Engineering, Dalian Jiaotong University, Dalian 116028, China*

Correspondence should be addressed to J. Y. Fei; fjy@djtu.edu.cn

Received 25 July 2022; Accepted 6 September 2022; Published 25 September 2022

Academic Editor: Chao Huang

The insulation resistance value is one of the important indexes for the safe operation of airport navigational lighting system. In this paper, a method based on deep learning and transfer learning is proposed to measure the insulation resistance value. To reduce the influence of high voltage environment and signal injection on the measurement accuracy, a multilayer LSTM model is established, in which the network convergence rate is accelerated by introducing a normalized layer in front of the first LSTM layer. Based on the constructed deep network, transfer learning is employed by sharing the weight parameters of the pretraining model to solve the problem of small data sample. The experimental results demonstrate that the proposed method can effectively improve the measurement accuracy of the insulation resistance value.

1. Introduction

The airport navigational lighting system (ANLS) is necessary visual navigation that aids to ensure the safe take-off and landing of the aircraft in special circumstances [1–3]. In the lighting system, almost all the lights are installed outdoors, and the power supply cables are mostly buried in the field. Due to the environmental factors such as ground temperature, humidity, and mold, the phenomenon of water branches is caused, which makes the cables insulation worse [4–7]. If not handled on time, it will cause a short circuit between the cable core and the ground, resulting in a light bulb failure among the grounding points, which will threaten the safety of aircraft take-off and landing to varying degrees. Thus, the working status of the ANLS can be judged by the insulation resistance value of the ANLS circuit cable.

General airport lighting stations are equipped with 500 V and 2500 V hand shaking megohm meter. 500 V is used for periodic inspection of insulate on resistance of low voltage electrical equipment or insulation fault detection, and 2500 V is mainly used for periodic inspection of insulation resistance of light circuit or insulation fault detection. All of these require manual testing, which cannot guarantee personal safety, and there are cases of untimely fault detection. This leads to intelligent insulation resistance testing technology.

The existing insulation measurement technology is typically classified into balance bridge method and signal injection method [8–11]. The basic principle of measuring insulation resistance with balanced bridge method is to connect two individual resistors with DC high voltage line under the Earth ground and use resistors and switches to form a bridge to obtain the differential current loops and corresponding voltages, thereby converting the insulation resistance value. Papers [12, 13] adopt the balanced bridge method, which have a similar application method. It mainly checks the resistance between the positive and negative DC bus and the ground equivalent potential node to determine the insulation state. These methods require electronic switches or high voltage relays and current or voltage sensors. With the advantages of simple measurement, the method still exist some deficiencies. The most notable defects are low sensitivity, reliability, and mechanical wear caused by mechanical switch relays, as well as the lack of consistency and accuracy of the circuit caused by the multiple operational amplifiers. Moreover, the resistance

can only be measured without current. Online measurement will not only cause measurement misalignment but also damage the measuring instrument or even endanger personal safety if the voltage of the measured line is higher than the test high voltage of the instrument. Therefore, this method is not suitable for online monitoring. The basic principle of measuring insulation resistance by signal injection method is to inject a voltage signal of a certain frequency between the insulated cable and the ground and calculate the insulation resistance by measuring the feedback signal. Paper [14] establishes a series battery pack model composed of first-order resistor-capacitance (RC) circuit cell units. Based on this series battery pack model, the insulation monitoring model of low voltage and low frequency signal injection method is designed. The insulation monitoring method in paper [15] is based on the measurement of the differential current of each power supply line. When the insulation resistance is reduced, the leakage current flows to ground, and the differential current in the line can be measured to prevent the occurrence of faults. Paper [16] proposes a method to dynamically change the reference measurement resistance according to the DC bus voltage. This method uses an intermediate value as the reference resistance prediction test value and dynamically changes the reference resistance value according to this prediction value to obtain the test result of insulation resistance. However, the accuracy of insulation resistance measured by these methods is limited by high voltage, interference, and other factors, and it is difficult to determine the selection of injected signals in the signal injection method. Moreover, the ripple of the DC system will be increased, which will affect the quality of power supply and sampling accuracy.

Deep learning has been widely applied in some fields with the in-depth development. In recent years, it has also been applied in circuits. Aiming at the problems of low accuracy and weak robustness of traditional fault line selection method for small current grounding system, a new method based on improved GoogLeNet was proposed in paper [17]. In this paper, the zero-sequence current signal is mapped to a two-dimensional time-frequency diagram by wavelet transform, and the small current grounding data set is prepared. Then, the fault line selection of the small current grounding system is carried out by using the GoogLeNet model. In paper [18], deep learning is applied to diagnose the interturn short circuit problem for permanent magnet synchronous motors. An efficient and accurate method based on a conditional generative adversarial net and an optimized sparse auto encoder is proposed to detect the interturn short circuit problem for permanent magnet synchronous motors. The experimental results indicate that the fault diagnosis accuracy achieves 98.9%. The successful application of deep learning in circuits shows the practicability and applicability of deep learning in the field of circuits. Therefore, a method of the insulation resistance value is proposed based on deep learning and transfer learning. Finally, the experimental measurement results demonstrate that this method can measure the insulation resistance value effectively and the relative error with the calibration value is controlled within 0.6%. Moreover, three different insulation resistance measurement methods are compared in this paper, including switching time-sharing method, deep learning without transfer learning method, as well as the deep learning and transfer learning method proposed in this paper. Compared with the traditional switching time-sharing method, the results show that the measurement accuracy of the deep learning method is greatly improved. At the same time, compared with the measurement method using deep learning alone, the measurement accuracy of insulation resistance value is improved by at least 1% based on the deep learning and transfer learning method proposed in this paper. It can be concluded from the data that the insulation resistance value calculation method based on deep learning and transfer learning proposed in this paper can effectively improve the detection effect of the insulation resistance value cable.

The main contributions of this paper are summarized as follows:

(1) Transfer learning is applied for the first time to solve the problem of small sample size of insulation resistance measurement under high voltage by sharing the pretraining model weight

(2) A multilayer LSTM model is built to minimize the impact of the high voltage environment and injected voltage signal on the insulation resistance measurement accuracy

The remainder of this article is set out as follows. Section 2 introduces the structure of the ANLS and the method for measuring the basic value of insulation resistance. Then, the measurement error of insulation resistance is analyzed. In Section 3, the idea of deep learning and transfer learning method proposed in this paper is described, which mainly analyzes the acquisition of source domain data and the construction of the source domain model. After comparing the effects of the LSTM model and GRU model in this method, the multilayer LSTM model architecture is finally selected. In Section 4, the experimental setup and results are also introduced and this section carries on the comparative experiment and the analysis of the three measurement methods. Finally, Section 5 draws the conclusion and the outlook.

2. Insulation Resistance Measurement Structure of ANLS and Error Analysis

2.1. Structure of ANLS. ANLS is composed of low voltage power distribution system (including substation, standby diesel generator set, relay protection, and switching device), constant current dimmer (including booster transformer), cables, and lamps distributed on the runway of the airport (including isolation transformers and bulbs) [19, 20]. A schematic diagram of a circuit in the ANLS is shown in Figure 1. In order to avoid the impact of single bulb damage on the power supply of the whole circuit, the bulbs in the ANLS are connected in series 1 : 1 through the isolating

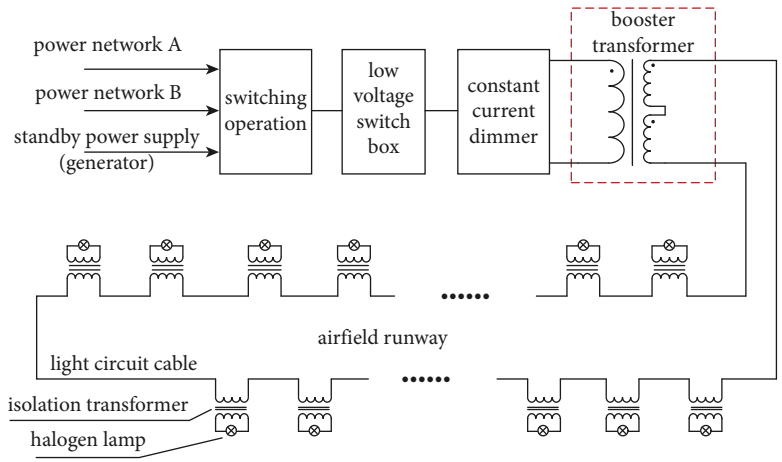


Figure 1: A circuit of the ANLS.

transformer in light circuit cable [21–23]. The constant current dimmer adjusts the effective value of the output voltage according to the change of the load size so that the current in the load loop remains constant at the light level that meets the set brightness level.

2.2. Insulation Resistance Basic Value Acquisition. Based on the switching time-sharing detection method proposed by the author Zhang in previous papers [24], the basic value of insulation resistance was obtained. This method not only reduces the influence of alternating current on measurement but also reduces the influence of cable water tree effect on measurement results. The better deep network model can be trained by using higher precision raw data.

As shown in Figure 2, the insulation resistance detection device has two terminals. The "a" terminal connected with the current sampling resistor R_I is directly grounded and the "b" terminal connected with the current limiting protection resistor R_0 is connected to the cable in the ANLS circuit. The resistance value R_X of the insulation resistance based on the switching time-sharing detection method is calculated as shown in the following equation:

$$R_X = \frac{U_d}{(I_{d+a} - I_a)} - R_0 - R_I. \quad (1)$$

Then, R_X is taken as the basic measurement value of cable insulation resistance for subsequent processing.

In Figure 2, U_d is the 24 V low voltage DC precision power supply, R_0 is the current limiting protection resistor with a resistance value of 1MΩ, R_I is the current sampling resistor with a resistance value of 1 kΩ, C_I is the filter capacitor, KF is the relay, R_X is the unknown equivalent insulation resistor, and U_a is the unknown equivalent AC voltage source.

2.3. Error Analysis of Cable Insulation Resistance Measurement. 30 kVA constant current dimmer is taken as an example. When the dimmer is turned on, considering the worst case of the measuring circuit (refers to the highest voltage under which the AC working voltage is 100% added

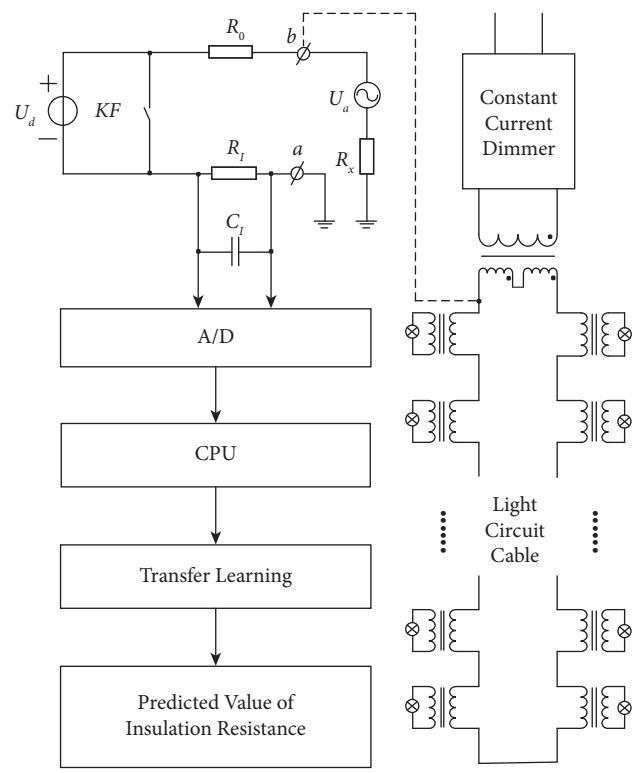


Figure 2: Real-time monitoring structure of insulation resistance.

to the insulation resistance), the maximum effective value of the output AC voltage is 4545 V. According to the circuit theory, when the insulation resistance is in the normal range, R_0 and R_I can be ignored, and then the maximum voltage that the insulation resistance R_X bears is as follows:

$$V_{\max} = U_d + 1.1 \times \sqrt{2} \times 4545V. \quad (2)$$

When the ANLS is running, it is assumed that the stray external interference current is I_g. As shown in Figure 2, the current flowing through the insulation resistance in the ideal state is shown in equation (3) and in the case of interference is shown in equation (4).

$$I_R = I_0, \tag{3}$$

$$I_R' = I_0 + I_g. \tag{4}$$

At this point, the absolute error of insulation resistance measurement is as follows:

$$D = R_x' - R_x = \frac{U_d}{\left(I_0 + I_g\right)} - \frac{U_d}{I_0}. \tag{5}$$

The relative error is as follows:

$$D_R = \left|\frac{D}{R_x'}\right| \times 100\% = \left|\frac{R_x'}{R_x - 1}\right| \times 100\% = \left|\frac{1}{I_0/I_g + 1}\right| \times 100\%. \tag{6}$$

According to equation (6), the greater the interference current I_g is, the greater the relative error D_R is. In order to enhance the anti-interference ability, the existing means mostly use 500 V or higher excitation power U_d. When the cable is selected, the insulation resistance R_X is certain. Only when the value of U_d is larger, I_0 will be larger correspondingly, the influence of the interference current I_g on the error will be relatively small, and the anti-interference ability will increase accordingly. However, the voltage sustained by the insulated cable at this time is at least 500 V higher than that without the installation of online detection device, which increases the electric field strength sustained by the insulated cable. The injection of large voltage signal in the detection circuit leads to the breakdown and leakage of the cable instead of the normal working circuit, which shortens the service life of the cable. Therefore, some scholars began to reduce the injection voltage in order to reduce the influence of electric field intensity. However, with the decrease of the injection voltage, the relative error of the system measurement increases. The selection of the injection signal threshold has not been clearly explained, and whether the injection of the signal will affect the measurement results of the insulation resistance value is not known.

At the same time, the stray current I_g cannot be determined and the interference current generated by AC alone and the interference current generated by AC and DC together cannot be determined to be the same value numerically, so the interference caused by stray current cannot be eliminated simply through the difference method. The aging degree of insulated cables caused by external environment cannot be deduced by precise mathematical formula.

Based on the above discussion, this paper proposes a method to measure the insulation resistance value based on deep learning and transfer learning. The LSTM model is used to train the relationship between the input and output of cable insulation resistance under low voltage condition, and the weight parameters are shared with the model under high voltage condition through transfer learning so as to improve the measurement accuracy of the model. The normalization layer is introduced in front of the first LSTM layer to accelerate the convergence rate of the network.

3. Insulation Resistance Measurement Based on Deep Learning and Transfer Learning

3.1. Overall Structure. This paper proposes a method of insulation resistance measurement for the ANLS circuit cable based on deep learning and transfer learning; the technical process is shown in Figure 3.

The specific operation steps are as follows:

Step 1. When the constant current dimmer is turned off, the standard resistance with known resistance value R_{ys} is used to simulate the working condition of the AMLS circuit cable insulation resistance under low voltage. Then, the basic measurement value of the resistance R_{xs} is measured by switching time-sharing method.

Step 2. A multilayer LSTM network framework is constructed, and R_{xs} is taken as the input value and R_{ys} as the output value and the network convergence speed is accelerated by introducing a normalized layer in front of the first LSTM layer. Then, it is trained and denoted as Net_s, and the network was evaluated.

Step 3. When the constant current dimmer is turned on, the standard resistance with known resistance value R_{yt} is parallel with the circuit cable, and the change of the insulation resistance value of the cable in the ANLS is simulated. Then, the basic measurement value of the resistance R_{xt} is measured by switching time-sharing method.

Step 4. The input layer and the LSTM hidden layer of Net_s are kept, that is, the weight parameters of Net_s network are kept, and the full connection layer and the output layer of Net_s are rebuilt. The normalization layer is also added to accelerate the convergence rate.

Step 5. R_{xt} is taken as the input value and R_{yt} as the output value. Then, it is trained and denoted as Net_t, and the network is evaluated.

Step 6. The insulation resistance value is measured using Net_t.

3.2. Source Domain Data Acquisition. Firstly, the source domain data of insulation resistance are measured under low voltage working environment. When there is a fault in the system, the specific performance is that the insulation resistance value changes. Therefore, a precision resistance A with known resistance value is made in parallel in the circuit, as shown in Figure 4. By closing the miniature toggle switch, a series of resistors with different resistance values can be obtained to simulate the insulation resistance of the cable. Connect resistor A to the "Target Plate" as shown in Figure 5.

Place the device at room temperature and turn on the power. Then, turn on the switch of resistor A in Figure 4 and measure according to the steps in Section 2.2 to obtain a set of data. After cooling off the power and the equipment, measurements are made again after turning on the power and repeated for 6 times to form source domain data $\{R_{xs}, R_{ys}\}_1, \{R_{xs}, R_{ys}\}_2, \ldots, \{R_{xs}, R_{ys}\}_6$. The six sets of sequence data are divided into training set and test set according to 4 : 2.

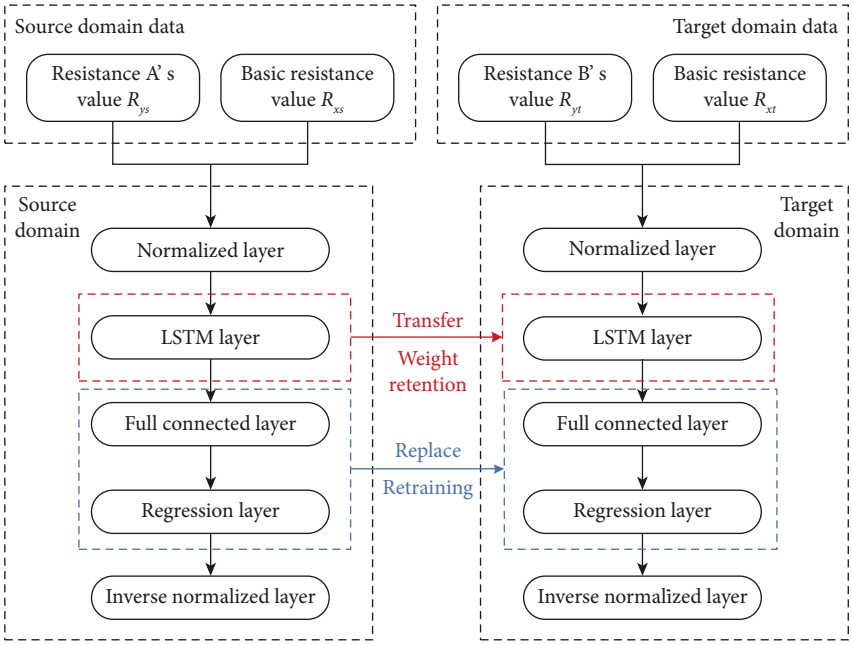


Figure 3: Technical process.

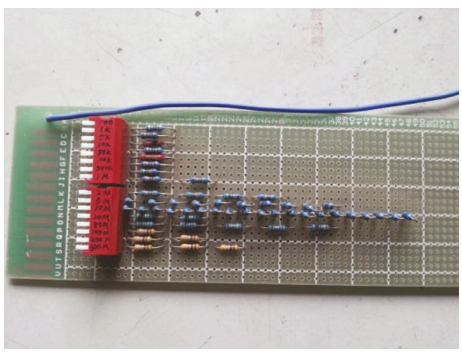


Figure 4: The standard resistance A.

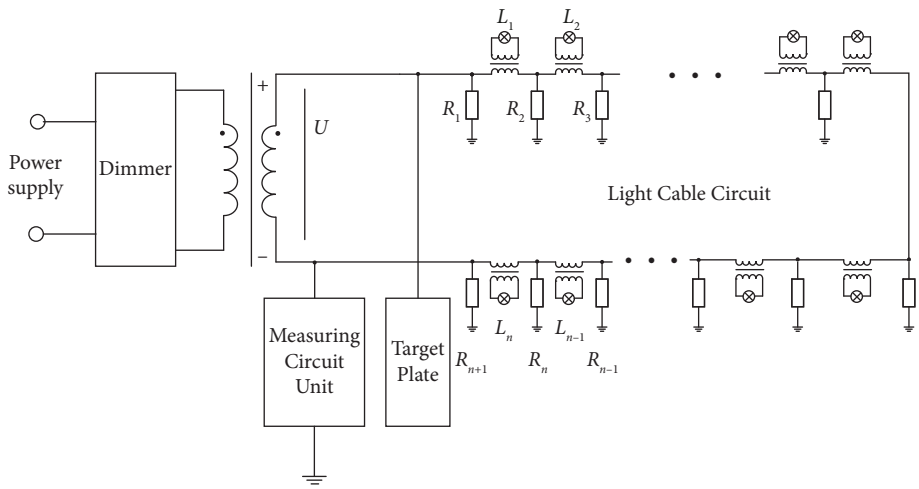


Figure 5: Wiring method of the experimental device.

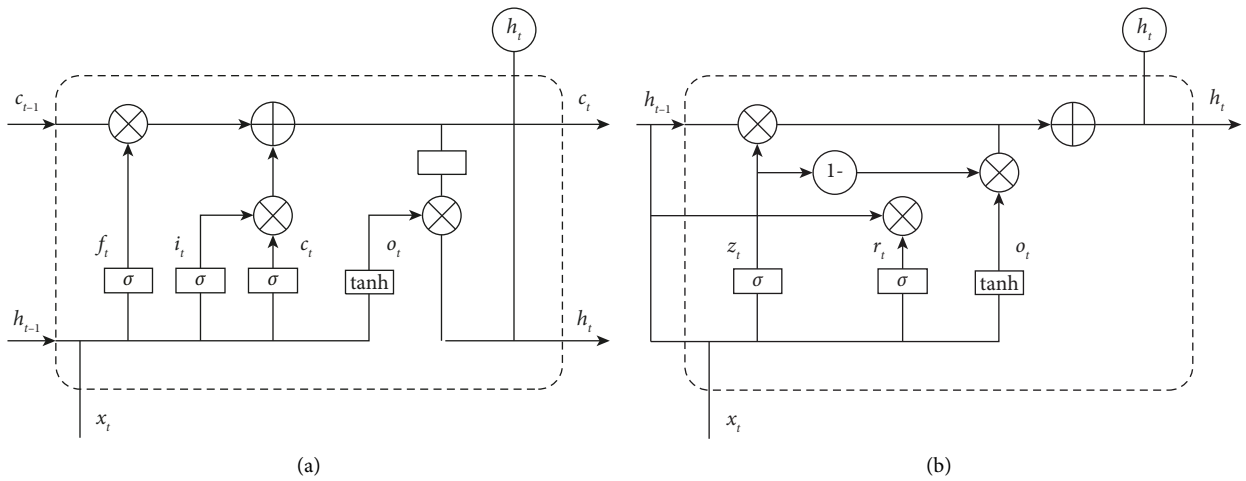


Figure 6: LSTM and GRU unit structure. (a) LSTM. (b) GRU.

3.3. Multilayer LSTM Network Construction and Performance Analysis. With the rise of artificial intelligence, some deep learning models have been proposed specifically for data prediction problems, such as the long short-term memory (LSTM) network and gate recurrent unit (GRU) network. The LSTM network was first proposed by Hochreiter and Schmidhuber in 1997. LSTM network is designed to solve the long-term dependence problem of deep learning, which is improved by recurrent neural network (RNN) [25, 26]. Now, it is widely used in the fields of text generation, speech recognition, and image generation description [27]. The GRU network is a type of recurrent neural network (RNN) [28, 29]. It is also proposed to solve the problems of long-term memory and gradients in back propagation, just like the LSTM. The internal network structure of LSTM and GRU is basically the same, and the biggest difference is shown as follows: LSTM network structure takes output gate o_t, input gate i_t, and forget gate f_t as the internal state, while GRU combines the latter two into update gate z_t, deletes the output gate, and adds reset gate r_t, which takes update and reset gate as the internal state. The details of the unit structure for both are shown in Figure 6.

These two deep learning models are substituted into the method of this paper, which is used to train the relationship between the input and output of cable insulation resistance of the source domain to reduce the influence of high voltage environment and signal injection on the measurement accuracy.

3.3.1. LSTM Network. Firstly, the LSTM network is brought in, and in order to accelerate the convergence rate of the model, a normalized layer is added in front of the first LSTM layer. The structural design is shown in Figure 7.

The LSTM model designed in Figure 7 is used as the pretraining model. Among them, there are 200 hidden layers, 1000 times of training, and the learning rate is 0.001. The training effect is shown in Figure 8. The pretraining network is evaluated, and the prediction effect of the test set was shown in Figures 9–11.

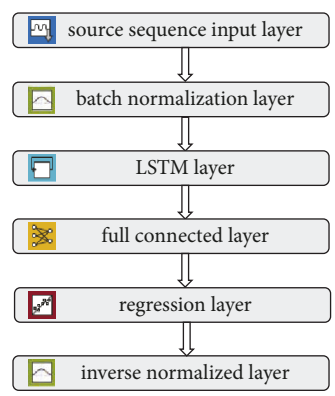


Figure 7: Source domain training model based on LSTM.

Figure 9 shows the measurement results of insulation resistance values corresponding to the two test sets. The error of the measured value compared with the real value and the corresponding relative error are shown in Figures 10 and 11. As can be seen from Figures 9–11, the effect of the pretraining model is effective, and the relative error of the insulation resistance value obtained on the test set can be controlled within 2.4%.

3.3.2. GRU Network. Similar to the method based on the LSTM network, a multilayer GRU model is also established, and a normalization layer is added to accelerate the convergence speed of the model. 200 hidden layers were set in the GRU network and trained for 1000 times, the learning rate was set to 0.001. The training effect is shown in Figure 12. The pretraining network is evaluated, and the prediction effect of the test set is shown in Figures 13–15.

Figure 13 shows the measurement results of insulation resistance values corresponding to the two test sets. The error of the measured value compared with the real value and the corresponding relative error are shown in Figures 14 and 15.

As can be seen from Figures 14 and 15, the relative error of the insulation resistance value obtained on the test set of

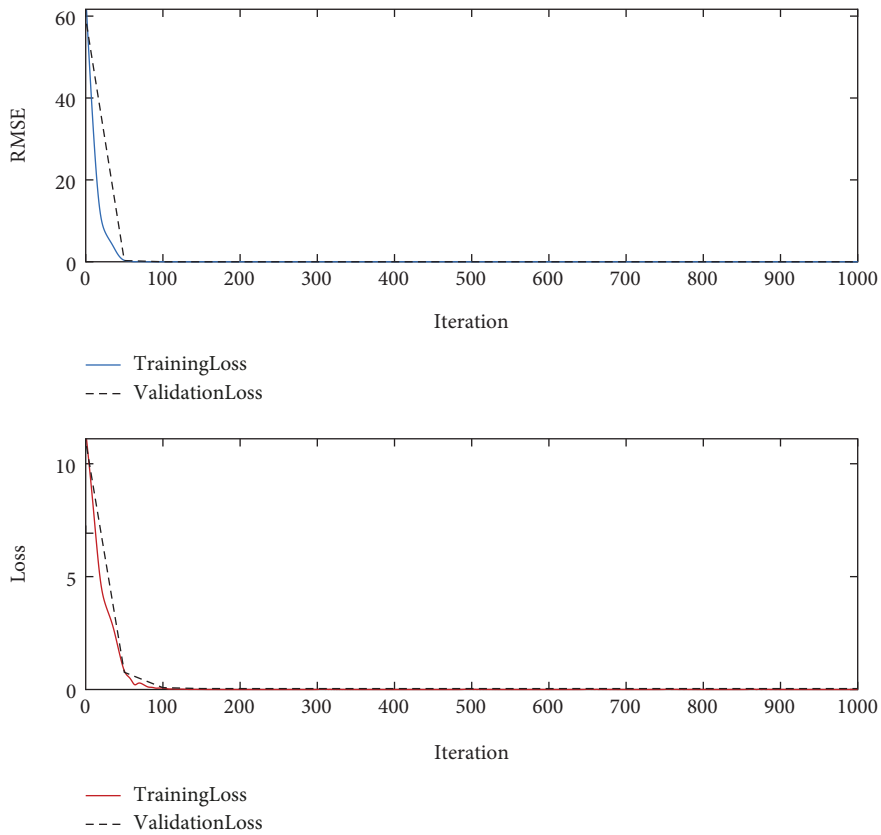


Figure 8: Pretraining effect of the source domain model by LSTM.

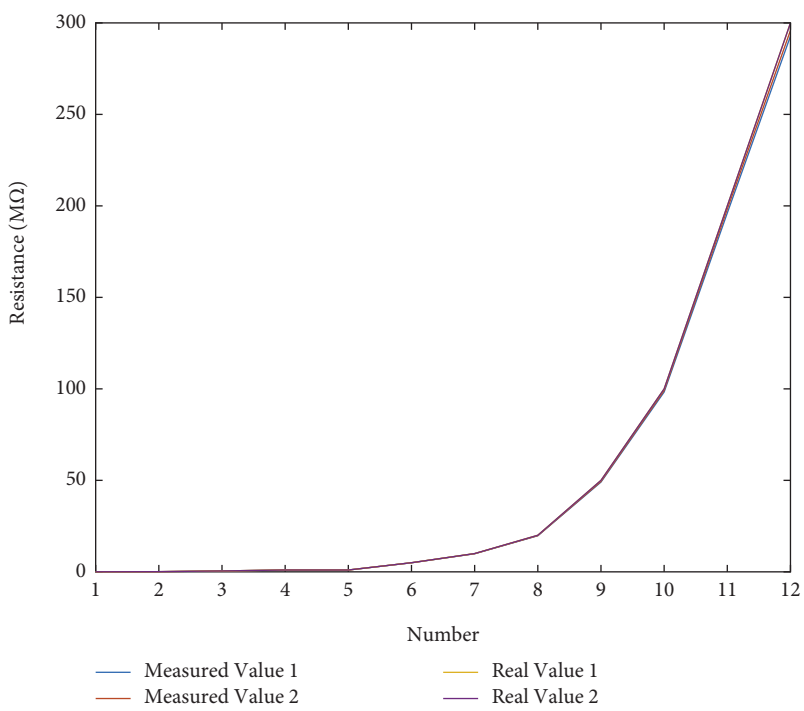


Figure 9: Measured values of insulation resistance.

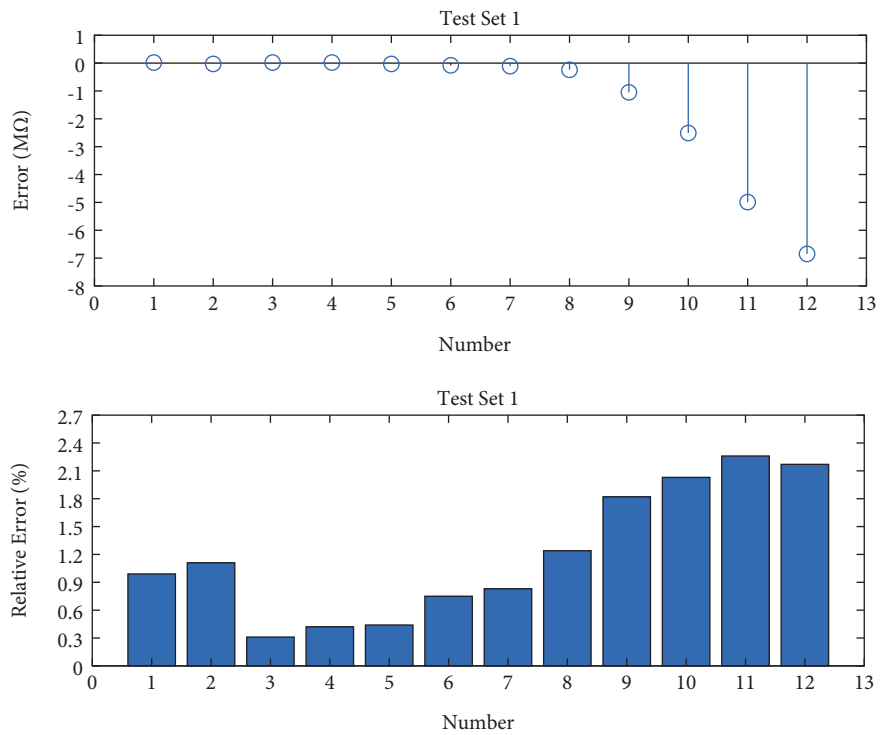


Figure 10: Diagram of insulation resistance measurement error and relative error in test set 1.

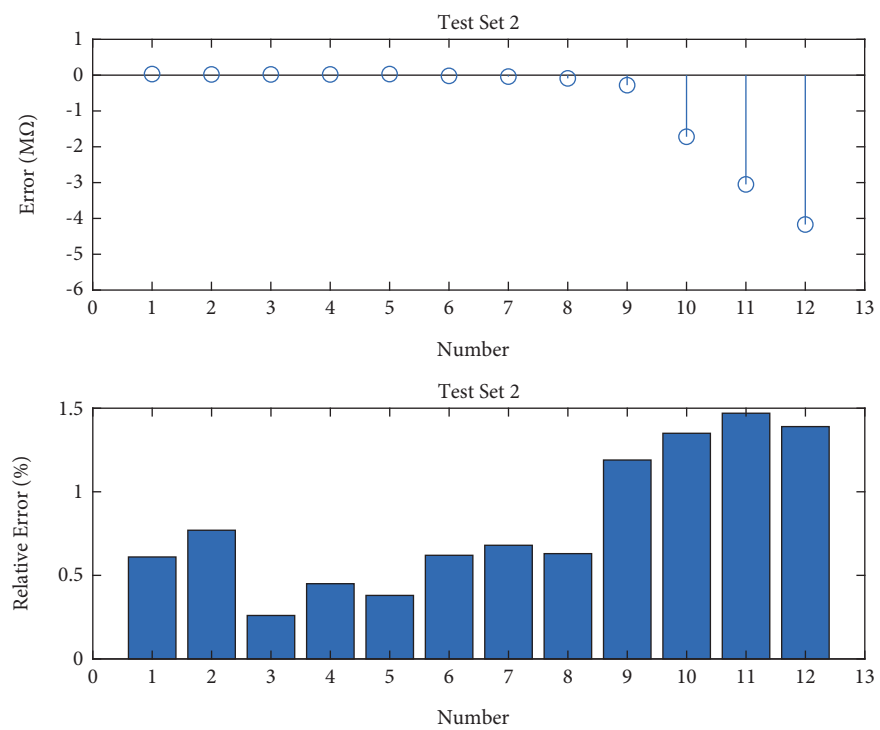


Figure 11: Diagram of insulation resistance measurement error and relative error in test set 2.

the model trained by GRU network is controlled within 3.5%. Compared with the LSTM model, it can be seen that the precision of insulation resistance measurement results of GUR network fluctuates greatly. Although the relative error is small in the section with low insulation resistance value, the relative error increases suddenly in the section with high insulation resistance value and the measurement effect is unstable. Therefore, a multilayer LSTM model is selected as the basic network architecture in this paper.

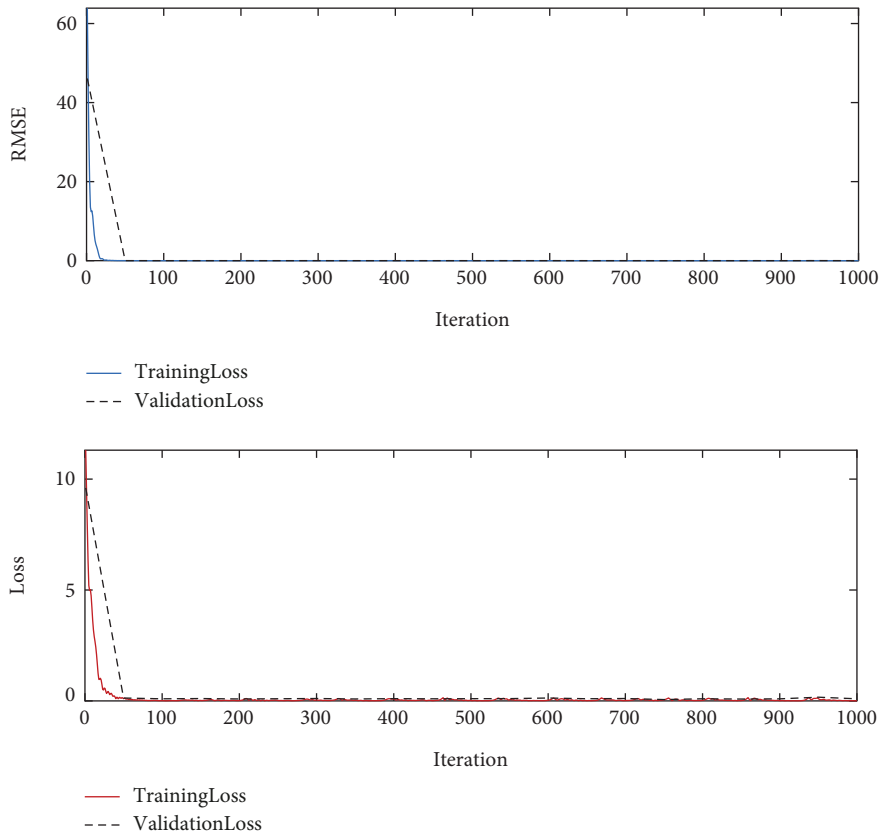


Figure 12: Pretraining effect of the source domain model by GRU.

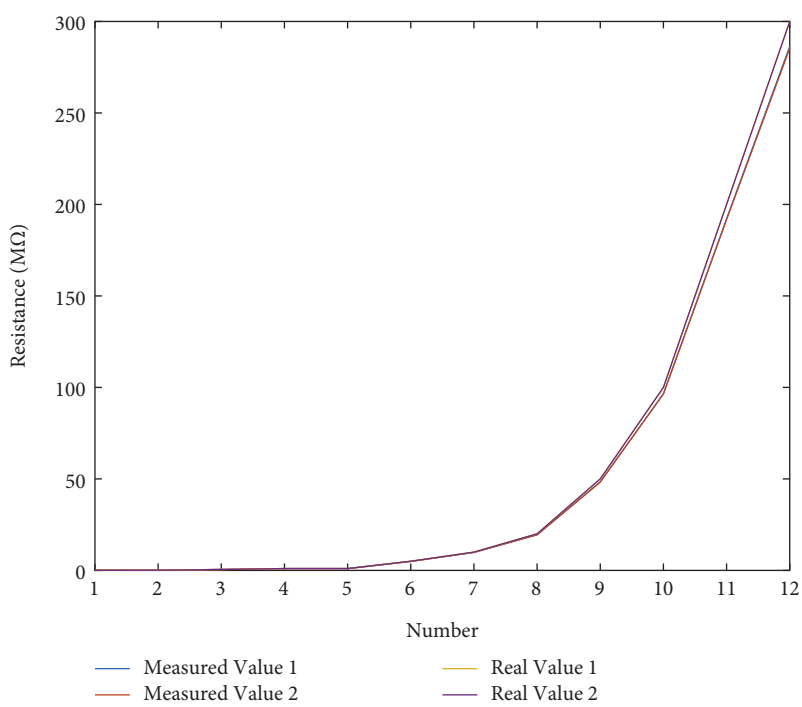


Figure 13: Measured values of insulation resistance.

3.4. Establishment of Transfer Learning Model. Transfer learning is a kind of machine learning approach, in which a pretrained model is reused in a task with another target model. The pretrained model is set as the source domain Ds, the training process of the pretrained model is the learning task Ts, the target model is set as the target domain Dt, and

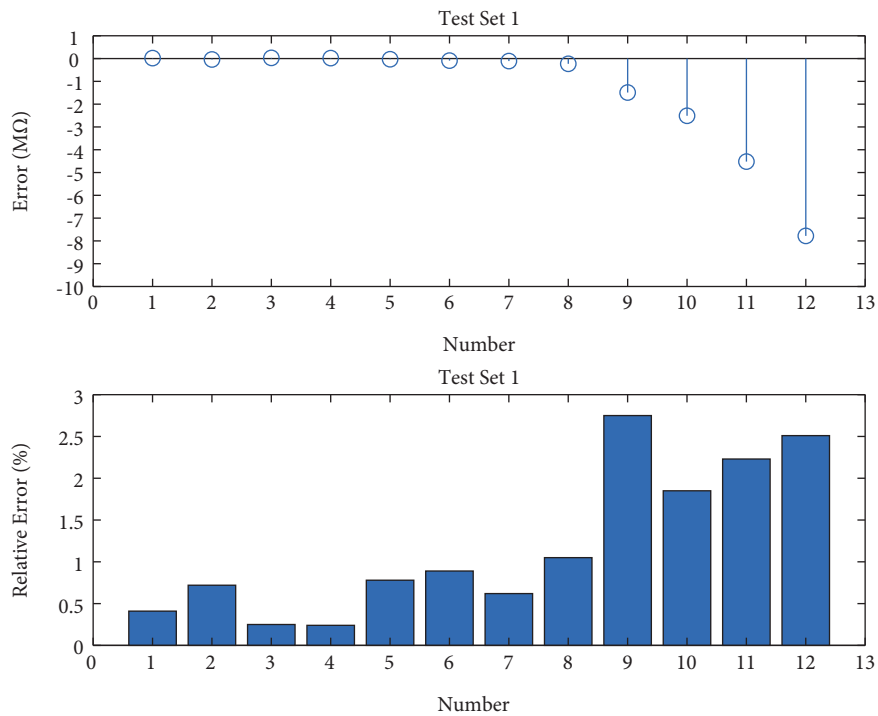


Figure 14: Diagram of insulation resistance measurement error and relative error in test set 1.

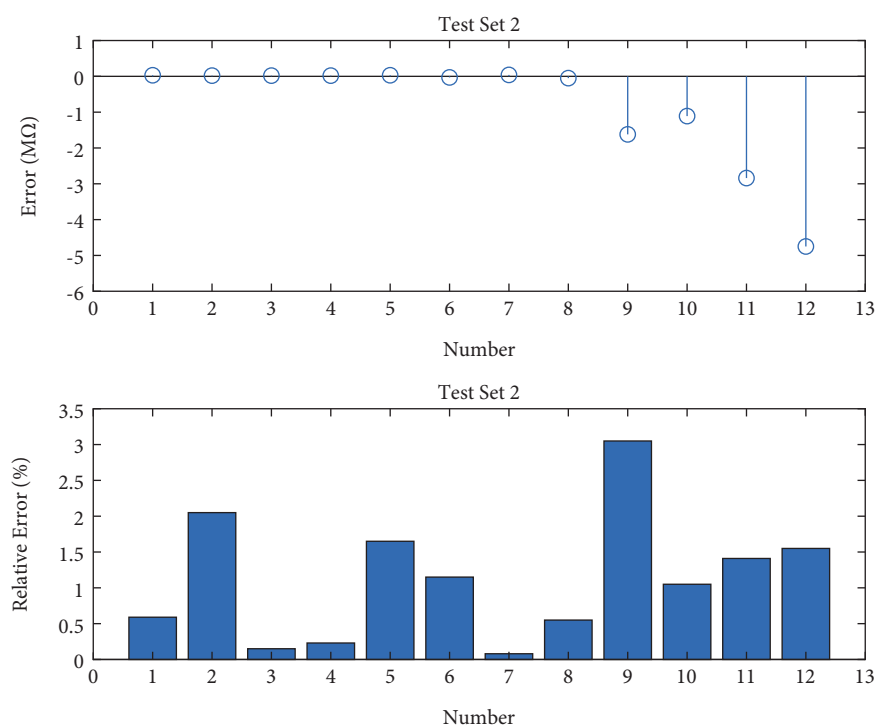


Figure 15: Diagram of insulation resistance measurement error and relative error in test set 2.

the training process of the target model is the learning task Tt . The source domain and the learning task are used to predict the rules of the target domain. It's worth noting that Ds ≠ Dt and Ts ≠ Tt .$D_T$and a learning task $T_T$on the target domain, $D_S$and$T_S$are used to learn the prediction function $f(\cdot)$ on the target domain and $D_S \neq D_T$ or $T_S \neq T_T$. Its core is to find the similarity or some mapping relationship between source domain D_S and target domain D_T [30, 31]. This paper adopts the parameter-based transfer learning method, and its basic idea is shown in Figure 16.

In the source domain, the LSTM model Net_S is trained with the source domain data, and the weight parameters of the lower

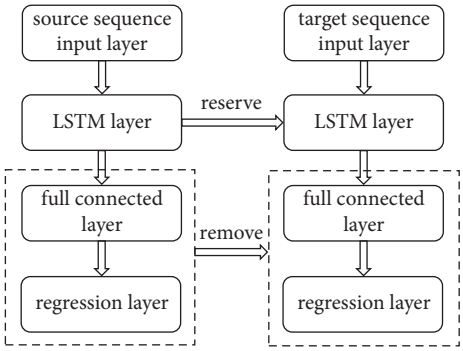


Figure 16: Schematic diagram of the parameter-based transfer learning idea.

layer of Net_S are fixed and applied to the LSTM model Net_T in the target domain. At the same time, the full connection layer and the regression layer of Net_T are rebuilt and the parameters are adjusted according to the input data of the target domain.

In this paper, the measurement of insulation resistance of the ANLS cable under low voltage condition is taken as the learning task in the source domain, and the measurement of insulation resistance of the ANLS cable under high voltage condition is taken as the learning task in the target domain. Resistance measurements at low voltage are relatively safe and readily available. The pretraining model is established after a lot of training, and the weight parameters of the pretraining model are applied to the target domain through the parameter-based transfer learning, which saves the training time and improves the measurement accuracy.

4. Measurement System Performance Analysis

4.1. Target Domain Data Acquisition. The purpose of transfer learning is to apply the model trained in advance to the target domain. In order to analyze the measurement performance of this design, an experimental platform of the ANLS is built. Figure 17 is a sinusoidal constant current dimmer, and Figure 18 is an ANLS circuit, which can be used to simulate high voltage conditions and actual working conditions (in addition to loop length).

In order to simulate the change of the ANLS circuit cable insulation resistance value, the target resistance B with known resistance value is made, as shown in Figure 19. It is connected in parallel to the cable of the light circuit, and different resistance values can be obtained by rotating the switch to simulate the change of insulation resistance.

Connect target resistance B to the "Target Plate" as shown in Figure 5. Place the device at room temperature and turn on the power. Then, turn on the switch of the target resistance B in Figure 19 and measure according to the steps in Section 2.2 to obtain a set of data. After cooling off the power and the equipment, measurements are made again after turning on the power and repeated for 6 times to form target domain data $\{R_{xt}, R_{yt}\}_1, \{R_{xt}, R_{yt}\}_2, \ldots, \{R_{xt}, R_{yt}\}_6$. The six sets of sequence data are divided into training set and test set according to 4 : 2.

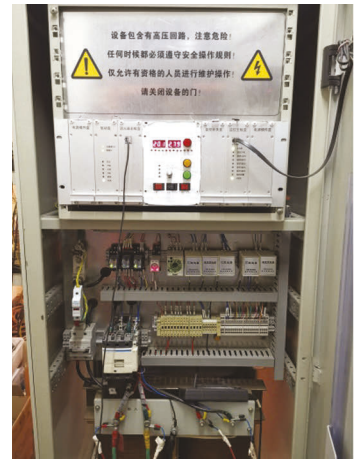


Figure 17: Sinusoidal constant current dimmer.

Figure 18: An ANLS circuit.

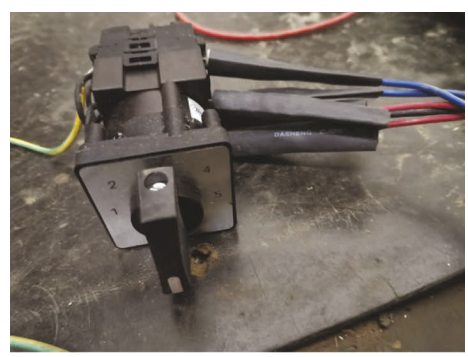


Figure 19: The target resistance B.

4.2. Target Domain Construction and Performance Analysis. Input the prepared target domain data into the model for training. The training results are shown in Figure 20, and the measurement results on test sets are shown in Figure 21.

Figure 21 shows the measurement results of insulation resistance values corresponding to the two test sets. The error of the measured value compared with the real value and the corresponding relative error are shown in Figures 22 and 23. As can be seen from Figures 21–23, the effect of the pretraining model is effective, and the relative error of the

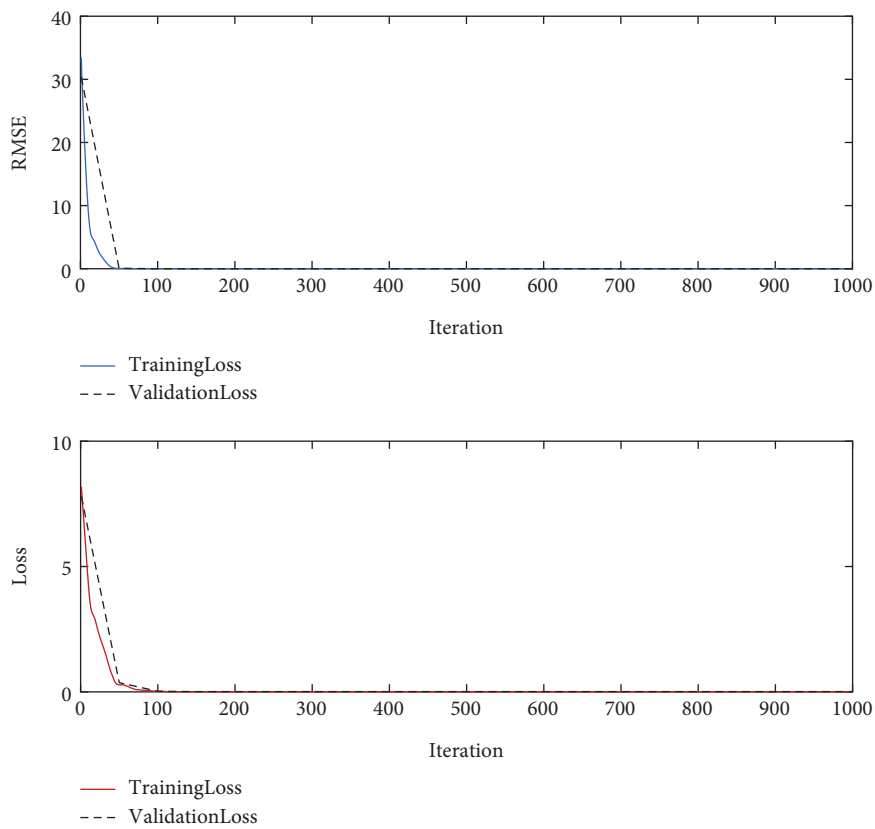


Figure 20: Model training results of target domain.

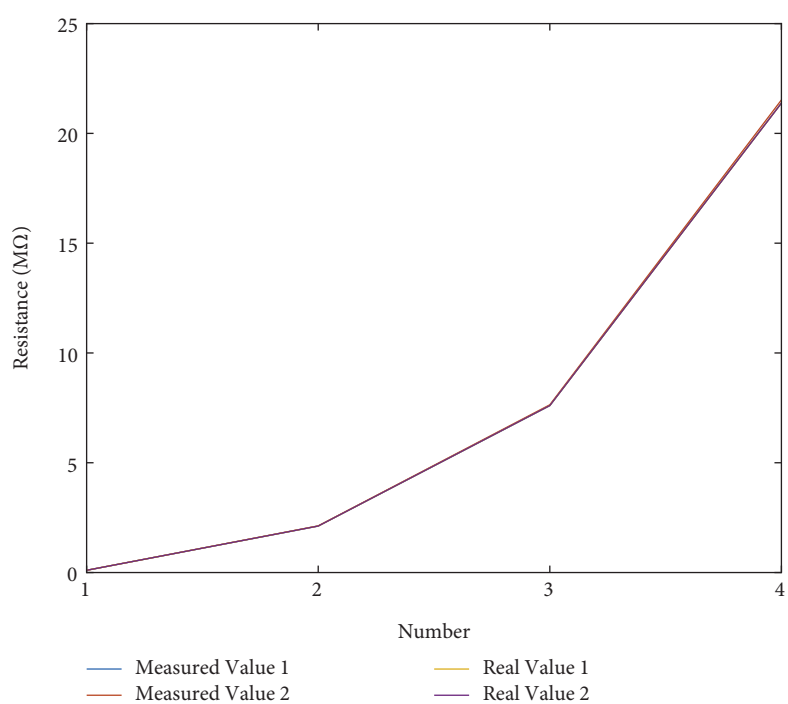


Figure 21: Measured values of insulation resistance.

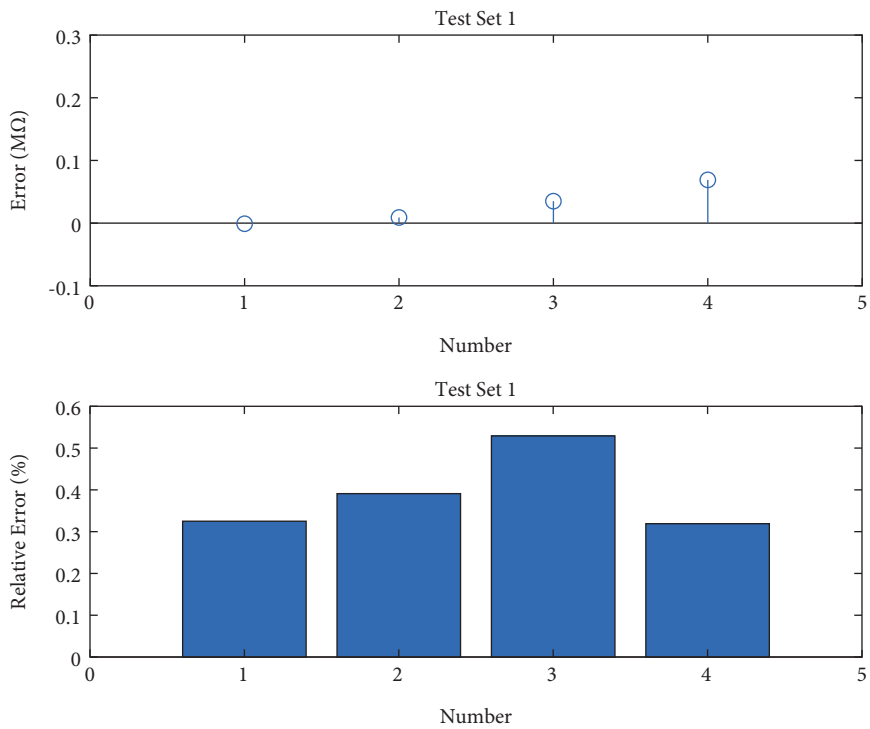


Figure 22: Diagram of insulation resistance measurement error and relative error in test set 1.

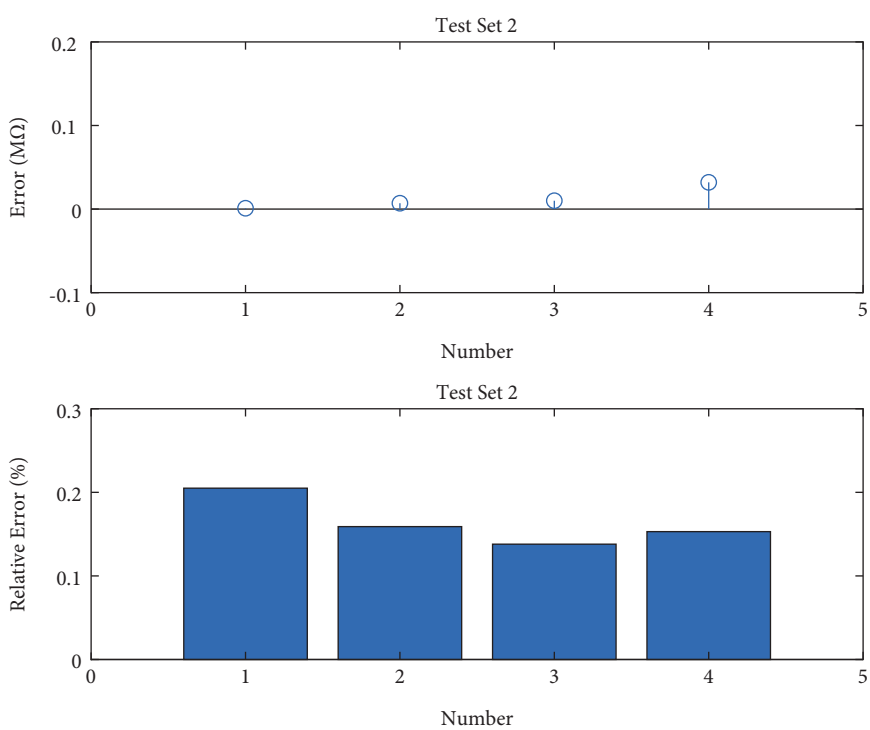


Figure 23: Diagram of insulation resistance measurement error and relative error in test set 2.

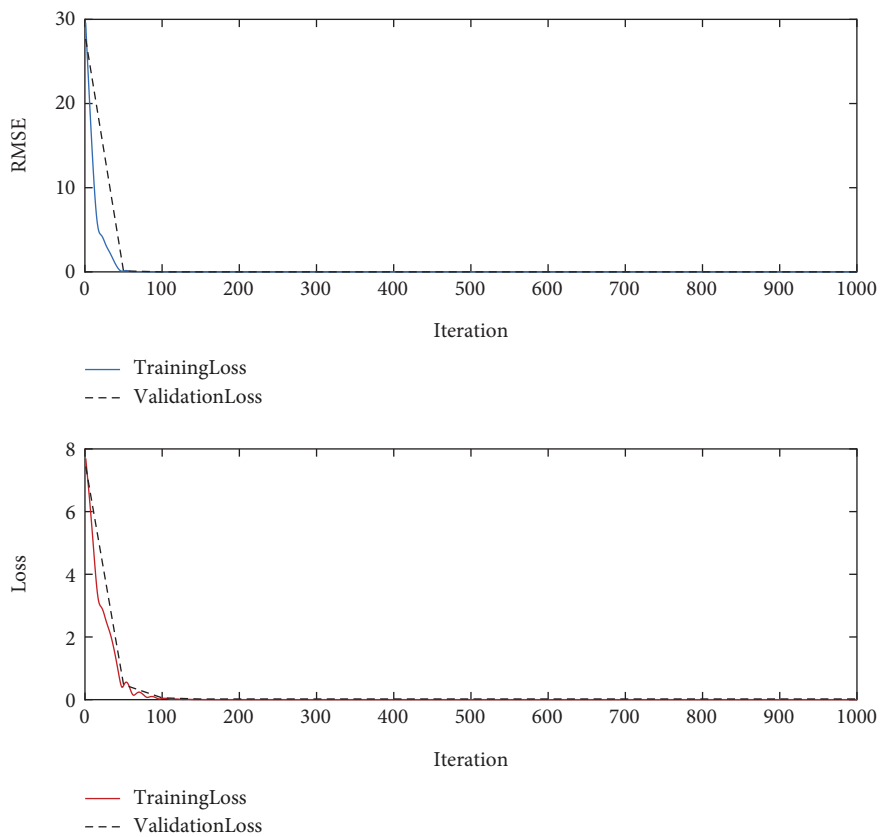


Figure 24: Training effect of the LSTM model without transfer learning.

insulation resistance value obtained on the test set can be controlled within 0.6%.

4.3. Overall Performance Comparison. In order to analyze the performance of the system, the insulation resistance of the ANLS circuit cable is tested by many ways. The target resistance B in Figure 14 is selected as the reference object. In this section, insulation resistance measurement results of three different methods, including switching time-sharing method, deep learning without transfer learning, and deep learning and transfer learning proposed in this paper, will be compared. Also, connect target resistance B to the "Target Plate" as shown in Figure 5. The change of insulation resistance value is simulated by rotating the switch under high voltage conditions for measurement.

Method 1 (switching time-sharing method). The insulation resistance value of the ANLS circuit cable is measured directly by the method in Section 2.2.

Method 2 (deep learning without transfer learning). Place the device at room temperature and turn on the power. Then, turn on the switch of the target resistance B in Figure 19 and measure according to the steps in Section 2.2 to obtain a set of data R_x. R_x is taken as the input data and the true value of the target resistance B is taken as the output data. The LSTM model in Figure 7 is used for training, and the model training result is shown in Figure 24. The model is directly applied to the measurement without subsequent transfer learning.

Table 1: Insulation resistance measurement results of the three methods.

True value of insulation resistance (MΩ)	Switching time-sharing method (MΩ)	Deep learning without transfer learning (MΩ)	Deep learning and transfer learning (MΩ)
0.1027	0.1040	0.1020	0.1025
2.1189	2.0943	2.1331	2.1222
7.6065	7.3363	7.6748	7.6173
21.3790	19.9578	21.5793	21.4121

Method 3 (deep learning and transfer learning). According to the method presented in this paper, the insulation resistance value is measured according to the steps in Section 3.1.

Results of the three measurement methods are shown in Table 1. To observe the measurement effect more intuitively, the insulation resistance values obtained by the three methods are compared and their relative errors are shown in Figure 25.

It can be seen from Figure 25 that the relative error of insulation resistance value obtained by the deep learning and transfer learning method is the smallest among the three measurement methods. When the insulation resistance value is small, the relative error of the three measurement

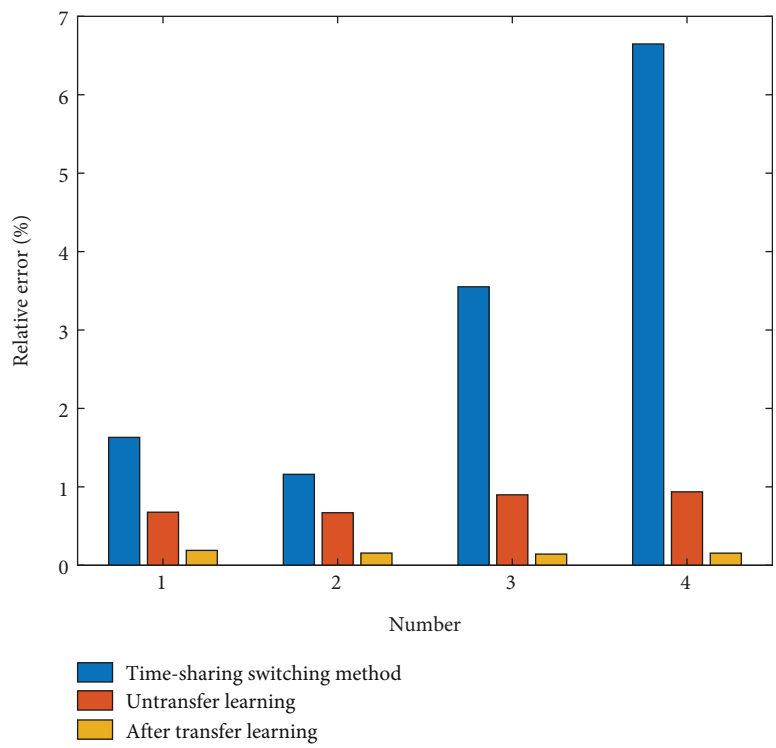


Figure 25: Relative error comparisons of insulation resistance values.

methods is relatively small. When the insulation resistance value of the cable to be measured increases, the measurement effect of the three methods becomes different. The switching time-sharing method has the largest relative error, while the deep learning and transfer learning method has the smallest relative error. The cable insulation resistance of the ANLS is generally extremely high under the actual working conditions. Therefore, the method proposed in this paper can improve the measurement accuracy of insulation resistance. Especially for high voltage and high insulation, the measurement effect is more stable. At the same time, the introduction of transfer learning can share the pretrained model parameters under low voltage conditions with the online measurement model under high voltage conditions, which saves the training time and increases the safety factor in the detection process and also solves the problem of small sample data. To sum up, the relative error of the deep learning and transfer learning method proposed in this paper is small and controlled within 0.6%, which indicates that this method can better reflect the real-time situation of the circuit cable insulation resistance and has a better indication effect on the current working state of the ANLS.

5. Conclusions

In this paper, deep learning and transfer learning are firstly applied in the measurement of insulation resistance in the ANLS. Aiming at reducing the influence of high voltage environment and voltage injection signal on the measurement accuracy of insulation resistance, a multilayer LSTM model is established, which is used for pretraining, and a normalized layer is added in front of the first LSTM layer to accelerate the data convergence rate. Then, based on the pretrained LSTM model, the weight parameter of the pretrained model is shared through transfer learning to solve the problem of small sample data. Simulation and experimental results show that the influence of the injection voltage signal and stray current on the measurement results is reduced. It has obvious advantages for the measurement of high insulation resistance value, and the relative error can reach less than 0.6%. Moreover, the measurement of insulation resistance value of the ANLS circuit cable has not formed an industry standard. The application of deep learning and transfer learning in this paper provides a new idea for the measurement method.

There are still new challenges: the weight of the pretrained model is shared for transfer learning in this paper, and the feature-based transfer learning method can be further considered. The data distribution under low pressure and high pressure conditions is mapped to the same feature space by establishing the joint distribution function of source domain data and target domain data to enhance domain adaptability.

Data Availability

The processed data used to support the findings of this study have not been made available because the data also form part of an ongoing study.

Conflicts of Interest

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

Acknowledgments

This work was supported in part by the Doctoral Research Start-up Foundation of Liaoning Province under Grant 2019-BS-040.

References

[1] N. K. Chawda, H. W. Schuett, and R. Wiechert, *Devices, Methods, and Systems for Alternating Current Circuits for Airfield Lighting*, Patent application 2018/0249540A1, U.S.A, 2018.

[2] Y. G. Yu, *Airfield Lighting Sustem. Patent Application 9131584B2*, U.S.A, 2015.

[3] J. Pablo Conti, "Airport of the future," *Engineering & Technology*, vol. 10, no. 1, pp. 66–68, 2015.

[4] T. Kurihara, E. Tsutsui, E. Nakanishi et al., "Investigation of water tree degradation in dry-cured and extruded three-layer (E-E Type) 6.6-kV removed XLPE cables," *Electrical Engineering in Japan*, vol. 198, no. 4, pp. 37–50, 2016.

[5] K. Steppe, "The potential of the tree water potential," *Tree Physiology*, vol. 38, no. 7, pp. 937–940, 2018.

[6] IEEE standard, *International Standard Norme International*, 2019.

[7] C. Zhang, C. Li, L. Nie, Z. Jing, and B. Han, "Research on the water blade electrode method for assessing water tree resistance of cross-linked polyethylene," *Polymer Testing*, vol. 50, pp. 235–240, 2016.

[8] P. Olszowiec, *Insulation Measurement and Supervision in Live AC and DC Unearthed Systems*, p. 15, Springer International Publishing, Berlin, 2014.

[9] R. Polansky and M. Polanská, "Testing of the fire-proof functionality of cable insulation under fire conditions via insulation resistance measurements," *Engineering Failure Analysis*, vol. 57, pp. 334–349, 2015.

[10] V. Verdingovas, S. Joshy, M. S. Jellesen, and R. Ambat, "Analysis of surface insulation resistance related failures in electronics by circuit simulation," *Circuit World*, vol. 43, no. 2, pp. 45–55, 2017.

[11] Y. Shen, A. Liu, and G. Cui, "Design of online detection system for insulation resistance of electric vehicle based on unbalanced bridge," in *Proceedings of the 2019 IEEE Innovative Smart Grid Technologies-Asia (ISGT Asia)*, pp. 537–542, IEEE, Chengdu, China, May 2019.

[12] Y. H. Chiang, W. Y. Sean, C. Y. Huang, and L. H. Chiang Hsieh, "Adaptive control for estimating insulation resistance of high-voltage battery system in electric vehicles," *Environmental Progress & Sustainable Energy*, vol. 36, no. 6, pp. 1882–1887, 2017.

[13] X. Jian and G. Hong, "An intelligent dual-voltage driving method and circuit for a common rail injector of heavy-duty diesel engine," *IEEE Access*, vol. 5, pp. 27681–27689, 2017.

[14] C. Song, Y. Shao, S. Song et al., "Insulation resistance monitoring algorithm for battery pack in electric vehicle based on extended Kalman filtering," *Energies*, vol. 10, no. 5, p. 714, 2017.

[15] N. Shtabel, L. Samotik, and E. Mizrakh, "Fluxgate direct current sensor for real-time insulation resistance monitoring," in *Proceedings of the 2019 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM)*, Sochi, Russia, March 2019.

[16] X. Jian and G. Hong, "An online adaptive monitoring system for insulation resistance of electric vehicles," in *Proceedings of the 2019 IEEE 2nd International Conference on Information Systems and Computer Aided Education (ICISCAE)*, pp. 37–41, Dalian, China, September 2019.

[17] S. Hao, X. Zhang, R. Ma, and G. Jie, "Fault line selection method for small current grounding system based on improved GoogLeNet," *Power System Technology*, vol. 2021, pp. 1–9, Article ID 678120, 2021.

[18] Y. Li, Y. Wang, Y. Zhang, and J Zhang, "Diagnosis of interturn short circuit of permanent magnet synchronous motor based on deep learning and small fault samples," *Neurocomputing*, vol. 442, pp. 348–358, 2021.

[19] T. J. Donnelly and S. D. Pekarek, "Modeling and control of an LED-based airfield lighting system," in *Proceedings of the IEEE Power and Energy Conference at Illinois (PECI)*, pp. 1–5, Champaign, IL, USA, April 2018.

[20] I. V. Komlev and Komlev, "On associated data and power channels in the cable ring of an airfield lighting system," *Russian Electrical Engineering*, vol. 90, no. 5, pp. 412–416, 2019.

[21] X. Ma, T. Li, J. Li, and F. Chen, "Airport-specific light-guiding taxiing guidance sign," in *Proceedings of the 2020 IEEE 3rd International Conference on Electronics Technology (ICET)*, pp. 207–2111, Chengdu, China, May 2020.

[22] Y. Yan, S. D. Sudhoff, J. Nadel, and D. W. Gallagher, "Modeling and stability analysis of a fixture-centric airfield lighting System," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 55, no. 6, pp. 3479–3491, 2019.

[23] M. Liu, "Design of airport navigation aid lighting system based on LED," *International journal of science*, vol. 5, pp. 102–105, 2018.

[24] Y. Zhang, *The Design and Implementation for Real-Time Online Monitoring Unit of Insulation Resistance in Airfield Lighting Circuit*, M. S. thesis, Dalian Jiaotong University, Dalian China, 2015.

[25] S. Li, X. You, and S. Liu, "Multiple ant colony optimization using both novel LSTM network and adaptive Tanimoto communication strategy," *Applied Intelligence*, vol. 51, no. 8, pp. 5644–5664, 2021.

[26] X. Wu, Z. Du, Y. Guo, and H. Fujita, "Hierarchical attention based long short-term memory for Chinese lyric generation," *Applied Intelligence*, vol. 49, no. 1, pp. 44–52, 2019.

[27] T. Bogaerts, A. D. Masegosa, J. S. Angarita-Zapata, E. Onieva, and P. Hellinckx, "A graph cnn-lstm neural network for short and long-term traffic forecasting based on trajectory data," *Transportation Research Part C: Emerging Technologies*, vol. 112, pp. 62–77, 2020.

[28] Y. Zhang, Y. Xin, Z. Liu, and M. Guijun, "Health Status Assessment and Remaining Useful Life Prediction of Aero-Engine Based on BiGRU and MMoE," *Reliability Engineering & System Safety*, vol. 220, 2022.

[29] A. Flores, H. Tito-Chura, and V. Yana-Mamani, "An ensemble GRU approach for wind speed forecasting with data augmentation," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 6, pp. 569–574, 2021.

[30] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345–1359, 2010.

[31] M. Wang and W. Deng, "Deep visual domain adaptation: a survey," *Neurocomputing*, vol. 312, pp. 135–153, 2018.

Hindawi
Mathematical Problems in Engineering
Volume 2022, Article ID 3986247, 10 pages
https://doi.org/10.1155/2022/3986247

Research Article

Optimization of Skipping Rope Training Method Based on Chaotic Logistics

Han Li and Yuxia Wang

Department of Physical Education, Ordos Institute of Technology, Ordos 017000, Inner Mongolia, China

Correspondence should be addressed to Yuxia Wang; wangyuxia@oit.edu.cn

Received 24 June 2022; Revised 13 July 2022; Accepted 16 July 2022; Published 23 August 2022

Academic Editor: Long Wang

Skipping training belongs to a competitive event, which can not only lose weight but also improve the physical quality of trainers. At present, there are some problems in skipping training, such as unreasonable plan and unsatisfactory actual training effect. The original rough set method cannot solve the analysis of multivariate data in skipping training, and the ability to evaluate the training effect of skipping is poor. Aiming at the problems existing in the training of skipping rope, this paper puts forward a skipping rope training method based on a chaotic logistics algorithm, aiming at improving the training effect of skipping rope. First, the chaos theory is used to classify the training data. Different data classifications correspond to different training results, which eliminates irrelevant information in the training scheme and reduces the amount of single data analysis. Finally, the data after classification are optimized in a single stage, and the optimal results of different training items are obtained and summarized. After the MATLAB test, the chaotic logistics algorithm is superior to the rough set method in accuracy and convergence speed and meets the actual training needs. Therefore, the chaotic logic analysis method proposed in this paper is suitable for the evaluation of skipping training.

1. Introduction

At present, the state has increased its support for sports training. Skipping rope training is a national competitive activity, which has the dual effects of competition and leisure. As a national sport, skipping rope can not only improve the physical fitness of the masses but also belong to a leisure and entertainment project. At present, the training method of skipping rope is unreasonable, which cannot play the role of skipping rope. Therefore, looking for an effective exercise method is an urgent problem to be solved at present, which has very important practical significance. Ant colony algorithm and regression analysis methods cannot carry out continuous analysis and cannot achieve massive data analysis. Chaotic regression analysis method, which can reduce the amount of data analysis, more accurate data analysis, for skipping training method analysis, has a very obvious analysis advantage. Chaos theory is used to classify and analyze the data of skipping rope training, and logistic regression analysis is used to determine the main factors affecting skipping rope training. The integration of chaos theory and logistic regression can give full play to the advantages of massive data analysis in the early stage and accurate identification of later results. According to the survey data of skipping competition in 2021, skipping training accounts for 12.6% of sports competition, while the evaluation accuracy of skipping competition effect is only 45.3%, and some skipping competition training schemes are contrary to science. Domestic scholars have applied artificial normal, artificial speech, artificial portrait, and other algorithms to skipping training, aiming at providing data support for skipping training analysis [1]. On the basis of previous studies, this paper constructs a chaotic logistics algorithm to evaluate the training effect of skipping rope and find out its existing problems and shortcomings.

Because the data of the skipping training effect mainly exist in semistructured form and present nonlinear relationship, rough set cannot accurately evaluate skipping

training. Chaotic logistics algorithm has better semi-structured data analysis ability, analyzes the mapping characteristics of various factors in skipping training, and carries out repeated iteration and self-training [2] to improve the accuracy of skipping training. Some scholars combine Apriori with chaos theory to build a prediction model of skipping rope training [3] and carry out actual case analysis. The results show that the combined analysis of chaos theory and the Apriori algorithm is better, and all indexes are significantly better than the Apriori algorithm [4]. The reason is that the chaotic logistics algorithm belongs to a comprehensive analysis method, which integrates training mechanism, classifies a series of problems in skipping training, and realizes comprehensive chaos analysis. Therefore, based on the chaotic logistics algorithm, this paper evaluates the training effect of skipping rope [5]. Chaos logistics algorithm research is mainly from the following aspects to improve (1) training parameters adjustment and training strategy change. There are also scholars to skipping training evaluation strategy research and compare three skipping training strategies, and the results show that the selection of appropriate strategy can achieve higher training results in a short time [6]. (2) Some scholars put forward an adaptive logical logistics algorithm, which integrates K-mean classification into logistic logistics, but this algorithm cannot classify the data of skipping rope training [7], and the evaluation speed of training results is slow [8]. From the point of view of time delay, some scholars use a logical logistics algorithm to analyze the training results of skipping rope and get the results in a dynamic and random way [9], but the accuracy of the results of this algorithm is cross. (3) Compared with classical theories, such as Bayesian theory and rough set, chaotic logistics algorithm has strong data sensitivity [10], which is suitable for massive data points. The specific survey results are shown in Table 1.

From Table 1, we can see that the research trend of skipping training is increasingly obvious, and it is increasing year by year. The data for 2007 were the smallest, followed by 2021 and finally 2010. However, the number of studies gradually increased in 2011, 2016, and 2020. On the whole, the research on skipping training methods is above year by year and presents ups and downs of the development trend. Therefore, the research demand on skipping training methods is increasing year by year, and it has become a research hotspot. All this research has very important theoretical and practical significance. All the above studies show that the single algorithm is not ideal in terms of calculation accuracy, search time, and overall convergence. In this paper, chaos theory and logics algorithm are integrated, and the chaos logics algorithm is proposed to analyze the results of skipping training and verify them.

Table 1: Research situation of skipping rope training.

Time	Number of studies	Time	Number of studies
2006	45.92	2014	55.10
2007	15.31	2015	68.37
2008	61.22	2016	85.71
2009	61.22	2017	32.65
2010	27.55	2018	82.65
2011	81.63	2019	56.12
2012	67.35	2020	80.61
2013	52.04	2021	20.41

2. Concepts Related to Research

2.1. Chaos Theory. Chaos theory was first put forward in the early 19th century [11], which belongs to a statistical analysis method, mainly solves the problems of data description, data analysis, and result prediction, and is widely used in ships, aviation, and other fields. Chaos theory is embodied in the persistent description of data. Different points are randomly selected for analysis. Its principle and structure are as follows: let g_{it} be a data set in i multidimensional space and t time, and any x belongs to g_{ij}, and then, the data can be analyzed by chaos theory.

Suppose 1: $(P_j(x) = \sum_{i,t=1}^{n}\{x_{it}^{j} \cdot k \cdot \varphi(k)\})$, $(\varphi(k) = C(x,\varepsilon))$, and then, $(P_j(x))$ is the best data set. Among them, $(C(x,\varepsilon))$ is the adjustment coefficient of x and ε, ε is the eigenvalue in logistic logistics at this time $(x_{it} = \sin(\tan\pi t/\rho))$, and ρ is the least common multiple set, and then, at this time $(\sin(\tan\pi t/\rho)^{j} < \sum_{i,j,t=1}^{n} x_{it}^{j} \cdot k)$, the data link in chaos theory is the shortest. The related theories of logistic logistics are as follows:

Theorem 1. *If the characteristic function of any data value is $(P_n(x))$, then the calculation formula of $(\varphi(x \cdot k))$, $(f(\cdot) \in B_t)$, and the evaluation result function is shown in the following formula:*

$$\int_{x \in G_{ij}} f(x)dx < Q(f(\widehat{x})) \cdot \varphi(k). \tag{1}$$

Among them, $(Q(f(x)))$ is the overall evaluation results of $f(\cdot)$.

Theorem 2. *If the following conditions $f(x)$ and $(f(x) < \varepsilon)$ are satisfied, any function value in logistic logistics is within the integral of the function $f(x)$, and the error is less than $(lin(1/x))$.*

Theorem 3. *Assume that the eigenvalues x_{it} of any data are normally distributed on X and Y axes, then the formula for calculating the deviation of any point is shown in the following formula:*

$$D(x, f(x)) = \text{lin}\sqrt{\log\left(P_j(x)\right)}. \tag{2}$$

Among them, $D(x, f(x))$ is in the set [0, 1].

From the above analysis, it can be seen that using the approximate integral of chaos theory, the relationship between the amplitude of evaluation results and data results can be obtained, and the massive training data in space can be reduced. Therefore, the functions processed by chaos theory can be used for logistics analysis, which does not provide a good theoretical basis for the calculation of evaluation results and reduces the influence of unstructured data. From Theorem 3, we can see that the error of data in chaos theory is $lin(1/x)$, and the average error is

$lin(1/\sqrt{x} \cdot {}^{t}\sqrt{\log(P_j(x))})$, which shows that the single-phase error has nothing to do with the average error and needs to be controlled separately. Chaos theory and logistics algorithm are combined to control the amplitude of the evaluation results, which verifies the ability of chaos theory to control the training data and the superiority of this method. In this paper, this method is used for reference to optimize the training data of skipping rope and analyze it continuously.

2.2. Logistic Model. In this paper, chaos theory is chosen to optimize logistics model. This method belongs to a classification method, which is chaotic and adaptive, and can realize self-training [12]. Logistics analysis constantly revises the membership relationship between classification set and complete set. Logistics model is defined by the "IF" judgment mode, and its classification process is as follows.

If $x_i \in g_{ij}$ and $M(x) = C(x_i, x_{i-1})$, then $y = \sum_{i=1}^{n} y_i = \sum_{i=1}^{n} M(x_i) \cdot \lambda(x_i)$.

Among them, $\lambda(x_i)$ is a classification function, g_{ij} is classification set, $M(x)$ is constraint function, $C(x_i, x_{i-1})$ is relationship between any adjacent data, and x and y are classification results. The training part of chaos theory is clear, the training data is chaotic, and the evaluation result data is clear, so we can use chaos theory to reason the evaluation results and get the result evaluation combination. Assuming that any training variable is x_i, the membership relationship between each training variable x_i and the result evaluation variable y_i can be obtained by constraint rule M as shown in the following formula:

$$C \cdot g_{ij} = \sum_{i,j=1}^{n} \exp\left\{\frac{\left(x_i - c_{ij}\right)}{b_{ij}}\right\}. \tag{3}$$

Among them, c_{ij} and b_{ij} are the eigenvalues of membership functions, C is a relational function, and g_{ij} is a chaotic set. According to the membership relationship, the classification calculation formula can be obtained, as shown in the following formula:

$$C \cdot g_{ij} = \delta \cdot \sum_{i,j,k=1}^{n} g_{ij}^{k}(x), \tag{4}$$

where δ is the classification adjustment coefficient, and k is the derivative coefficient. According to formulas (3) and (4), the evaluation value of the classification result can be obtained, as shown in the following formula:

$$y = \frac{\left[\delta \cdot \sum_{i,j,k=1}^{n} g_{ij}^{k}(x) \cdot f\left(P_j(x)\right)\right]}{\max \sum_{i,j=1}^{n} g_{ij}^{k}(x)}, \tag{5}$$

where $\max \sum_{i,j=1}^{n} g_{ij}^{k}(x)$ is the maximum value of $g_{ik}^{k}(x)$, which represents the maximum value of the sum of all $g_{ij}^{k}(x)$, and $f(P_j(x)$ represents the total scheme of skipping training, which is an arbitrary skipping scheme. Chaos theory not only shortens the processing time of skipping training data but also increases the amount of single data processing. Therefore, the model constructed in this paper can classify a large number of training data [13] and analyze them continuously to form a continuous data link. In formula (5), the value is y, the relation function is C, and the classification coefficient is δ.

3. Construction of Skipping Training Model

3.1. Initial Skipping Training Data. The chaotic logistics model constructed in this paper can be used for multidimensional and holistic analysis, reduce single-phase convergence, improve the search ability of feature data, and shorten the analysis time of training results. In the evolution strategy of data link in initial skipping training and multidimensional skipping training, convergence threshold and weighting factor are used to promote the association evolution of different data values and get the best result.

The preprocessing of data by chaos theory: according to the classical logistics algorithm, if the data link is randomly distributed in the initial skipping training and the uncertainty is very strong, the data link calculation in skipping training will fall into single-phase optimum, which will reduce the evaluation time of the overall training effect and increase the error rate of the evaluation results, in order to expand the scale of data link and increase the types of data in skipping training. In this paper, using the classical logistics principle for reference, the number of data links in skipping training of initial data is expanded, and the diversity of data links in skipping training is improved. The results are shown in Figures 1 and 2.

Among them, the data in Figure 1 are the skipping training data processed by chaos theory, which is scattered in the whole space and does not change greatly with the increase of data volume. At the same time, the position of data is relatively stable, showing a planar situation, which shows that chaos theory maps data and highlights the numerical attributes of data.

As can be seen from Figure 2, the distribution of data in space is relatively discrete, and it changes up and down, indicating that data not only have numerical attributes but also have other attributes. In the process of skipping rope training, only numerical attributes are analyzed. Therefore, the chaotic regression analysis method not only extracts the data numerically but also integrates other attributes into the mapping, resulting in better data analysis results. Figures 1 and 2 show the multidimensional initial chaos theory treated by random method and chaos theory method, respectively, and the number of segments of chaos theory is 60. By comparison, it is found that the data link generated by the random method is messy and nondirectional in skipping training. The data link in skipping training processed by chaos theory is more concentrated and directional. According to Theorems 1 and 2 of the logistics algorithm, the data set constructed by this algorithm has nothing to do with spatial dimension and is suitable for multidimensional data processing. Moreover, every time the skipping training data are obtained, the distribution effect of the initial data link is the same, and the stability of the data set is high. To sum up, chaos theory is selected to process the initial data link.

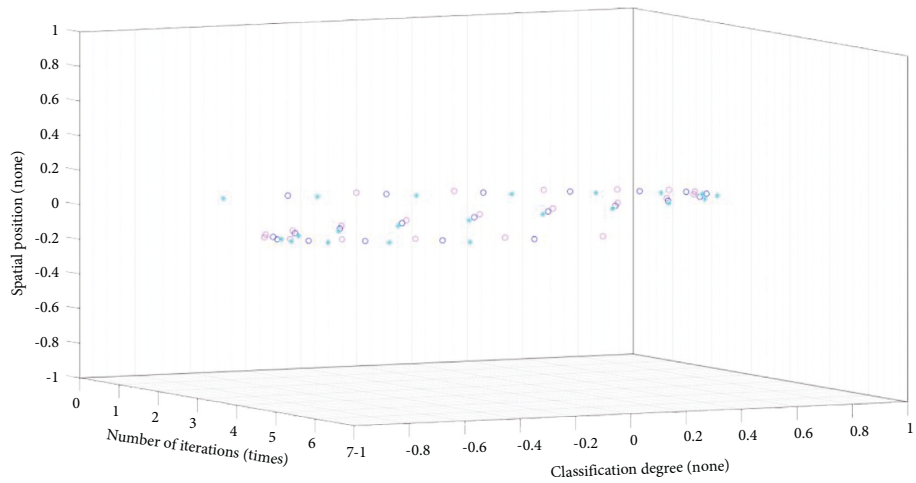


Figure 1: Random data constructed by chaos theory.

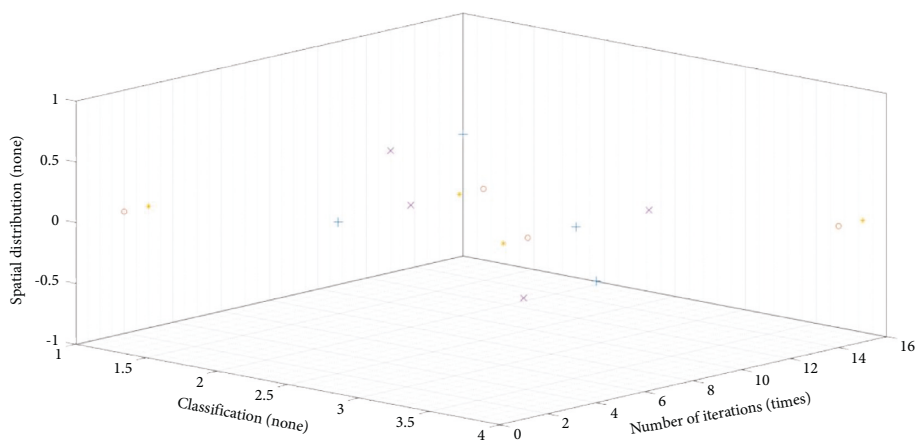


Figure 2: Data distribution constructed by chaotic logistics.

The synergy action between multiple strategies: the ability to evaluate the overall training effect is an important measure of the algorithm. In the initial stage of evaluation, the algorithm pays attention to the overall search and then progresses to the single-phase search [14]. In order to improve the evaluation efficiency of the algorithm, this paper introduces a multistrategy association method, which makes the stacks search according to different data links in skipping training and realizes overall and single-phase search. At present, the standard logistics algorithm is adopted, and there are other improved algorithms.

(1) Self-data model is shown in the following formula:

$$y_{ij}(T) = \omega \cdot y_{ij}(t) c_1 \cdot r_1 \cdot \frac{\left[\delta \cdot \sum_{i,j,k=1}^{n} g_{ij}^{k}\left\{x(t) \cdot f\left(P_j[x(t)]\right\}\right.\right.}{\max \sum_{i,j=1}^{n} g_{ij}^{k}[x(t)]}. \tag{6}$$

(2) Population data schema is shown in the following formula:

$$y_{ij}(T) = c_2 \cdot r_2 \frac{\left[\delta \cdot \sum_{i,j,k=1}^{n} g_{ij}^{k}\left\{x(t) \cdot f\left(P_j[x(t)]\right\}\right.\right.}{\max \sum_{i,j=1}^{n} g_{ij}^{k}[x(t)]}. \tag{7}$$

(3) Data judgment mode is shown in the following formula:

$$y_{ij}(T) = n \cdot \Delta \left| \frac{\left[\delta \cdot \sum_{i,j,k=1}^{n} g_{ij}^{k}\left\{x(t) \cdot f\left[P_j[x(t)]\right\}\right.\right.}{\max \sum_{i,j=1}^{n} g_{ij}^{k}[x(t)]} \right|. \tag{8}$$

(4) A broad data schema is shown in the following formula:

$$y_{ij}(T) = \omega \cdot y_{ij}(t) \frac{\left[\delta \cdot \sum_{i,j,k=1}^{n} g_{ij}^{k}\left\{x(t) \cdot f\left(P_j[x(t)]\right\}\right.\right.}{\max \sum_{i,j=1}^{n} g_{ij}^{k}[x(t)]}. \tag{9}$$

Among them, c_1 and c_2 are the correlation coefficient of different strategies, r_1 and r_2 are the correlation coefficient of different strategies, ω is the average weight of different

strategies, and n is the number of different strategies. In this paper, the standard logistics function is improved in two aspects. On the one hand, the search scope is expanded. For every iterative evaluation, one chaos theory form will be randomly selected from the above five forms for self-evaluation of chaos theory. Chaos theory evaluation can avoid single-phase optimal evaluation results and expand the search range. Random selection of evaluation forms can keep the diversity of analysis data and increase the possibility of obtaining single-phase optimal results. On the other hand, the convergence is improved. In order to balance the whole search ability and single-phase search ability of chaos theory, nonlinear adjustment coefficient α and linear weight ω are added to improve the evaluation speed, as shown in the following formula:

$$\alpha = \frac{\mathrm{Line}_t^d - 1}{\mathrm{Line}_T^D}, \tag{10}$$

where e is the evaluation parameter, t is the evaluation time, T is the maximum time, d is the iteration number, and D is the maximum iteration number. In the initial stage, α shows a decreasing trend, so the overall optimal result can be evaluated. In the middle and late stages of evaluation, the number gradually increases, and the single-phase optimal result evaluation can be carried out. The linear weight is evaluated as shown in the following formula:

$$w = \sum_{d=1}^{D}\left[\left(w_{\max} - \sum_{t=1}^{T}\Delta w_i\right)\right]^d, \tag{11}$$

where $w_{\max}$ and $w_{\min}$ are the maximum and minimum weights, and the results of d, t, D, and T are the same as above.

The strategy selection in the process of data analysis: the chaotic logistics algorithm adopts different operations and strategies for different massive training data and adjusts corresponding training parameters to realize distributed association of multiple massive training data, so as to complete the association evolution process. In the model, chaos theory data are divided into five dimensions, and each massive training data represent subspace [15]. In each iteration, different massive training data evolve simultaneously. After an iterative evaluation is completed, the adaptation results of different massive training data are compared, and the position of the overall optimal results of each massive training data is recorded. Then, each submassive training data is progressively trained to the overall optimal result, and the optimal position of submassive training data is obtained in the simplest way, so as to improve the speed and efficiency of search and evaluation.

3.2. Rope Skipping Training Judgment Method Based on Chaotic Logistics Algorithm. The basic idea based on chaos theory is the association evolution of multitime massive training data, and the initial results and threshold results are adjusted and optimized [16] to obtain the optimal solution and reduce the evaluation result rate in skipping training. The specific steps are shown in Figure 3.

Step 1. Determine the massive training data of chaos theory and the data link structure in skipping training, and determine the data structure of chaos theory according to the data characteristic demand. The initial weight results and threshold results of the whole data as a whole are mapped to chaos theory. The massive training data of each chaos theory are weight results and threshold results, which determine the massive training data of chaos theory in this paper.

Step 2. Data initialization randomly initializes the related training parameters of chaos theory. Set the number of chaos theory and the correlation constant between different massive training data and the maximum iteration times.

Step 3. Generate the moderation function. The chaotic theory is used to generate the data link in the initial skipping training of the chaotic theory, and it is mapped to the chaotic theory as the initial weight result and threshold result. Through the self-training and training of formulas (1) ~ (7), the correlation coefficient of chaos theory is continuously obtained. Evaluate the accuracy of each chaos theory, and take the absolute result of its square sum as a moderate function.

Step 4. Search the optimal position of chaos theory and the optimal position of each subgenetic stack as a whole. The initial chaos theory is randomly divided into five subdata links in skipping training, the fitness ratio results are obtained, and the optimal overall position and the optimal position of each subdata link in skipping training are recorded.

Step 5. Iteration of optimal position and speed. In the evolution of data links in 5 seed skipping training, randomly select one for evolution, such as formula (7) ~ (10), and add nonlinear adjustment coefficients and linear weights, such as formula (11) and formula (12).

Step 6. Data links are associated and evolved simultaneously in each subskipping training. After an iterative evaluation is completed, the best overall evaluation position is selected, and the best position is shared with other data links in skipping training. In other subskipping training, the data link gradually iterates to this position to get the best position result.

Step 7. Determine whether to stop the iteration.

4. Case Analysis of Skipping Rope Training

4.1. Judgment of Chaotic Logistics Model. Using Sphere, Rastring, Ackley, and other functions, the classical chaotic logistics method is tested to verify the performance of the model proposed in this paper.

Sphere is the sole minimum result function [17] for the test model as a whole, as shown in the following equation:

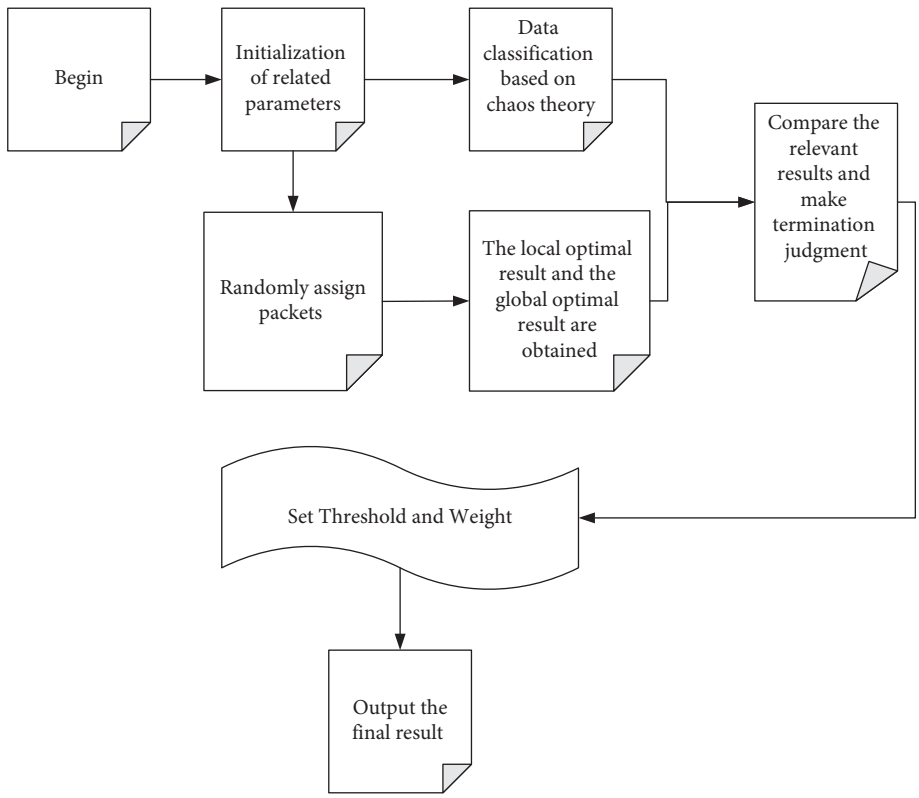


Figure 3: Evaluation process based on chaotic logistics algorithm.

$$F_1(x) = \sum_{i=1}^{n} \left[x_i^2\right]. \tag{12}$$

Rastring is based on the De Jong function and produces single-phase minimum results frequently through the cosine modulation transfer function to verify the practicability of the model solution [18], as shown in the following formula:

$$F_2(x) = \sum_{i=1}^{n} \left[x_i^2 - 10\cos(2\pi x_i) + 10\right]. \tag{13}$$

Ackley is a gradient optimization function for multidimensional points, testing the evaluation speed of multidimensional data to detect the overall convergence rate [19], as shown in the following equation:

$$F_3(x) = -20\exp\left\{-0.2\sum_{i=1}^{n}\sqrt{\frac{1}{n}\sum_{i=1}^{n}x_i^2}\right\} - \exp\left\{\frac{1}{n}\sum_{i=1}^{n}\cos(2\pi x_i)\right\} + 20 + e, \tag{14}$$

where n is the total number of indicators of evaluation data and x_i is any number of indicators. The result range of x_i is $\{-20, 20\}$ in Sphere, $\{-20, 20\}$ in Rastring, and $\{-20, 20\}$ in Ackley. For the convenience of evaluation, the number of chaos theory in this paper is $n = 50$, the maximum iteration number is $D = 20$, and the longest time is $T = 12$ min. The above three functions are tested, respectively. In order to improve the accuracy of data calculation, this paper uses the average value as the initial data, and the specific contents are shown in Table 2.

In order to further verify the results, different indexes of skipping training are compared, and the results are shown in Figures 4–6.

As can be seen from Table 1, compared with the logistics algorithm and chaos theory, the chaotic logistics algorithm proposed in this paper is closer to the overall optimal result. The chaotic logistics algorithm is superior to the other two algorithms in the aspects of standard deviation, average result, and range of taking result. Through the curve changes in Figures 4–6, we can see that the chaotic logistics algorithm has better stability and faster convergence speed. Therefore, the convergence speed, evaluation accuracy, and summation stability of chaotic logistics algorithm are better.

4.2. Preprocessing of Rope Skipping Training Data. The data set of skipping training includes physical fitness index, blood pressure, speed, endurance, and sensitivity [20]. After preliminary pretreatment, 982 rows of structured data and 12 rows of semistructured data were obtained, and other data were eliminated if they did not meet the requirements. In order to facilitate data analysis, the skipping training results are divided into five grades, which are accurate, normal, good, poor, and relatively poor. The processing results of data amount are shown in Table 3.

As can be seen from Table 2, the accuracy of all data results is better, all of which are more than 90%.

Table 2: Test results of different test functions.

Detection function	Algorithm	Physical strength	Endurance	Vital capacity	Overall training effect
Sphere	Chaotic logistics algorithm	91.84	94.90	93.88	98.2
	Logistics algorithm	98.98	99.90	99.88	
	Chaotic algorithm	94.90	98.98	96.94	
Rastring	Chaotic logistics algorithm	99.96	99.90	99.92	99.6
	Logistics algorithm	93.88	91.84	98.98	
	Chaotic algorithm	96.94	96.94	97.96	
Ackley	Chaotic logistics algorithm	99.90	99.88	99.86	95.3
	Logistics algorithm	91.84	94.90	94.90	
	Chaotic algorithm	98.98	98.98	96.94	

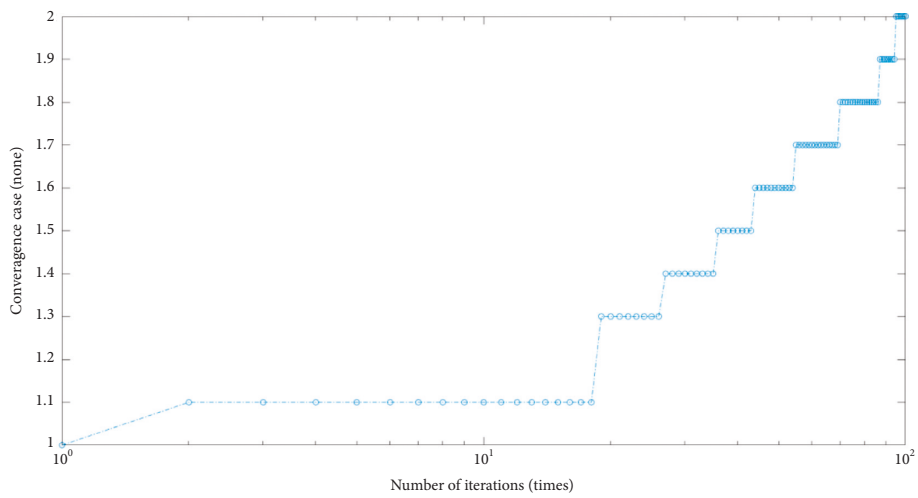


Figure 4: Strength test results.

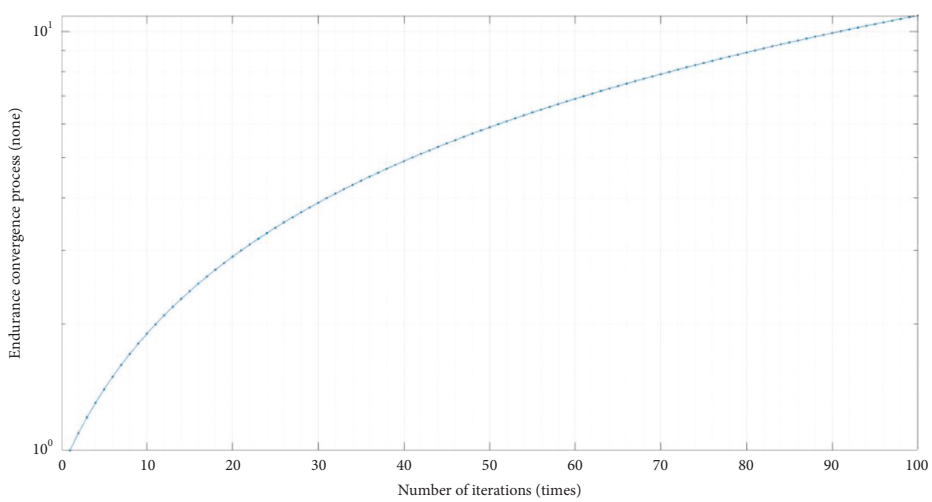


Figure 5: Endurance test results.

4.3. Test Results. In order to verify the chaotic logistics algorithm proposed in this paper, the results are compared with chaos theory and logistics algorithm, and the results are shown in Figure 7.

It can be seen from Figure 7 that the accuracy of the chaotic logistics algorithm is higher than that of the chaotic theory and logistics algorithm, but the error rate is lower, which shows that the evaluation of the chaotic theory and logistics algorithm is relatively stable, while the evaluation of chaotic theory and logistics algorithm is uneven. The average results of the above three algorithms are shown in Table 4.

It can be seen from Table 4 that the single chaos theory and chaos logistics method have the problems of insufficient accuracy and large variation of evaluation results in different

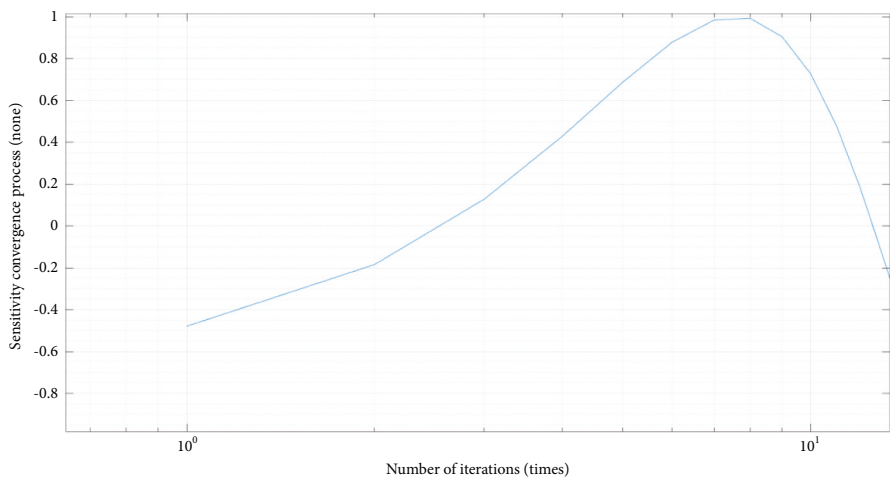


Figure 6: Sensitivity test results.

Table 3: Classification and quantity of skipping training.

Normal level	Amount of data	Accuracy
Accurate	123	93.91
Normal	543	91.21
Good	332	86.43
Poor	17	94.80
Comparatively poor	26	83.17

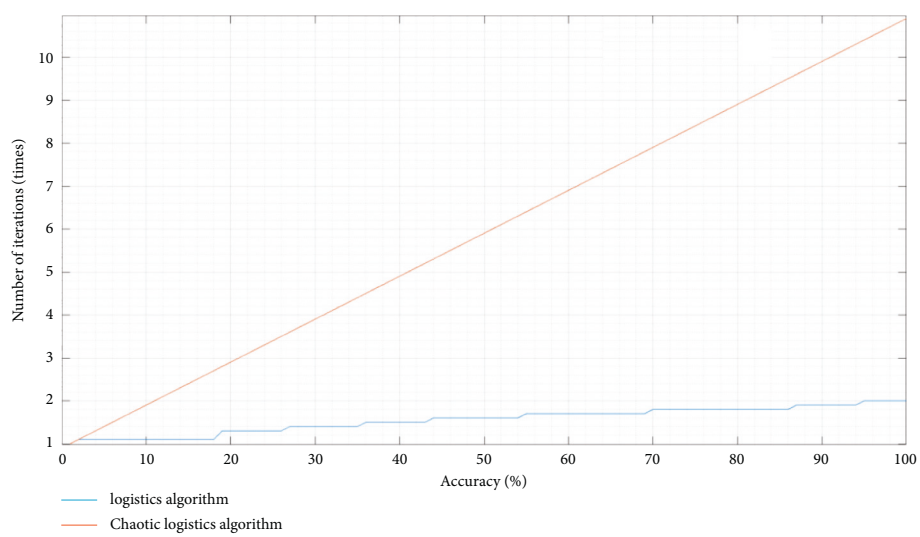


Figure 7: Test results of different algorithms.

Table 4: Comparison of prediction accuracy of different grades.

Algorithm	Accurate	Normal	Good	Poor	Comparatively poor
Logistics algorithm	77.23	69.22	77.28	78.34	75.94
Chaotic logistics algorithm	92.32	90.13	91.33	92.47	91.91
Chaotic algorithm	76.41	76.27	82.17	85.19	87.71

levels of skipping training prediction. Compared with chaos theory, the accuracy of the algorithm constructed in this paper is significantly improved. At the same time, the accuracy of the algorithm constructed in this paper is close to that of chaos theory, both of which are more than 80%, which is superior to chaos theory. In order to further verify

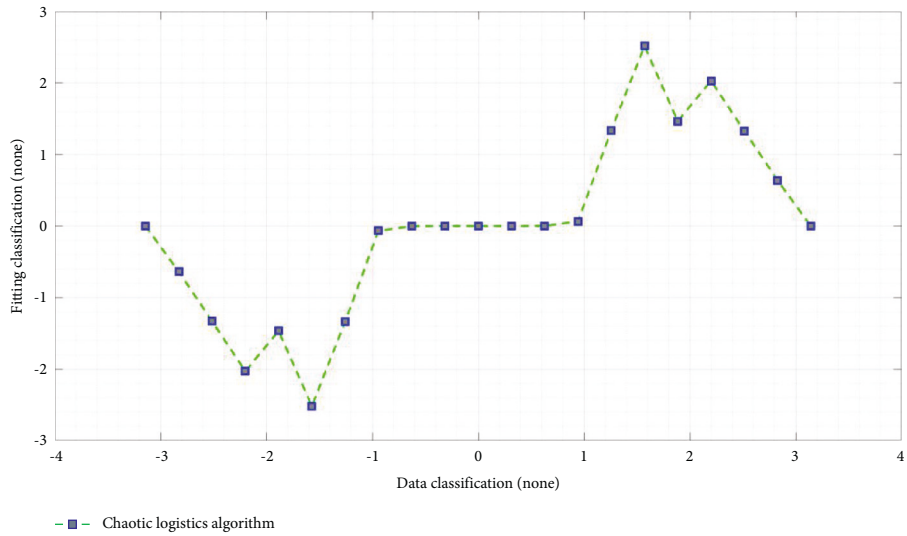


Figure 8: Fitting degree of chaotic logistics algorithm.

the superiority of the chaotic logistics algorithm, the optimal fitness results of different algorithms are compared, and the results are shown in Figure 8.

As can be seen from Figure 8, the optimization of fitness function results of chaotic logistics algorithm is more significant. The reason is that chaotic algorithm increases the adjustment, weight, and convergence factor of different massive training data. Consistent with relevant research results [21].

5. Conclusion

In this paper, chaos theory is proposed, and combined with the logistic evolution method, chaos theory is improved. At the same time, the threshold and weight of chaos theory are set, and the analysis model of skipping rope training is constructed. The results show that, compared with the classical chaos theory and chaos theory, the prediction accuracy and convergence of skipping rope training based on chaos theory and chaos theory are better, and it can be used to judge skipping rope training. However, in this model, the logistics strategy pays too much attention to training the overall effect evaluation ability, which leads to the relative decline of single-phase search ability and reduces the evaluation speed of the optimal solution. Therefore, in the future research, the logistic adjustment function will be added to improve the regression ability of the model.

Data Availability

The data supporting the conclusion of the article are shown in the relevant figures and tables in the article.

Conflicts of Interest

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

Acknowledgments

This work was supported by the Ordos Institute of Technology.

References

[1] Y. L. Yuan, X. K. Mu, X. Y. Shao, J. J. Ren, Y. Zhao, and Z. X. Wang, "Optimization of an auto drum fashioned brake using the elite opposition-based learning and chaotic k-best gravitational search strategy based grey wolf optimizer algorithm," *Applied Soft Computing*, vol. 123, Article ID 108947, 2022.

[2] L. H. Yu, L. Y. Xie, C. M. Liu, S. Yu, Y. X. Guo, and K. J. Yang, "Optimization of BP neural network model by chaotic krill herd algorithm," *Alexandria Engineering Journal*, vol. 61, no. 12, pp. 9769–9777, 2022.

[3] G. D. Ye, M. Liu, and M. F. Wu, "Double image encryption algorithm based on compressive sensing and elliptic curve," *Alexandria Engineering Journal*, vol. 61, no. 9, pp. 6785–6795, 2022.

[4] J. Y. Yao, Y. B. Sha, Y. L. Chen et al., "IHSSAO: an improved hybrid salp swarm algorithm and aquila optimizer for UAV path planning in complex terrain," *Applied Sciences*, vol. 12, no. 11, p. 5634, 2022.

[5] Y. J. Xian, X. Y. Wang, Y. Q. Zhang, X. P. Yan, and Z. Y. Leng, "A novel chaotic image encryption with FSV based global bit-level chaotic permutation," *Multimedia Tools and Applications*, vol. 23, no. 2, pp. 1–20, 2022.

[6] X. Y. Wang and Z. Y. Leng, "A dynamic image encryption algorithm based on improved ant colony walking path thought," *Sensing and Imaging*, vol. 23, no. 1, p. 17, 2022.

[7] P. F. Lv, X. S. Yu, J. N. Chi, and C. D. Wu, "Saliency detection via absorbing Markov chain with multi-level cues," *IEICE - Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E105.A, no. 6, pp. 1010–1014, 2022.

[8] L. Lian, "Reactive power optimization based on adaptive multi-objective optimization artificial immune algorithm,"

Ain Shams Engineering Journal, vol. 13, no. 5, Article ID 101677, 2022.

[9] Z. Li, X. R. Luo, M. J. Liu, X. Cao, S. H. Du, and H. X. Sun, "Wind power prediction based on EEMD-Tent-SSA-LS-SVM," *Energy Reports*, vol. 8, pp. 3234–3243, 2022.

[10] J. Li, T. S. Fu, C. F. Fu, and L. F. Han, "A novel image encryption algorithm based on voice key and chaotic map," *Applied Sciences*, vol. 12, no. 11, p. 5452, 2022.

[11] C. M. Li and X. Z. Yang, "An image encryption algorithm based on discrete fractional wavelet transform and quantum chaos," *Optik*, vol. 260, Article ID 169042, 2022.

[12] D. Kumar and M. Rani, "Alternated superior chaotic variants of gravitational search algorithm for optimization problems," *Chaos, Solitons & Fractals*, vol. 159, Article ID 112152, 2022.

[13] M. A. Khelifa, B. Lekouaghet, and A. Boukabou, "Symmetric chaotic gradient-based optimizer algorithm for efficient estimation of PV parameters," *Optik*, vol. 259, Article ID 168873, 2022.

[14] T. H. B. Huy, T. V. Tran, D. Ngoc, and H. T. T. Nguyen, "An improved metaheuristic method for simultaneous network reconfiguration and distributed generation allocation," *Alexandria Engineering Journal*, vol. 61, no. 10, pp. 8069–8088, 2022.

[15] X. L. Huang, Y. X. Dong, H. Y. Zhu, and G. D. Ye, "Visually asymmetric image encryption algorithm based on SHA-3 and compressive sensing by embedding encrypted image," *Alexandria Engineering Journal*, vol. 61, no. 10, pp. 7637–7647, 2022.

[16] B. H. He, B. Y. Jia, Y. H. Zhao, X. Wang, M. Wei, and R. Dietzel, "Estimate soil moisture of maize by combining support vector machine and chaotic whale optimization algorithm," *Agricultural Water Management*, vol. 267, Article ID 107618, 2022.

[17] M. Es-Sabry, N. El Akkad, M. Merras, A. Saaidi, and K. Satori, "A new color image encryption algorithm using multiple chaotic maps with the intersecting planes method," *Scientific African*, vol. 16, Article ID e01217, 2022.

[18] Y. Chen, S. C. Xie, and J. Z. Zhang, "A novel double image encryption algorithm based on coupled chaotic system," *Physica Scripta*, vol. 97, no. 6, Article ID 065207, 2022.

[19] A. A. Arab, M. J. B. Rostami, and B. Ghavami, "An image encryption algorithm using the combination of chaotic maps," *Optik*, vol. 261, Article ID 169122, 2022.

[20] M. R. Abuturab and A. Alfalou, "Multiple color image fusion, compression, and encryption using compressive sensing, chaotic-biometric keys, and optical fractional Fourier transform," *Optics & Laser Technology*, p. 151, 2022.

[21] H. G. Zhu, J. X. Ge, W. T. Qi, X. D. Zhang, and X. X. Lu, "Dynamic analysis and image encryption application of a sinusoidal-polynomial composite chaotic system," *Mathematics and Computers in Simulation*, vol. 198, pp. 188–210, 2022.

Hindawi
Mathematical Problems in Engineering
Volume 2022, Article ID 1852447, 15 pages
https://doi.org/10.1155/2022/1852447

Research Article

Classification of FinTech Patents by Machine Learning and Deep Learning Reveals Trends of FinTech Development in China

Hao Wang,[1] Xizhuo Chen,[2] Jiangze Du,[3] and Kin Keung Lai[4,5]

[1]*School of Public Finance and Public Administration, Jiangxi University of Finance and Economics, No. 169, East Shuanggang Road, Nanchang 330013, China*
[2]*School of Finance, Jiangxi University of Finance and Economics, No. 169, East Shuanggang Road, Nanchang 330013, China*
[3]*School of Finance, Jiangxi University of Finance and Economics, Research Centre of Financial Management and Risk Prevention, No. 169, East Shuanggang Road, Nanchang 330013, China*
[4]*International Business School, Shaanxi Normal University, No. 620 West Chang'an Street, Xi'an, China*
[5]*Department of Management Sciences, City University of Hong Kong, 83 Tat Chee Avenue, Kowloon Tong, Kowloon, Hong Kong SAR, China*

Correspondence should be addressed to Jiangze Du; jiangze.du@hotmail.com

Received 11 February 2022; Accepted 8 June 2022; Published 8 July 2022

Academic Editor: Long Wang

The development of the financial industry and its integration with information technology have promoted FinTech innovation. China is a major contributor to FinTech innovation, but few studies have systematically summarized FinTech innovation and development in China from the perspective of patents. This lacuna is attributable to the lack of a generally accepted classification of FinTech patents and the unavailability of classified Chinese FinTech patent text data. To fill this research gap, we developed a classification of Chinese FinTech patents and manually annotated a set of patent texts to train machine learning and deep learning models to classify massive Chinese patent application data and identify different types of FinTech innovations. Among the evaluated models, the character-level convolutional neural network (CNN) model and BERT model classified FinTech innovation most accurately. We used the character-level CNN to classify 20,529 Chinese FinTech patent applications from 2013 to 2020. The classified dataset was used to briefly analyze the history of FinTech innovation development in China and its future prospects.

1. Introduction

The essence of FinTech is to apply new technology to the financial field [1]. Cutting-edge technology can fundamentally change financial services by reducing transaction costs and improving transaction convenience and security [2]. The rapid development of FinTech is altering the mode of operation of the traditional financial system, and the integration of new technologies and financial systems has boosted global economic development. The continuous expansion of the scale of the FinTech market is the result of investments in research and development. Payment institutions are continuing to expand the application scope of FinTech by using new technologies such as big data, Internet of Things, and artificial intelligence, promoting financial industry reform and improving the operating efficiency of the financial system while controlling for risks [3]. Most notably, FinTech enables new types of lenders outside the traditional banking system to enter the financial market [4] and allows financial intermediary services to incorporate users' digital footprints into credit calculations and default rate prediction [5].

The launch of Alipay in 2004 is regarded as the beginning of this new financial industry in China. Alipay and WeChat Pay have since become important platforms for the huge flow of mobile payment funds in China. In the first half of 2020, mobile payments in China reached 196.98 trillion yuan, up 18.61% year on year, ranking first in the world. Despite the widespread use of FinTech in China's financial industry, the relevant literature has mainly focused on explaining the meaning of FinTech, the reasons for its emergence, and the impact of its level of development on regional economic development and corporate financing

constraints [6, 7]. A systematic and accurate overview of research on FinTech innovation in China is therefore lacking. The present study fills this research gap by identifying and classifying Chinese FinTech patents to analyze the technology layout and development status of FinTech in China. After reviewing previous research on the distribution of FinTech patents in the US and Europe [8–10], we screened patent application data from 2013 (the first year of Internet finance in China [11]) to 2020 using machine learning technology to classify the patent text. Patents related to FinTech were identified and classified according to different financial industry applications. A variety of machine learning algorithms and deep learning algorithms for identifying and classifying Chinese FinTech patents were tested, and the model with the best classification performance for patent classification combined with text filtering and manual identification was selected.

The rest of this study is organized as follows. In Section 2, we report related research progress. In Section 3, we introduce the databases used. The experimental procedure is explained briefly in Section 4, and the traditional machine learning models and deep learning models used in this study are described in Sections 5 and 6. We enumerate the classification results and evaluate model metrics in Section 7. In Section 8, we analyze the development process and orientation of FinTech innovation in China from the perspective of patents. Finally, we summarize the results of this study and outline future research directions in Section 9.

2. Related Work

With the boom of FinTech development, FinTech-related patent applications and licenses continue to grow and are becoming an important channel for enterprises to enter new markets. New technologies based on big data, cloud computing, and artificial intelligence are maturing, finding wider application in the financial industry and attracting attention from financial and academic circles. Journals such as *Review of Financial Studies* and *Journal of Management Information Systems* have featured FinTech innovation, and research perspectives include technologies applied in FinTech [12–14], opportunities and challenges faced by the traditional financial industry [15, 16], and corresponding risk controls and market regulations [17–19].

In the literature, innovation development is typically systematically summarized using annotated patent datasets and text-based machine learning approaches to classify patent databases [20–25]. Text classification approaches are mainly divided into traditional machine learning and deep learning approaches. The first step in text classification is to select the text classification category. Then, a small, representative segment of the unclassified data is manually labeled and divided into a training set, verification set, and test set. The training set and validation set are used to train and validate the model, and the test set is used to verify the accuracy of the model. The text classification algorithm treats a string of text as a sequence $W = [w_1, w_2, \ldots, w_T]$ composed of words or characters. In the data to be classified, each piece of data belongs to a certain category a_i, and set $A = \{a_1, a_2, \ldots, a_K\}$ contains all classification categories a_i. The model is calculated by the a posteriori probability of each category in a given text W:

$$P(a_i|W) = P(a_i|w_1, w_2, w_3 \ldots, w_T), \quad i = 1, 2, 3 \ldots k. \tag{1}$$

From this, the category of text M is calculated:

$$a = \arg\max_{a_i} P(a_i|w_1, w_2, \ldots, w_T). \tag{2}$$

Finally, the model with the best classification performance is chosen for unclassified data classification. Figure 1 shows commonly used text classification models.

In recent years, more and more researchers have tried to study FinTech innovation from the perspective of patents. Zhao et al. [26] directly take the number of patents applied by financial institutions as the proxy variable of FinTech innovation. Cojoianu et al. [27] use patents filed by FinTech startups to measure their innovation. However, these methods can only measure the FinTech innovation carried out by the financial sectors and cannot cover other sectors. Lee and Sohn [28] found that FinTech-related technological innovation could be found in patents, and the theme of FinTech innovation could be mined based on text analysis. However, due to the lack of annotated patents and corresponding datasets, there are few studies on FinTech patent identification based on machine learning or deep learning models.

Among available analyses, Chen et al. [8] manually annotated 1800 FinTech patents (in seven categories such as data analysis) from a US patent dataset from the BDSS database and used English text-based machine learning approaches for classification. The machine learning methods comprised a series of algorithms, including a linear support vector machine (SVM). The best performance on the test set was attained by the linear SVM, Gaussian SVM, and neural network model ensemble classifier based on patent text (accuracy on the test set: 82.6%; $F1$ score: 76.3%). In the subsequent classification of unlabeled data, 6,511 FinTech-related patents were extracted, of which 2,588 were from individuals and 3,923 were from companies. Descriptive statistics revealed that the largest number of FinTech patent applications in the United States was related to network security (1,179).

Xu et al. [9] used a random forest algorithm to identify and classify US FinTech patents from the Lens database covering the years 2014–2018. Compared with Chen et al.'s classification of FinTech patents, Xu et al. added lending as one of the categories and removed pure technology categories such as P2P. After labeling samples in seven categories, the best classification algorithm achieved an accuracy of 71.67% on a test set of 1800 FinTech patents. The largest category of licensed FinTech patents was related to mobile payments (682).

Caragea et al. [10] classified a large sample of 3850 manually annotated FinTech patents (in five categories: insurance, fraud, data analysis, investment, and payment) from patent databases under the jurisdiction of US and European laws using BERT, CNN, LSTM, and other methods. Compared with the seven categories used by Chen et al., Caragea et al.'s classification was more in line

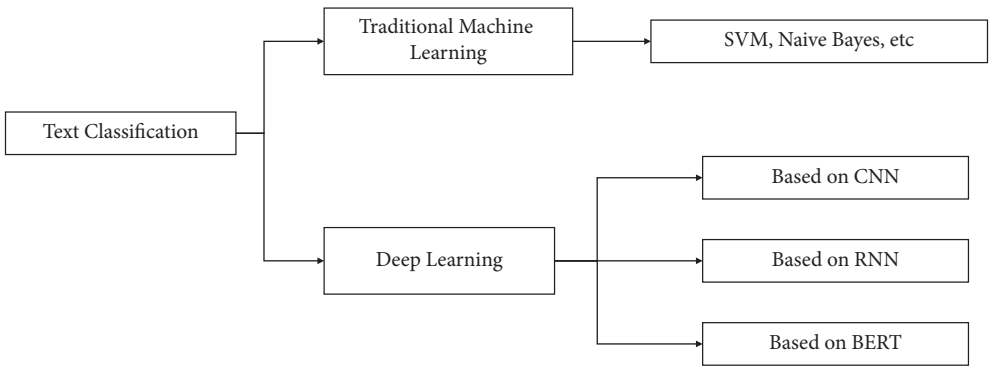


Figure 1: Commonly used text classification models.

with the definition of FinTech, as they abandoned the pure technology category and added insurance as a category to cover a large number of FinTech innovations in insurance and insurance risk prediction. Table 1 presents a comparison of the above three studies with our current work.

Chinese text classification has evolved from traditional machine learning algorithms to deep learning algorithms. Before classifying Chinese text, traditional machine learning and early deep learning algorithms first use classification modules such as Jieba to segment the text and remove stop words. Zhang et al. proposed character-level convolutional neural networks (ConvNets) for text classification. Comparisons of classification on a large number of datasets proved that the character-level CNN model performed better in the classification task than traditional word packets, n-grams, and TD-IDF variants. Based on this network structure, Google proposed the transformer structure and BERT model. The BERT model does not require word segmentation but operates directly on characters. For Chinese classification tasks, the word segmentation object in BERT is a single Chinese character.

3. Fintech Dataset

The patent application databases of various countries are not yet consolidated. Consequently, only the data for patents that are under the jurisdiction of the laws of a specific country and some patent data that are shared for the purpose of activity exchange can be retrieved from that country's official patent data. Previous studies of FinTech innovation have used patent data from the United States and Europe and omitted FinTech patents in China. Here, we used patent application data published by the China National Intellectual Property Administration from January 1, 2013, to December 31, 2020. Because FinTech development started relatively late in China compared with the United States, most patent applications have not yet been authorized, and patent application data are therefore used as the research object. Because an innovation subject applies for a patent once the innovation activities have achieved phased results, patent application data are more appropriate than patent licensing data for analyzing the boom in FinTech innovation in China.

The patent application dataset for China is very large, with more than 10 million patent applications in our selected time interval. To improve the classification efficiency, we referred to previous research on patent classification to preliminarily clean and filter the patent data. The official patent document classification and retrieval tool used by countries around the world is the International Patent Classification (IPC) compiled according to the Strasbourg Agreement on International Patent Classification signed in 1971. The major IPC categories are given in Table 2. The patent data of the China National Intellectual Property Administration also follow the IPC. In general, patents related to FinTech fall under categories *G* (Physics) and H (Electricity).

A preliminary screen of the 5,398,266 patent application data falling under categories G and H eliminated categories unrelated to finance or economics, such as nuclear physics. Then, we preliminarily used a machine learning algorithm to classify the patent data into finance-related and finance-independent patents after training the model on the general financial dataset. To ensure the integrity of the financial patents, we combined the patents that the different classification methods identified as potentially related to finance into a dataset containing 77,418 patent applications. This dataset is a collection of all patents roughly classified as related to FinTech. In the next steps, we used the manually annotated FinTech categories and different classification models to accurately classify the patent data.

4. Experimental Steps

FinTech encompasses the application of a variety of technologies in the financial industry. For example, information processing technology can accelerate the speed of information processing, enhance the ability of technology users to obtain information, reduce costs associated with the traditional reliance on human resources, and reduce redundant expenses. Although the technologies applied by different types of financial services overlap, their influences on China's financial development and corporate governance differ depending on the specific business context. In the earliest research on FinTech patents, Chen et al. divided FinTech patents into data analysis, Internet of Things, mobile payments, cybersecurity, blockchain, P2P, and robo-advising. Blockchain and Internet of Things do not fully

Table 1: Comparison of studies of the classification of FinTech patents.

	Dataset characteristics	Chen et al.	Xu et al.	Caragea et al.	Our work
1	Patent data source	BDSS	Lens	Orbis/Patsat	CNIPA
2	Year	2003–2017	2014–2018	2000–2017	2013–2020
3	Country	United States	America	America + Europe	China
4	Language	English	English	English	Chinese
5	Optimal classification algorithm	Ensemble classifier	Random forest	BERT	Character-level CNN
6	Number of patents manually labeled	1800	1800	3850	5400
7	Number of FinTech patents	6511	3602	25580	20529

Table 2: Major IPC categories.

Patent class	Meaning
A	Human necessities
B	Performing operations; transporting
C	Chemistry; metallurgy
D	Textiles; paper
E	Fixed constructions
F	Mechanical engineering; lighting; heating; weapons; blasting
G	Physics
H	Electricity

align with the general definition of FinTech, as these technologies are generally applied to specific financial industries such as insurance, lending, and data processing and thus have high overlap with other FinTech classification labels. Although there is no unified classification of FinTech patents, ongoing updates are gradually converging with the definition of FinTech. In this study, we referred to previous FinTech research [1–19] and research on the development of FinTech in China [29–31] that has analyzed and summarized the main characteristics of FinTech to divide the patent data into seven categories: FinTech-unrelated, payment, loan, insurance, security, data analysis, and investment. Table 3 provides the specific classifications.

Based on the categories in Table 3, representative patent texts were manually selected and labeled. The final labeled sample contained 5200 patents, including 2200 patents unrelated to FinTech and 3,000 FinTech-related patents (500 per category).

For a more intuitive presentation of the contents of FinTech patents, we randomly select one patent from each category as an example and report in Table 4. Meanwhile, considering that the patent abstract is too long to be a whole display, we extracted the contents that can reflect the characteristics and uses from the original abstract of patents and report in Introduction of Table 4.

The process of classification is illustrated schematically in Figure 2. The labeled sample was divided 8 : 1 : 1 into the training set, validation set, and test set, and traditional machine learning models (SVM, KNN, decision tree, and naive Bayes) and deep learning models (CNN, RNN, LSTM, and BERT) were used for text classification. To evaluate the performance of the different models in classifying Chinese patents, we used standard metrics such as precision (Pr), recall (Re), $F1$ score ($F1$), and accuracy (Acc):

$$\begin{aligned} \text{Precision} &= \frac{\text{TP}}{\text{TP} + \text{FP}}, \\ \text{Recall} &= \frac{\text{TP}}{\text{TP} + \text{FN}}, \\ F1 &= \frac{2 \cdot \text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}, \\ \text{Acc} &= \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}. \end{aligned} \tag{3}$$

The above metrics take values between 0 and 1, where values closer to 1 indicate higher classification accuracy and a superior model. The parameters in the above formulas include true positive cases (TP), false positive cases (FP), false negative cases (FN), and true negative cases (TN). A TP occurs when the model correctly predicts a positive category sample, whereas a TN means that the model correctly predicts a negative category sample. An FP occurs when the model incorrectly predicts that a negative category sample is positive, while an FN corresponds to the incorrect prediction of a positive category sample as a negative category sample.

5. Traditional Machine Learning Models

This section describes the traditional machine learning models evaluated for patent classification. All models used are supervised machine learning models. The typical steps for text-processing are segmenting words, removing pause words and low-frequency words, performing feature selection, converting string text to text represented by vectors, and finally classifying the text. In English text, the spaces between words act as natural boundaries, and word segmentation is not necessary. However, Chinese text must be segmented. Automatic segmentation of a sentence into reasonable words can be performed by a computer following semantic logic. In natural language processing, the word is the smallest unit, and the accuracy of word segmentation directly affects the results of text classification.

Feature selection is required for Chinese word segmentation. If all feature words in the text are used to represent the text, the dimension of the feature space will typically be greater than 100,000. Such a high-dimensional space will greatly reduce the calculation efficiency or even make completion of the calculation impossible. Words must be selected from the text to form a new feature space and achieve dimensionality reduction; words with very weak

Table 3: FinTech categories in China.

FinTech category	Interpretation
Data analytics	The use of big data, cloud computing, and other pieces of information technology to clean, filter, and analyze large amounts of data related to the financial industry collected through various channels such as mobile banking
Lending	The use of information technology to deal with deposit and loan business
Insurance	The application of information processing technology for the intelligent recommendation of insurance products, policyholder risk identification, and other insurance businesses
Payment	The use of the Internet of Things and other technologies to process mobile payment business, provide payment function interfaces, and make payments intelligent, convenient, and cashless
Security	The application of fingerprint verification, iris verification, and face recognition technology in financial industry and equipment security management to improve the security of financial activities
Investment	Quantitative investment by enterprises and intelligent portfolio recommendation based on the analysis of investors' risk preferences

Table 4: Examples of FinTech patents.

Category	Patent	Applicant	Introduction
Data analytics	Abnormal warning methods, devices, systems, equipment, and media of big data products	Shenzhen Qianhai Micro Public Bank Co., Ltd.	The invention relates to the field of big data technology of FinTech and converts the error information of big data products into error codes and error descriptions for the convenience of operation and maintenance personnel to view and improve efficiency
Investment	Generating method and device of financial market product trading report	Bank of China Limited	The invention discloses a method and device for generating a financial market product trading report, in which the transaction information of all target customers is collected within a specified period
Security	Method and device of bank user authentication based on behavior characteristics	Industrial and Commercial Bank of China Limited	The invention proposes a bank user authentication method based on behavior characteristics, which makes full use of the user identity to identify the behavior characteristics of all kinds of business in the bank and labels the behavior characteristics
Insurance	Method, device, and storage medium for calculating insurance probability	Ping An Technology (Shenzhen) Co., Ltd.	The invention calculates the common characteristics of the customers who buy insurance through the GBDT model and then injects the customer data of the target customers into the model to obtain the insurance probability of the target customers
Lending	The invention relates to a loan system data processing method	China Construction Bank Corporation	The invention discloses a loan system data processing method, device, device, and storage medium. By dividing the prebatch and the primary batch, batch errors can be exposed in advance during prebatch, preventing dirty data from being directly stored in the database
Payment	Payment methods, devices, and systems	Tencent Technology (Shenzhen) Co., Ltd.	The invention discloses a payment method, equipment, and system. It solves the problem of complicated operation of online payment process when the third-party merchants access the shopping platform application in the form of web page

contributions to the text can be omitted. For example, the adverb "de" appears in almost all texts and is a frequent and meaningless feature that is usually deleted as a stop word.

Humans understand text in the form of character coding, whereas a computer system requires binary coding. Consequently, Chinese text must be transformed into binary code to allow the computer to calculate the text information. One commonly used text representation model is the vector space model. However, the weight of many feature words in vector space is zero, leading to a less-than-ideal classification effect. For traditional machine learning classification, we use TD-IDF text representation. TF-IDF is a statistical method that evaluates the importance of a single word in a dataset or a corpus. The importance of a word increases with the frequency of its appearance in a document but decreases with the frequency of its appearance in the corpus.

In this study, we evaluated the following traditional machine learning models: SVM, k-nearest neighbor (KNN), naive Bayes, and decision tree.

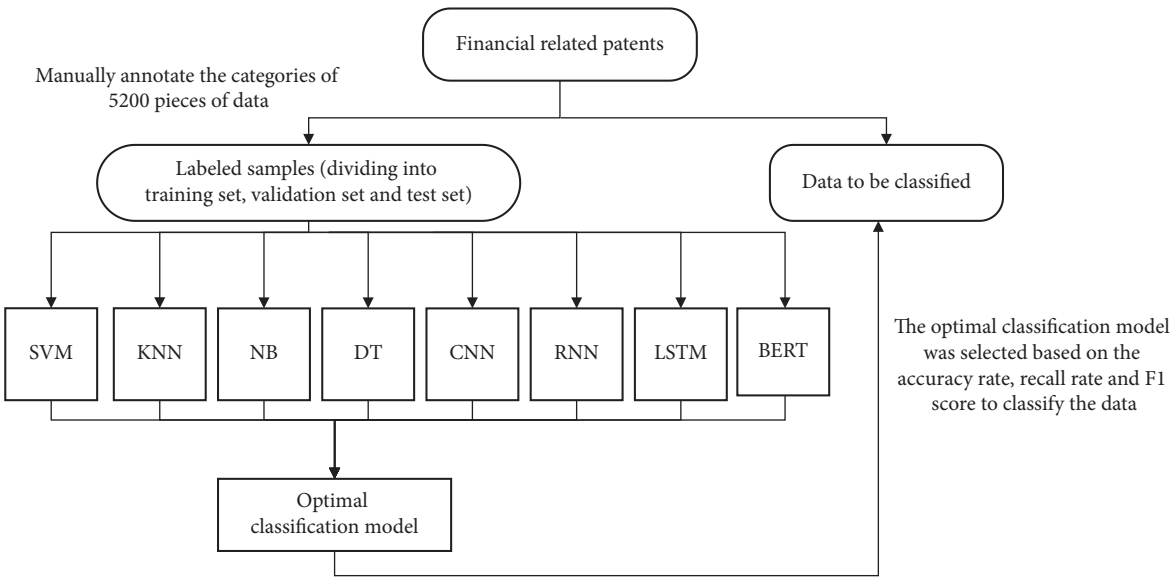


Figure 2: Patent data classification steps.

5.1. SVM. SVM is a kind of general feedforward neural network. The basic model finds the best separation hyperplane in the feature space that gives the largest interval between positive and negative samples in the sample training set. Depending on the sample's characteristics, the sample can be classified as linearly separable or linearly inseparable, as shown in Figure 3.

Linearly separable means that two kinds of samples can be separated by a straight line in a two-dimensional space. For a training sample $\{(x_i, d_i)\}$, where x_i is the ith sample of the input pattern and d_i is the corresponding expected response (target output), it is first assumed that the class represented by the subset $d_i = +1$ and the pattern represented by $d_i = -1$ are linearly separable. The hyperplane decision surface equation for separation is as follows:

$$f(x) = \omega^T x + b, \tag{4}$$

where x is the input vector, ω is the adjustable weight vector, and b is the offset. For a given weight vector ω and offset b, the goal of SVM is to find the hyperplane with the largest separation edge. Under this condition, the decision surface is called the optimal hyperplane.

A high-dimensional space can be separated by a high-dimensional function. Linearly inseparable means that the sample features are mapped to high-dimensional space by a Gaussian kernel function. Nonlinear features are transformed into linear separable features so that the sample can be processed by a linear separable method.

SVM can be used in a variety of supervised learning algorithms, including classification, regression, and anomaly detection, and has many advantages among traditional machine learning algorithms, such as high efficiency in high-dimensional space, a data dimension that is larger than the number of samples, and the ability to use a subset of the training set in the decision function (commonly known as a support vector), which can make efficient use of computer memory. However, SVM also has many disadvantages when it is used for a classification task. For example, if the number of features is much larger than the number of samples, overfitting can occur easily when selecting the kernel function for training. Overcoming this limitation usually requires regularization, among other means of dealing with overfitting.

5.2. KNN. KNN was first proposed by Cover and Hart in 1968. It is mature in both theory and research and is one of the simplest machine learning algorithms. The operation idea of this method is as follows: if most of the k-nearest samples in the feature space belong to a certain category, then the sample should also belong to this category. KNN determines the category of the samples to be classified based only on the category of the nearest one or several samples.

The drawback of this method is that the computational burden is very large, as the distance of each text to be classified from all known samples must be calculated to obtain the text's k-nearest neighbor points. The most common solution is to manually process the known sample points and remove the samples with little effect on classification in advance. The reverse KNN method can reduce the computational complexity of the KNN algorithm and improve the efficiency of classification. The KNN algorithm is more suitable for text classification of large sample sizes, as classification errors can occur easily when the sample size is small.

5.3. Naive Bayes. Bayesian methods use theories of probability and statistics to classify a sample dataset. Because the algorithm has a solid mathematical foundation, all Bayesian algorithms used for classification have higher accuracy and a lower error rate when the dataset is large. Bayesian methods combine prior probability and posterior probability to avoid the subjective bias of only using prior probability. Naive Bayes methods are a group of supervised learning algorithms

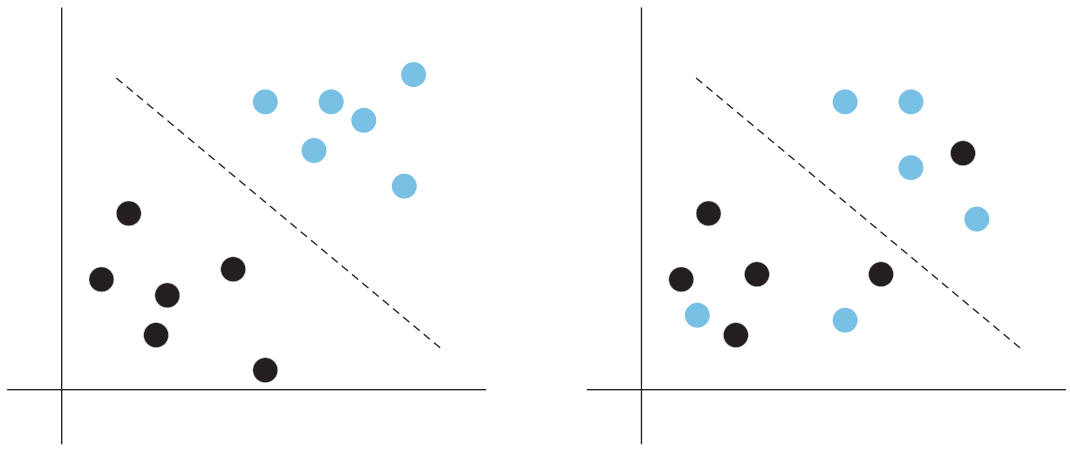


Figure 3: Linearly separable and linearly inseparable samples.

based on the Bayes theorem, which simply assumes that each pair of features is independent of each other. For a class y and x_1 to x_2 related eigenvectors of N, the naive assumption that each pair of features is independent of each other is used:

$$P(x_1|y, x_1, \ldots x_{i-1}, x_{i+1}, \ldots x_n) = P(x_i|y). \tag{5}$$

The maximum posterior probability can be used to estimate $P(x_i|y)$, which is the relative frequency of class y in the training set. The various naive Bayes classifiers differ mainly in the assumptions made when dealing with $P(x_i|y)$ distributions. Although the assumptions of the naive Bayes model are relatively simple, it works quite well in actual classification tasks, and only a random training sample is needed to estimate the required parameters. Naive Bayes learners and classifiers are much faster than other more complex methods. The decoupling of the conditional distribution of classification means that each feature can be estimated independently as a one-dimensional distribution. This in turn helps to mitigate the problems caused by dimensional disasters.

The Bayesian algorithm used in this study is a multinomial distributed naive Bayesian model, which is commonly used for text classification tasks and has shown excellent performance in previous experimental studies. The distribution parameter is determined by the $\theta_y = (\theta_{y1}, \ldots, \theta_{yn})$ vector of each class y, where n is the number of features (for text classification, the number of features is the size of the vocabulary) and θ_{yi} is the probability $P(x_i|y)$ of feature i in the sample belonging to class y. θ_y is estimated using the smoothed maximum likelihood estimation method to calculate the relative frequency:

$$\hat{\theta}_{yi} = \frac{N_{yi} + \alpha}{N_y + \alpha n}. \tag{6}$$

In the above formula, $N_{yi} = \sum_{x \in T} x_i$ is the number of times that feature i in training set T appears in class y, and $N_y = \sum_{i=1}^{n} N_{yi}$ is the sum of all features appearing in class y. The prior smoothing factor $\alpha \geq 0$ is applied to features that do not appear in the learning sample to prevent the occurrence of a model calculation result with probability 0. $\alpha = 1$ is called Laplace smoothing, while $\alpha < 1$ is called Lidstone smoothing.

5.4. Decision Tree. In traditional machine learning, a decision tree is a prediction model that represents a mapping relationship between object attributes and object values. Each node in the decision tree represents a particular object, each bifurcation path represents a possible value, and each leaf is the value of an object represented by the path taken from the root node to that leaf. A decision tree has a single output. To output complex values, multiple independent decision trees can be built to output different values. The decision tree technique is frequently used in data mining and can be used for data classification, prediction, and regression. For the task of classification, the decision tree model is constructed according to the given dataset in the training stage, and the most valuable feature segmentation node is selected from the root node. The test phase and prediction phase are based on the decision tree model constructed by training.

6. Deep Learning Model

Four deep learning models are evaluated in this study for patent classification: CNN, RNN, LSTM, and BERT.

6.1. CNN. The convolutional neural network (CNN) is a representative deep learning algorithm and a type of feedforward neural network that comprises convolutional computation and a deep structure. Because CNNs feature convolution kernel parameters within the neural network hidden layer and interlayer connection sharing of sparsity, learning pixel points and audio requires less computation, producing stable effects and data characteristics without additional requirements. Consequently, CNNs are widely used in image and text classification.

The CNN model used in this study is based on the implementation of TensorFlow on the Chinese dataset and classifies patent text based on character-level CNN. The

character-level CNN model encodes Chinese sentences into character sequences as input. Encoding is performed by specifying a Chinese word list of size M for input and quantifying each Chinese character using 1-of-M encoding ("one-hot" encoding). The sequence of characters is then converted into a sequence of m vectors of fixed length L. If the length of the paragraph is less than L, 0 is added after the sequence; if the paragraph length is longer than L, the field length greater than L is cutoff. Any character that is not in the word list (including whitespace characters) is encoded as a full zero vector. The character quantization order is reversed so that the most recent read of the character is always near the start of the output, which allows the full connection layer to associate weights with what the model reads.

The main operations of the model are shown in Figure 4. The first layer embeds text into a low-dimensional vector. The core of the model is the next layer, the CNN layer, which is the convolutional computing layer for calculating one-dimensional convolution. For example, assuming a discrete input function $g(x) \in [1, l] \longrightarrow \mathbb{R}$ and a discrete kernel function $f(x) \in [1, k] \longrightarrow \mathbb{R}$, the convolution $h(y)$ between f and g with stride d and migration constant c can be defined as

$$h(y) = \sum_{x=1}^{k} f(x) \cdot g(y \cdot d - x + c). \tag{7}$$

Next, we pool the maximum result calculated by the convolutional layer into a long feature vector and add dropout regularization. The convolution is again calculated given a discrete input function $g(x) \in [1, l] \longrightarrow \mathbb{R}$, stride d, and migration constant c. The max-pooling function $h(y)$ of $g(x)$ can be defined as

$$h(y) = \max_{x=1}^{k} g(y \cdot d - x + c). \tag{8}$$

The softmax layer is then used to classify the result.

6.2. RNN and LSTM. CNNs are unable to process sequence data due to the loss of sequence features during the pooling operation in the pooling layer. RNNs were introduced to alleviate this shortcoming. Text classification based on RNNs has developed rapidly in response to the need to process sequence data. As shown in Figure 5, the input vector of the neural network A containing several layers is x_t, and the output vector is h_t. The output calculated in the previous step is used as the input in the next step to form a chain structure of A through cyclic calculation.

However, RNNs inherit the flaw of deep architecture: the deeper the network, the more obvious the gradient explosion and gradient vanishing. Therefore, improvements of the structure of the RNN model have been proposed, such as LSTM (long short-term memory). As shown in Figure 6, in addition to the internal hidden state, LSTM adds a cell state. The information in the cell state is controlled by three hidden gates: the input gate, forget gate, and output gate. The forget gate determines what information to discard from the cell state, and the input gate determines how much new information to add to the cell state. This calculation directly affects the output cell state. The output gate determines the output value based on the cell state. These gates are used to determine the information in the previous cell that needs to be forgotten, update information, or output information for the current cell state.

6.3. BERT. BERT is a language model proposed based on Transformer structure that encodes tags and sentences into dense vector representations.

First, the BERT model needs to be pretrained on a large number of unlabeled language texts; the text marker sequence is used as the input. The first marker is a special marker represented by [CLS] (the output of (CLS) can be regarded as the semantic output of the whole text sequence), and the input sequence of BERT is composed of one or two sentences. Separated by another special tag [SEP], Transformer converts the embedding of the original input to the embedding of the context output. Figure 7 shows the operation structure of the BERT model when two sentences are entered. After preprocessing and fine-tuning the resulting model, the text-processing task is further pretrained to match the downstream task.

7. Classification Results

In this section, we present the classification performance of the different models and the classification results obtained using the optimal model. Table 5 provides the classification metrics for the different categories. The classification performances of character-level CNN and BERT were significantly better than those of the other classification models. As shown in Figure 8, the CNN had the best classification performance in terms of average score and Acc, and thus we chose CNN as the optimal model to analyze the data of 77,418 unclassified patents.

Applying the character-level CNN to patent text classification yielded a total of 20,529 FinTech patents. The steps of patent are given in Table 6.

After classifying 77,418 potentially FinTech-related patents and removing those belonging to the "FinTech-unrelated" category, we obtained a dataset containing 25,153 FinTech-related patents (belonging to six FinTech-related categories, such as payments). Removing FinTech patent applications filed abroad left a total of 20,529 FinTech patents. After excluding patents filed by individuals, universities, and so on, the total number of FinTech patents filed by companies (including banks) was 18,562. A total of 5,450 companies (including unlisted companies) applied for FinTech patents. A total of 81 banks (including branches) applied for 2009 FinTech patents.

8. Analysis of FinTech Patent Applications

In this section, we use the classified FinTech patents to briefly analyze the development of FinTech innovation in China. The dataset contains a total of 20,529 patents covering all FinTech patents filed in China between 2013 and 2020 based on classification by the deep learning model.

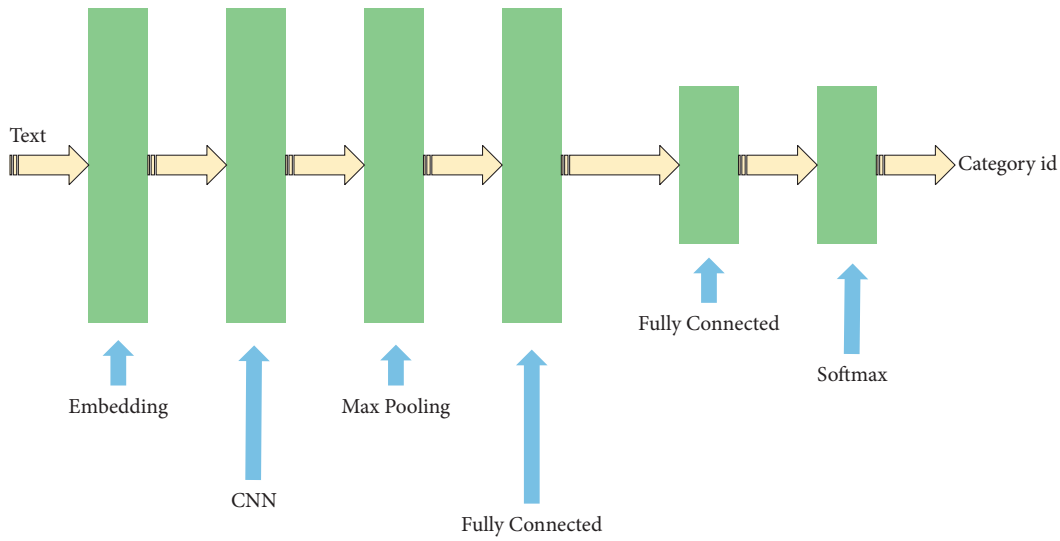


Figure 4: Convolutional neural networks for text classification.

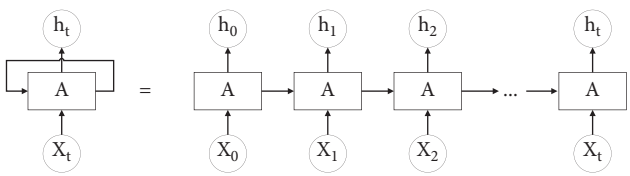


Figure 5: Recurrent neural network (RNN).

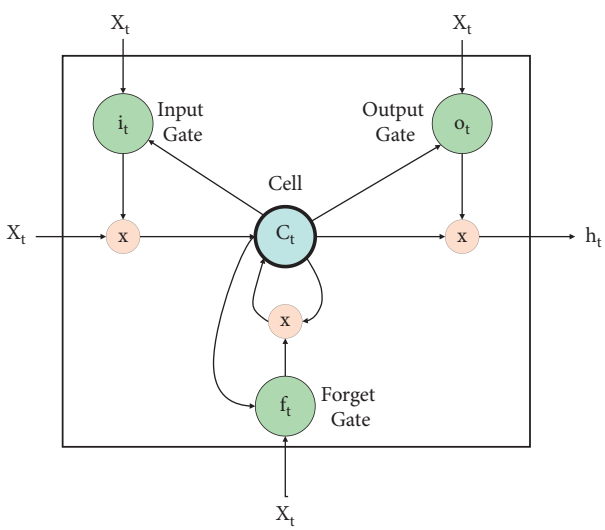


Figure 6: Long short-term memory (LSTM).

8.1. Number of New FinTech Patents. As shown in Figure 9, the number of FinTech patent applications in China was 862 in 2013 and increased each year from 2013 to 2018. Between 2018 and 2020, the number of patents remained high and stable, but the shares of different patent types changed significantly, reflecting the vigorous development of FinTech innovation. The growth in the number of FinTech patent applications in the payment category peaked (1029) in 2018 and then slowed. This shift may reflect the increasing maturity of mobile payment technology and, as a result, limited room for innovation. At present, the main innovations in payment FinTech are improving existing mobile payment systems and making mobile payments more convenient and faster by using Internet of Things technology. The number of data analytics FinTech patent applications rose each year, heralding the era of big data and cloud computing technology and financial industry integration. The future directions of development of financial science and technology will be efficiently handling large databases, upgrading financial service models, and promoting the development of the financial industry.

Figure 10 shows regional differences in the number of FinTech patent applications in China. FinTech-related patent applications from Guangdong and Beijing accounted for 52.47% of all FinTech patent applications in China. Overall, the number of FinTech patent applications was higher in economically developed regions than in less-developed regions. Provinces with more than 1000 applications included Guangdong, Beijing, Shanghai, Jiangsu, and Jiangsu. Figure 11 shows the proportions of the annual number of FinTech patent applications in these five provinces compared with all other provinces.

Comparing the proportions of FinTech patent applications in various provinces from 2013 to 2020 provides insights into the development process of FinTech in China. FinTech innovation gradually concentrated in the developed provinces. The proportions of FinTech patent applications from Beijing, Shanghai, Guangzhou, Jiangsu, and Zhejiang increased over the analyzed period, and in 2020, Beijing surpassed Guangzhou to lead China in the number of FinTech patent applications. Along with this growth, the proportion of payment patents decreased, whereas the proportion of data analytics patents increased, as shown in Figure 11. Accordingly, we can speculate that the change in patent application structure at the provincial level reflects the shift in the focus of FinTech development to the analysis and prediction of user behavior from big data. Beijing and Guangdong have advantages over other

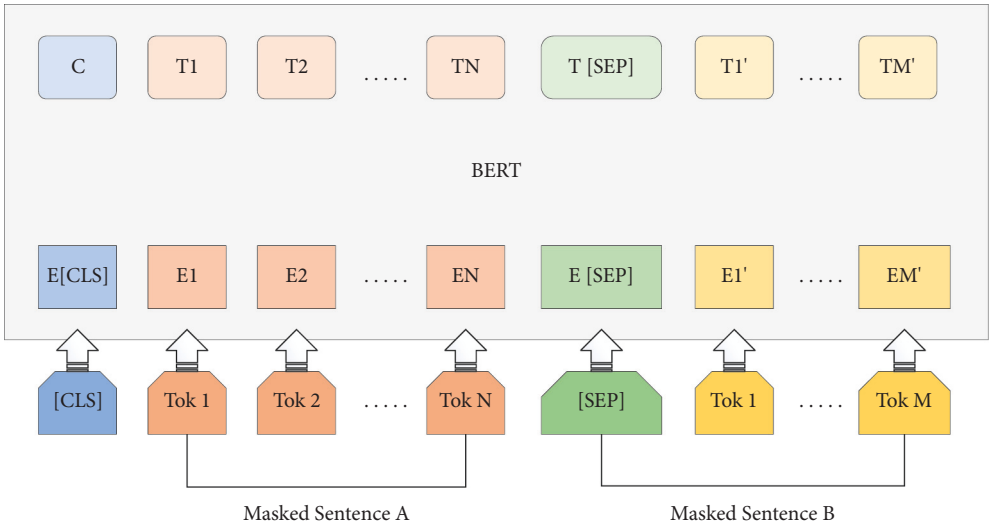


Figure 7: Pretraining BERT model structure.

Table 5: The classification performance of the different models on the test set.

Category	Metric	SVM	KNN	NB	DT	CNN	RNN	LSTM	BERT
FinTech-unrelated	Pr	0.74	0.54	0.47	0.62	0.81	0.47	0.44	0.84*
	Re	0.67	0.76	0.34	0.6	0.87	0.32	0.56	0.9*
	F1	0.71	0.63	0.40	0.61	0.84	0.38	0.50	0.87*
Insurance	Pr	0.58	0.45	0.46	0.34	0.82*	0.42	0.37	0.81
	Re	0.30	0.56	0.22	0.32	0.90	0.30	0.34	0.95*
	F1	0.39	0.50	0.30	0.33	0.86	0.35	0.35	0.87*
Data analytics	Pr	0.57	0.63	0.67	0.63	0.78*	0.67	0.65	0.74
	Re	0.73	0.76*	0.54	0.71	0.58	0.54	0.59	0.55
	F1	0.65	0.69*	0.59	0.67	0.67	0.59	0.62	0.63
Lending	Pr	0.92*	0.62	0.85	0.80	0.84	0.62	0.87	0.86
	Re	0.68	0.72	0.44	0.74	0.98*	0.42	0.66	0.96
	F1	0.78	0.67	0.58	0.77	0.91*	0.50	0.75	0.91*
Payment	Pr	0.76	0.62	0.40	0.59	0.90*	0.62	0.66	0.90*
	Re	0.74	0.78	0.24	0.70	0.86*	0.42	0.61	0.78
	F1	0.75	0.69	0.30	0.64	0.88*	0.50	0.63	0.83
Investment	Pr	0.90*	0.84	0.78	0.82	0.82	0.88	0.76	0.77
	Re	0.72	0.64	0.64	0.74	0.82*	0.60	0.74	0.82*
	F1	0.80	0.73	0.70	0.78	0.82*	0.71	0.75	0.8
Security	Pr	0.67	0.46	0.42	0.45	0.89*	0.44	0.62	0.86
	Re	0.67	0.59	0.33	0.47	0.80*	0.28	0.49	0.66
	F1	0.67	0.52	0.37	0.46	0.85*	0.35	0.55	0.75
Average	Pr	0.72	0.64	0.55	0.63	0.84*	0.59	0.65	0.83
	Re	0.72	0.61	0.54	0.63	0.83*	0.57	0.64	0.80
	F1	0.71	0.61	0.52	0.63	0.83*	0.55	0.64	0.81
	Acc	0.72	0.61	0.53	0.63	0.84*	0.56	0.64	0.83

The significance of the symbol "*" means that the model performs significantly better than others in the patent classification.

provinces in both development environment and technology, and the development of big data and cloud computing technology is relatively mature. Consequently, it is easy for these provinces to master this technology and apply it to the financial industry to lead the development of FinTech.

Figure 12 presents the numbers of FinTech patent applications by financial and nonfinancial sectors. In the financial sectors, banks are the main drivers of FinTech innovation due to the size and complexity of their financial operations. In 2019 and 2020, the number of FinTech patent applications filed by banks, which had few applications compared with nonfinancial sectors before 2019, began to rise rapidly. Prior to 2019, most FinTech patent applications came from Bank of China Limited and Shenzhen Qianhai WeBank Limited. Although an important part of the financial industry, banks were initially not sensitive to the development of financial technology and had a low level of

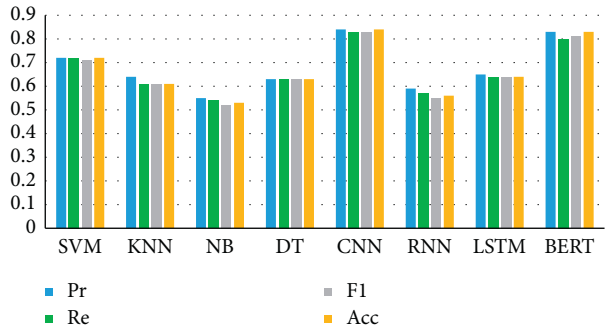


Figure 8: Average metrics of the different models.

Table 6: Complete FinTech patent classification process.

Filtering and sorting steps	Number of remaining samples	Number of remaining samples
All patents under entries G and H filed in China from 2013 to 2020 in the database		5398266
Remove pure technical patent data from IPC classification	4072571	1325745
Filter nonfinancial patents via machine learning	1248327	77418
Classify text using CNN to remove FinTech-unrelated patents	52265	25153
Remove patents filed abroad	4624	20529
Number of remaining FinTech patent applications	20529	
FinTech patents		
Individual patent applications	1967	
Nonindividual patent applications	18562	

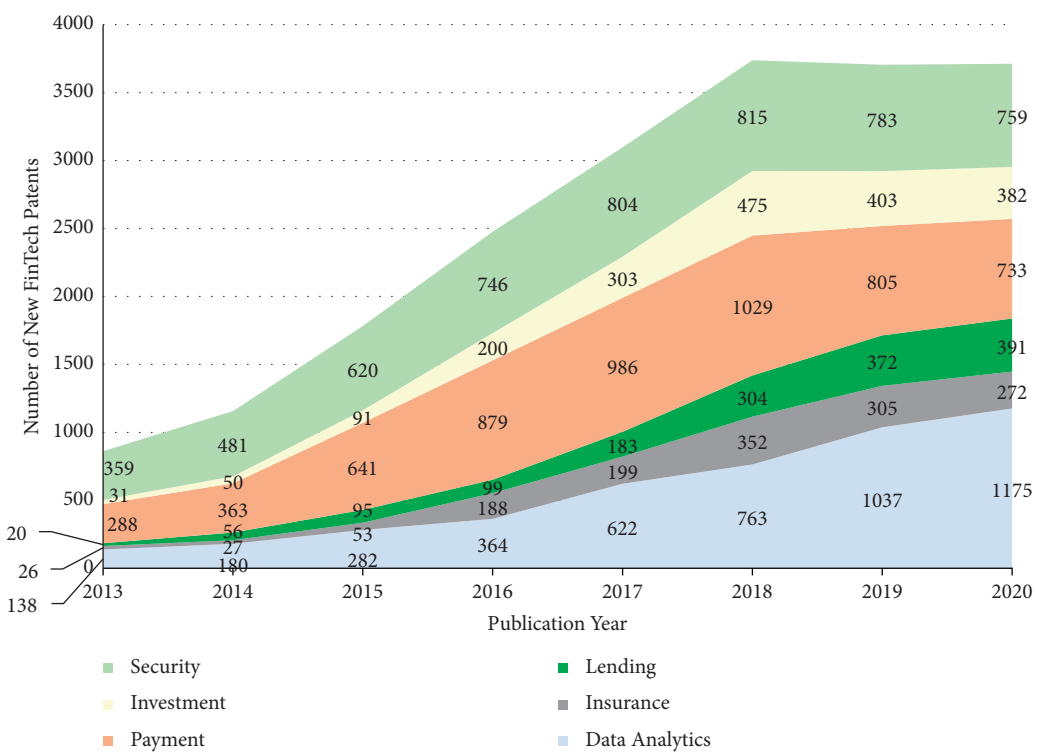


Figure 9: Annual number of new FinTech patent applications in China.

innovation. To some extent, this explains why Alipay and WeChat, rather than banks, currently dominate China's mobile payment business. In recent years, the traditional banking business has been strongly impacted by the development of FinTech. Various online lending institutions have entered traditional banking business activities such as

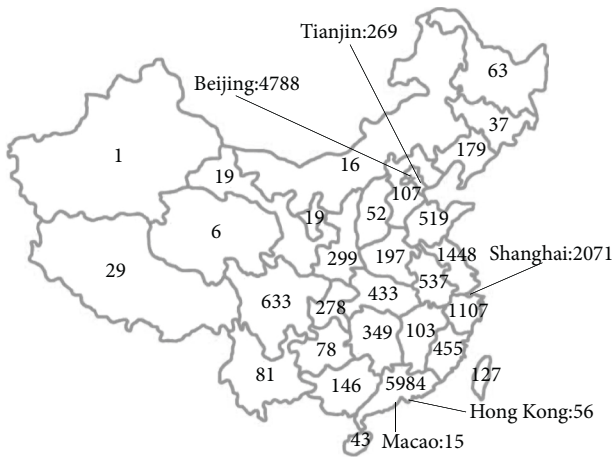


Figure 10: The distribution of FinTech patent applications at the provincial level in China.

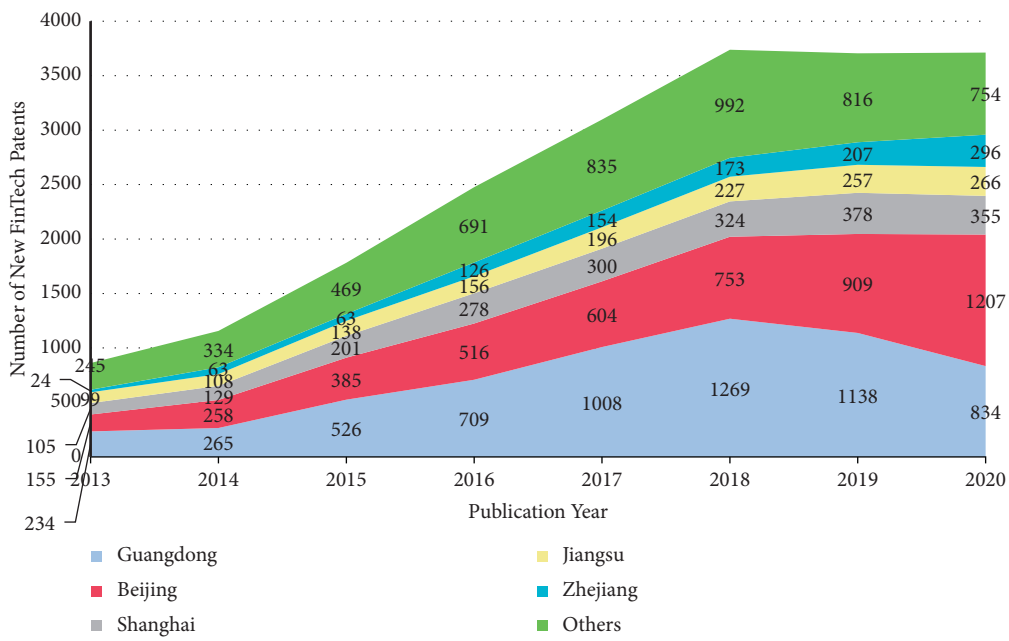


Figure 11: Proportions of the annual number of FinTech patent applications by province in China.

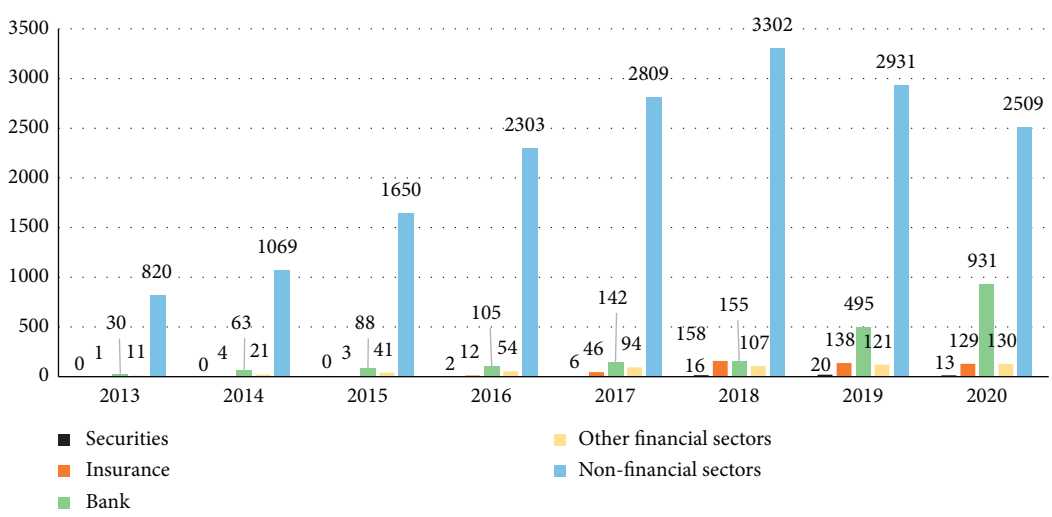


Figure 12: Annual number of FinTech patent applications in the financial and nonfinancial sectors.

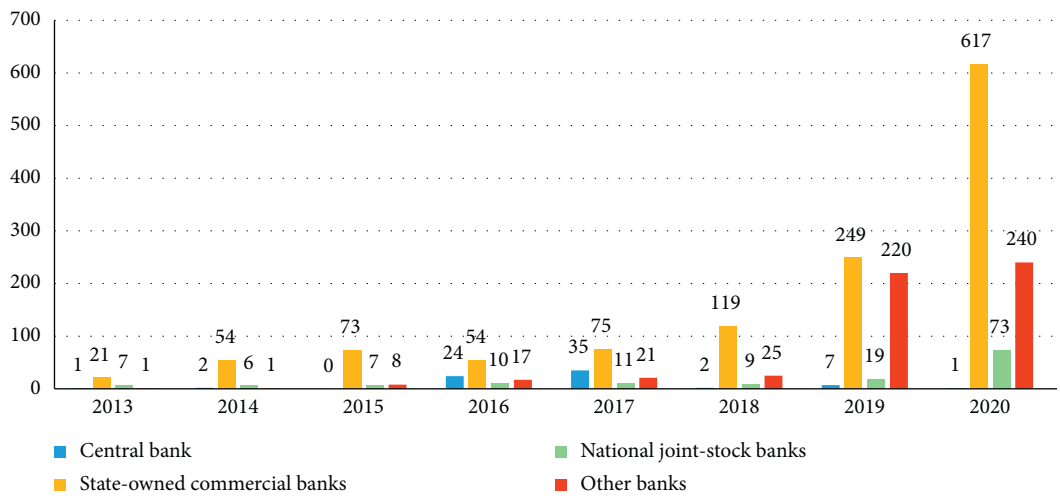


Figure 13: Annual number of FinTech patent applications of banks.

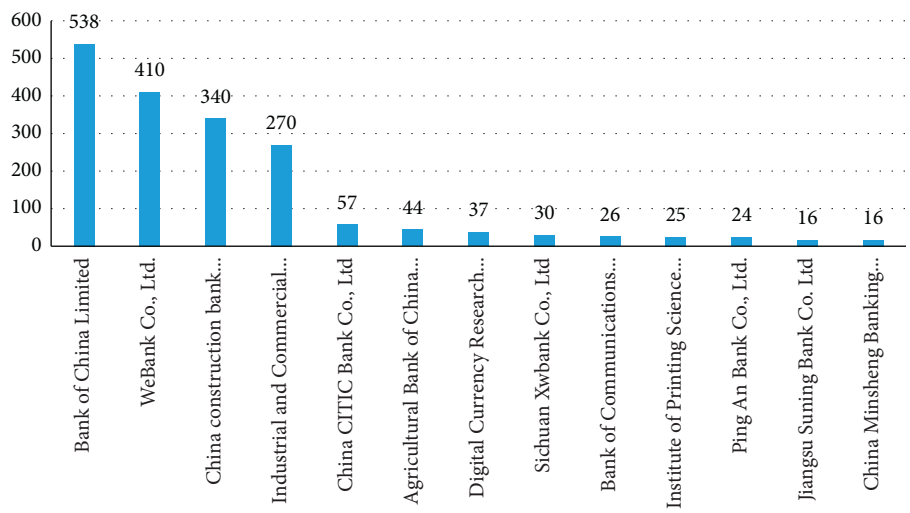


Figure 14: Banks with the highest numbers of FinTech patents.

lending. In response, banks have intensified their FinTech innovation efforts. The next round of banking competition may center on FinTech innovation. Relevant research on banking competition suggests that this industry trend may impact the development of local enterprises and the allocation of credit resources.

In China, the banking industry can be divided into central bank, state-owned commercial banks, national joint-stock commercial banks, and other banks. Figure 13 shows the number of FinTech patent applications per year for each type of bank. The central bank has a certain number of FinTech patent applications between 2016 and 2017, most of which are from the digital currency research institute of the central bank, which are products of China's digital currency strategy. The FinTech innovation of large state-owned banks has developed rapidly in recent years, which may be driven by their huge financial business and the need for digital transformation under internal and external pressures. What is more, from Figure 14, we can see that, among other banks, WeBank accounted for the lion's share of FinTech patent applications, and FinTech innovation has a high concentration.

9. Conclusions and Future Work

Based on an analysis of the connotation and function of FinTech as well as prior research related to FinTech, this study divided Chinese FinTech patents into six categories: payment, lending, insurance, security, data analysis, and investment. Using data from all patent applications published by China's State Intellectual Property Office between 2013 and 2020, we selected and classified FinTech-related patents using traditional machine learning and deep learning approaches. First, we used text filtering and manual annotation to obtain an annotated dataset for model training; then, the training and test datasets were used with different machine learning and deep learning models to compare the models' effectiveness in text categorization.

Character-level CNN gave the best classification performance and was selected to classify the patent data, ultimately resulting in the identification of 20,529 FinTech-related patents.

The classified data were then used to discuss the current situation of FinTech innovation in China. The number of annual patent applications for payment FinTech reached its peak in 2018. The high intensity of innovation led to prosperity in the mobile payment business, in line with the development of payment systems and the huge scale of business in China. Currently, data analytics patent applications account for the largest share of FinTech patent applications, indicating the future direction of FinTech.

There are distinct regional differences in the number of FinTech patent applications in China, with more applications in developed provinces than in less-developed provinces. Moreover, the concentration of new FinTech patent applications in Beijing, Shanghai, and Guangzhou is increasing each year. At the microlevel, China's banks entered FinTech innovation relatively late. Since 2019, the traditional banking industry has caught up with the FinTech challenge, and the number of FinTech patent applications from banks has grown rapidly.

The main contribution of this study is that we provide a way to search for Chinese FinTech patents in order to research on FinTech innovation while also helping the financial firms and regulators to understand the latest developments in FinTech. This study is the first study on the classification of Chinese FinTech patents. The dataset we constructed can deepen the research on FinTech innovation in China from the macrolevel to the microlevel; to be specific, it could be used to study where, when, and why FinTech innovations emerge and what their impact is on firms investing in them.

In future research, we plan to further associate FinTech patent application data with corporate microdata, industry mesodata, and economic macrodata to more deeply study the impact of FinTech innovation on the allocation of financial resources and corporate governance in China.

Data Availability

The data used to support the findings of this study are included within the supplementary information file.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (NSFC no. 72103083).

Supplementary Materials

The dataset contains all FinTech patent information, including patent name, applicant, application date, public date, number, address, and classification. (*Supplementary Materials*)

References

[1] I. Goldstein and W. Jiang, "To FinTech and beyond," *Review of Financial Studies*, vol. 32, no. 5, pp. 1647–1661, 2019.

[2] M. Harrist, "How FinTech is powering the global economy," Forbes, 2017.

[3] A. Fuster, M. Plosser, P. Schnabl, and J. Vickery, "The role of technology in mortgage lending," *Review of Financial Studies*, vol. 32, no. 5, pp. 1854–1899, 2019.

[4] H. Tang, "Peer-to-peer lenders versus banks: substitutes or complements?" *Review of Financial Studies*, vol. 32, no. 5, pp. 1900–1938, 2019.

[5] T. Berg, V. Burg, A. Gombović, and M. Puri, "On the rise of FinTechs: credit scoring using digital footprints," *Review of Financial Studies*, vol. 33, no. 7, pp. 2845–2897, 2020.

[6] C. Haddad and L. Hornuf, "The emergence of the global FinTech market: economic and technological determinants," *Small Business Economics*, vol. 53, no. 1, pp. 81–105, 2019.

[7] P. Utami and B. Basrowi, "Management of zakat payment based on FinTech for the good corporate governance improvement," *Eastern Journal of Economics and Finance*, vol. 4, no. 2, pp. 41–50, 2019.

[8] M. A. Chen, Q. Wu, and B. Yang, "How valuable is FinTech innovation?" *Review of Financial Studies*, vol. 32, no. 5, pp. 2062–2106, 2019.

[9] L. Xu, X. Lu, G. Yang, and B. Shi, "Identifying FinTech innovations with patent data: a combination of textual analysis and machine-learning techniques," in *Proceedings of the International Conference on Information*, pp. 835–843, Springer, Shanghai, China, September 2020.

[10] D. Caragea, M. Chen, T. Cojoianu, M. Dobri, K. Glandt, and M. George, "Identifying FinTech innovations using BERT," in *Proceedings of the 2020 IEEE International Conference on Big Data (Big Data)*, pp. 1117–1126, IEEE, Atlanta, Georgia, US, December 2020.

[11] J. Xu, "China's internet finance: a critical review," *China and World Economy*, vol. 25, pp. 78–92, 2017.

[12] Kamdjoug, R. E. Bawack, and J. G. Keogh, "Bitcoin, Blockchain and FinTech: a systematic review and case studies in the supply chain," *Production Planning & Control*, vol. 31, no. 2-3, pp. 115–142, 2020.

[13] T. Nakashima, "Creating credit by making use of mobility with FinTech and IoT," *IATSS Research*, vol. 42, no. 2, pp. 61–66, 2018.

[14] W. Du, S. L. Pan, D. E. Leidner, and W. Ying, "Affordances, experimentation and actualization of FinTech: a blockchain implementation study," *The Journal of Strategic Information Systems*, vol. 28, no. 1, pp. 50–65, 2019.

[15] I. Romānova and M. Kudinska, "Banking and FinTech: a challenge or opportunity?" in *Contemporary Issues in Finance: Current Challenges from across Europe*Emerald Group Publishing Limited, Bingley, United Kingdom, 2016.

[16] Y. J. Shin, "FinTech: ecosystem, business models, investment decisions, and challenges," *Business Horizons*, vol. 61, no. 1, pp. 35–46, 2018.

[17] M. Demertzis, S. Merler, and G. B. Wolff, "Capital markets union and the FinTech opportunity," *Journal of financial regulation*, vol. 4, no. 1, pp. 157–165, 2018.

[18] G. Buchak, P. Matvos, and A. Seru, "FinTech, regulatory arbitrage, and the rise of shadow banks," *Journal of Financial Economics*, vol. 130, no. 3, pp. 453–483, 2018.

[19] A. V. Thakor, "FinTech and banking: what do we know?" *Journal of Financial Intermediation*, vol. 41, Article ID 100833, 2020.

[20] C. J. Fall, A. Törcsvári, K. Benzineb, and G. Karetka, "Automated categorization in the international patent classification," *Acm Sigir Forum*, vol. 37, no. 1, pp. 10–25, 2003.

[21] K. Benzineb and J. Guyot, "Automated patent classification, current challenges in patent information retrieval," in *Current Challenges in Patent Information Retrieval*, pp. 239–261, Springer, Berlin, Heidelberg, 2011.

[22] L. Aristodemou and F. Tietze, "The state-of-the-art on Intellectual Property Analytics (IPA): a literature review on artificial intelligence, machine learning and deep learning methods for analysing intellectual property (IP) data," *World Patent Information*, vol. 55, pp. 37–51, 2018.

[23] C.-H. Wu, Y. Ken, and T. Huang, "Patent classification system using a new hybrid genetic algorithm support vector machine," *Applied Soft Computing*, vol. 10, no. 4, pp. 1164–1177, 2010.

[24] S. Li, J. Hu, Y. Cui, and J. Hu, "DeepPatent: patent classification with convolutional neural networks and word embedding," *Scientometrics*, vol. 117, no. 2, pp. 721–744, 2018.

[25] L. Xiao, G. Wang, and Z. Yang, "Research on patent text classification based on word2vec and LSTM," in *Proceedings of the 2018 11th International Symposium on Computational Intelligence and Design (ISCID)*, pp. 71–74, IEEE, Hangzhou, China, December 2018.

[26] J. Zhao, X. Li, C.-H. Yu, S. Chen, and C.-C. Lee, "Riding the FinTech innovation wave: FinTech, patents and bank performance," *Journal of International Money and Finance*, vol. 122, no. 2022, Article ID 102552, 2022.

[27] T. F. Cojoianu, G. L. Clark, A. G. F Hoepner, V Pažitka, and D. Wójcik, "Fin vs. tech: are trust and knowledge creation key ingredients in fintech start-up emergence and financing?" *Small Business Economics*, vol. 57, no. 4, pp. 1715–1731, 2021.

[28] W. Lee and S. Sohn, "Identifying emerging trends of financial business method patents," *Sustainability*, vol. 9, no. 9, p. 1670, 2017.

[29] L. Chen, "From FinTech to finlife: the case of FinTech development in China," *China Economic Journal*, vol. 9, no. 3, pp. 225–239, 2016.

[30] Y. Shim and D.-H. Shin, "Analyzing China's fintech industry from the perspective of actor-network theory," *Telecommunications Policy*, vol. 40, no. 2-3, pp. 168–181, 2016.

[31] C. Leong, B. Tan, F. T. C. Tan, and Y. Sun, "Nurturing a FinTech ecosystem: the case of a youth microloan startup in China," *International Journal of Information Management*, vol. 37, no. 2, pp. 92–97, 2017.

Hindawi
Mathematical Problems in Engineering
Volume 2022, Article ID 2674425, 14 pages
https://doi.org/10.1155/2022/2674425

Research Article

A Cross-View Gait Recognition Method Using Two-Way Similarity Learning

Y. J. Qi,[1,2] **Y. P. Kong**,[1,3] **and Q. Zhang**[3]

[1]*School of Mechanical and Electrical Engineering, Xi'an University of Architecture and Technology, Xi'an, Shaanxi 710055, China*
[2]*School of Business, Northwest University of Political Science and Law, Xi'an, Shaanxi 710063, China*
[3]*School of Information and Control Engineering, Xi'an University of Architecture and Technology, Xi'an, Shaanxi 710055, China*

Correspondence should be addressed to Y. P. Kong; kongyp@xauat.edu.cn

Received 25 March 2022; Revised 26 April 2022; Accepted 28 April 2022; Published 23 May 2022

Academic Editor: Long Wang

Gait recognition is a powerful tool for long-distance identification. However, gaits are influenced by walking environments and appearance changes. Therefore, the gait recognition rate declines sharply when the viewing angle changes. In this work, we propose a novel cross-view gait recognition method with two-way similarity learning. Focusing on the relationships between gait elements in three-dimensional space and the wholeness of human body movements, we design a three-dimensional gait constraint model that is robust to view changes based on joint motion constraint relationships. Different from the classic three-dimensional model, the proposed model characterizes motion constraints and action constraints between joints based on time and space dimensions. Next, we propose an end-to-end two-way gait network using long short-term memory and residual network 50 to extract the temporal and spatial difference features, respectively, of model pairs. The two types of difference features are merged at a high level in the network, and similarity values are obtained through the softmax layer. Our method is evaluated based on the challenging CASIA-B data set in terms of cross-view gait recognition. The experimental results show that the method achieves a higher recognition rate than the previously developed model-based methods. The recognition rate reaches 72.8%, and the viewing angle changes from 36° to 144° for normal walking. Finally, the new method also performs better in cases with large cross-view angles, illustrating that our model is robust to viewing angle changes and that the proposed network offers considerable potential in practical application scenarios.

1. Introduction

Gait recognition is an emerging biometric recognition approach used to identify people by their walking patterns. Unlike biometric identifiers such as faces, fingerprints, and irises, gaits have unique advantages, such as being noncontact, noninvasive, and hard to hide and forge, because of their characteristics during continuous cyclical movements. Early methods of gait feature analysis were used in the medical field, such as for the early diagnosis of Parkinson's disease [1], cerebral palsy [2], and other diseases. In recent years, an increasing number of cameras have been installed for security surveillance and crime forensics. Gait recognition is the most suitable human identification technology for long-distance monitoring. Another competitive advantage is that subjects also do not have to actively collaborate with other related equipment. Therefore, gait recognition offers unique advantages in terms of video surveillance. For example, the United Kingdom and Denmark have used gait information as criminal evidence. Despite the advantages of gait recognition, one of its greatest challenges involves removing factors that do not relate to human identity. These unavoidable influencing factors include walking mannerisms, such as walking speeds, clothing and carried objects, and walking environmental factors, such as camera viewing angles, road surfaces, human body occlusion, and illumination conditions. Such scenarios cause large intraclass gait feature fluctuations due to their lack of robustness. In the past 10 years, many works have researched robust gait feature extraction and have achieved promising results. In practical application scenarios, viewing angle changes are the most difficult factors due to the nonrigid movements of

the human body. Therefore, the most challenging problem for cross-view gait recognition remains an active research topic.

The gait recognition problem has been commonly studied from the perspective of computer vision, as computer vision is not intrusive with respect to this subject. The first gait recognition task is to build a model for describing gait characteristics. Appearance models form a popular current modeling method that integrates dynamic and static gait pattern information from video or image sequences into a two-dimensional (2D) feature image. Gait energy images (GEIs), motion silhouette images (MSIs), and motion history images (MHIs) are commonly used in appearance models. These appearance-based methods extract view-invariant gait features directly from planar images and achieve good results when the encountered viewing angle changes are relatively small. When the viewing angle between two subjects increases, the performance of these models decreases drastically because the features obtained from a 2D image can abundantly express gait information, but they weakly illustrate spatial change relationships. Gait is a characteristic of the human body as it moves periodically in three-dimensional (3D) space. Regardless of how often the viewing angle changes, the structure of the human body can satisfy the view-invariance characteristic. Studies in [3] showed that human joints can express gait changes. The mutual restraint relationships between joints can express the gait element characteristics, such as limb motion ranges, step lengths, foot angles, and center of gravity positions. The vector relationship formed between joints is also rotation invariant. Therefore, gait feature modeling through 3D joint poses is more in line with the spatial motion characteristics of the human body than appearance-based modeling.

Most methods based on deep learning extract gait invariant features and then measure the similarity levels for judgment purposes. When there are many sample categories and the number of samples in each category is relatively small, it is difficult to use the classification algorithm to achieve good results. At the same time, gait recognition is often necessary to directly judge the similarity of input pairs rather than solving classification problems in practical scenarios. From the perspective of gait verification, some scholars have designed methods that can directly discriminate gait similarity. To directly determine the similarity of a pair of GEIs, Zhang [4] developed an end-to-end gait measurement method based on a Siamese network. In [5], a convolutional neural network (CNN) based gait similarity measurement network was proposed by using the feature fusion technique. The paper concluded that the developed network can extract the gait feature differentials when two comparison GEIs are input directly.

Considering that the human body moves in three-dimensional space, the kinematic relationships between joints exhibit view invariance. In this work, we propose a novel two-way similarity learning method for gait recognition based on 3D human body poses. Using the constraint relationship between joints in 3D space, a view-robust 3D motion constraint matrix is constructed from the two dimensions of space and time. Then, an end-to-end similarity learning network is proposed to automatically learn the different features of input pairs. The network structure by which the spatiotemporal difference features of 3D gaits are learned is a hybrid of long short-term memory (LSTM) and residual network 50 (ResNet-50). The contributions of our work can be summarized as follows:

(1) Focusing on the overall motion constraint relationship between joints rather than the local physical motion characteristics of a single joint, a 3D gait feature model that can represent the spatiotemporal motion constraints between joints is constructed.

(2) We develop an end-to-end two-way spatiotemporal similarity learning network using LSTM and ResNet-50 to perform nonlinear feature modeling on joint motion constraint difference features and spatial action difference features and directly predict the similarity of input pairs.

(3) Compared with some state-of-the-art methods, the proposed 3D gait constraint model achieves better viewing angle robustness and cross-perspective recognition stability through learning temporal and spatial differences features.

The rest of the paper is organized as follows: In Section 2, the related works on gait recognition are reviewed. In Section 3, the 3D constraint modeling method and the proposed 3D gait similarity learning network framework are introduced. Finally, the proposed method is evaluated in Section 4, and the conclusion of this paper is presented in Section 5.

2. Related Work

The first task required for gait recognition is to build a gait model that can express gait characteristics from sequence images. Then, given a probe model and gallery models, abstract high-level gait features are obtained via machine learning. As a simple implementation, the feature similarity between each probe and gallery image pair is calculated. Finally, the gallery closest to the probe is selected as the probe identity. According to the gait model type utilized, gait recognition methods can be divided into two typical categories: appearance-based and model-based methods. The former methods extract gait representations directly from videos, and the latter model the underlying structure of the human body. Our work in this paper falls within the latter category.

2.1. Appearance-Based Methods. Appearance-based methods first obtain dynamic and static information from periodic and continuous image sequences according to specific rules. There are two types of apparent methods: template- and sequence-based methods. The template-based method constructs a two-dimensional gait feature image (also called an energy-like image) that is generated by superimposing this obtained information. Different energy-like images require specific methods for expressing the temporal and spatial gait characteristics and offer specific

strengths. Commonly used images are GEIs, MSIs, and MHIs. GEIs are the most commonly used images for appearance-based models. Sequence-based methods use pedestrian silhouette sequences as input to extract features. After this, the discriminative gait feature representation of the model is obtained by machine learning. In recent decades, appearance-based methods have produced many satisfactory results. As mentioned before, one of the difficulties for appearance models is that appearances change due to viewing angle changes. The following three approaches are commonly used to obtain view-invariant gait features.

Traditional feature extraction methods with constant viewing angles include the VTM (view transformation model), CCA (canonical correlation analysis), and subspace learning.

VTM projects gait features from one view into another. Makihara et al. [6] first proposed transforming the view of a gait template. A VTM is constructed from view-independent and object-independent matrices based on singular value decomposition (SVD). Kusakunniran et al. [7] introduced truncated SVD to avoid oversizing and overfitting the model. On this basis, the following methods further optimize the VTM method: introducing support vector regression (SVR) to model nonlinear correlation [8] and constructing a shared high-level virtual path to project a single model for different viewpoints on a single canonical view [9]. CCA projects each gait feature pair into two subspaces with maximal correlations. Bashir et al. [10] mined the low-dimensional geometric structure of feature data by CCA and learned the projection matrix of a specific view. Xing et al. [11] aimed to solve the inconsistency of CCA projections when processing two high-dimensional data sets and proposed a complete CCA method to improve the performance of the original CCA technique. Because some weakly correlated or noncorrelated information may exist in global gait features, Kusakunniran et al. [12] applied CCA to groups of features obtained from the motion coclustering of global gait features. Subspace learning projects a different view of a common subspace to measure similarity. Xu et al. [13] presented multiview max-margin subspace learning (MMMSL) to learn a common subspace for associating gait data across different views. Connie et al. [14] used dual-core principal component analysis (PCA) to perform coefficient expansion to establish a nonlinear subspace, forming a Grassman manifold to describe multiview gait features. The structure of a nonlinear subspace can more appropriately retain the gait characteristics when the viewing angle changes. The above methods alleviate the impact of view changes on gait recognition accuracy to a certain extent. However, they are still not effective in solving the highly nonlinear correlations among gait features under different views.

After 2015, deep learning methods began to appear in the gait recognition field [15–19]. CNN-based methods [9, 20–22] combine feature extraction and gait classification into one task. Furthermore, the nonlinear mapping and hierarchical feature extraction capabilities of deep networks dramatically reduce the impact of view changes on gaits. Yan et al. [23] extracted high-level gait features and introduced a multi-task learning model for state prediction during gait recognition. This method achieves better performance than gait recognition alone. Yu et al. [17] proposed a model based on an autoencoder to construct a VTM model for cross-view gait recognition. They fed walking states to the first two-layer encoder of the model, and each subsequent layer was converted by an 18° angle to reconstruct a GEI that was not affected by clothing and carried objects. In actual application scenarios, it is necessary to directly judge the similarity of two subjects while their identities are unknown. Some methods develop end-to-end gait judgment networks to determine the similarity of input pairs directly. Among these, the Siamese network is the most popular. It is a valuable tool that is significant for evaluating the similarity of pairs of different people. Zhang et al. [4] constructed an end-to-end gait similarity measurement-based Siamese network. Two convolutional subnetworks with the same structure extract the GEI features and judge the matching degrees of GEI pairs directly through a decision-making layer. Inspired by the Siamese network, Wu et al. [5] constructed gait difference feature learning networks for input pairs at the data level, feature level, and decision level and analyzed their performance. It was concluded that the data-level fusion network learns better gait difference features.

Considering that gait is a continuous movement characteristic, some scholars have also used silhouette sequences as inputs to achieve view invariance. The typical reconstruction method is based on LSTM [24] or a 3D-CNN [25, 26]. These methods fully use the temporal and spatial gait information and achieve better recognition results. Chao et al. [27] claimed that gait information has a cross-spatiotemporal correlation. They fed a disordered set of gait silhouette images to the GaitSet network to automatically learn the spatial structures and positional relationships of gaits. This method significantly improved the gait recognition rates of other approaches and has been regarded as a performance benchmark in the field of cross-view recognition. Fan et al. [28] focused on the motion characteristics of different body parts using the part-based model GaitPart to extract the spatiotemporal micromotion features of different parts from the silhouette sequence and achieved a high recognition rate. Hou et al. [29] used the GLN (gait lateral network) to learn both discriminative and compact representations from silhouettes. Then, the feature pyramid combining the silhouette-level and set-level features extracted by different stages was merged with the lateral connections in a top-down manner. Sequence-based methods introduce richer spatial information and add timing information, which improves the recognition rate compared to template-based methods.

2.2. Model-Based Methods. As mentioned before, model-based approaches model the structure or movement of the human body because of their insensitivity to human shapes or appearances. Kastaniotis et al. [30] used the human body skeleton and joint data obtained from a low-cost Kinect sensor to recognize different people. The results showed that

body joints contain sufficient information for human identification. However, in video surveillance application scenes, precisely restoring body structures from videos captured under uncontrolled conditions becomes a research difficulty [31–33].

Because gaits satisfy view invariance in 3D space, some methods attempt to use 3D imaging equipment or reconstruct 3D gait models of the human body in a collaborative multicamera environment. These approaches face problems when complicated camera parameter adjustments and modeling calculations are attempted. Therefore, accurate body part tracking becomes the bottleneck of model-based methods. Since 2016, the advent of human body pose estimation methods has made it easy to directly estimate human skeleton information from images or videos. These approaches have also advanced the model-based gait recognition technology research. Liao et al. [34] extracted 2D joint poses from video sequences by OpenPose2D and constructed a pose-based temporal-spatial network (PTSN) to obtain the spatiotemporal change characteristics of human joint sequences. In [35], pose-based LSTM (PLSTM) was used to reconstruct the view of a 12-joint heatmap sequence to reduce the impact of view changes. This method cannot wholly create a joint heatmap of gait sequences with more than three cross views simultaneously. However, while 2D poses are not view invariant, 3D poses are. With the maturity of OpenPose3D technology, the estimation of 3D skeleton data from 2D skeleton data in real time has become more accessible. Recently, Liao et al. [36] further used OpenPose3D to establish a motion constraint model for joint poses and extracted 3D joint motion features using a CNN. This method achieved a recognition rate of more than 80% for cross views of the same walking state. As mentioned above, model-based methods, especially 3D modeling methods, are reasonable gait recognition solutions.

Despite their good performance, the moving background and occlusion will have a certain impact on the apparent modeling in the actual scene. The 2D pose model is essentially a flat expression method for gait features that is also not view invariant. Therefore, human body pose estimation possesses a greater tolerance for interference from the external environment, which shows its great potential in actual gait recognition scenario application. Therefore, there remains much room to improve model-based methods.

3. Proposed Method

The gait is a periodic movement of joints driven by limbs, and it exhibits the characteristics of a constant viewing angle in three-dimensional space. Based on this, a visually robust 3D motion constraint matrix is constructed to express the motion constraint change relationship between joints. Starting from the requirements of gait verification tasks in video surveillance scenes, an end-to-end two-way similarity learning method is constructed to model and express the spatial and temporal differences of 3D motion constraints and directly judge the similarity of input pairs. The framework of the new method is shown in Figure 1.

The overall method consists of three components: a 3D joint coordinate matrix extraction (JCME) module, a 3D gait constraint matrix generation (GCMG) module, and a two-way similarity learning network (TSLN). The JCME module first automatically extracts 3D joint coordinates from continuous video frames using OpenPose3D and generates a joint coordinate matrix. After the JCME module, the GCMG module processes the obtained joint coordinate matrix to generate a 3D gait constraint matrix based on the time and space dimensions. This matrix expresses the spatiotemporal constraint relationships between joints. Finally, the 3D motion constraint matrix pairs are fed into the TSLN. The two branches of the TSLN extract the difference features of the motion constraint and action constraint. The two types of features are merged in a bottom-to-top manner to produce the final similarity output. The details of each process are described separately next.

3.1. 3D Joint Coordinate Matrix Extraction Module. To establish a 3D gait model independent of the viewing angle, the first task of the new method is to obtain the 3D coordinates of the human joints. The role of the JCME module is to generate a 3D coordinate matrix. Pedestrian videos are directly fed to the JCME module. Here, the pretrained OpenPose3D model is used to estimate the 3D coordinates of 14 joints (head, neck, LShoulder, LElbow, LWrist, RShoulder, RElbow, RWrist, LHip, LKnee, LAnkle, RHip, RKnee, and RAnkle). Positions and joint names can refer to the COCO output format of OpenPose from the video frames. The 3D coordinates of the i-th joint in the m-th frame are denoted by $N_i^m = (x_i^m, y_i^m, z_i^m)$. Then, a 3D coordinate matrix is developed to express the mutual joint position relationship in 3D space. The row vectors of the 3D coordinate matrix are composed of the 14 joint coordinate values in the X, Y, and Z directions of each frame.

3.2. 3D Gait Constraint Matrix Generation Module. Earlier work showed that joint motion is sufficient to identify different subjects [37]. The research focused more on the local physical dependencies between joints while ignoring the implicit correlations among joint motions. From the perspective of medical research, gait elements such as the joint angle, step length, step width, position of the body gravity center, and step frequency are the characterizing gait factors. They compose an organic whole that interacts and constrains the movement of the joints driven by the limbs. The motion constraint relationship between the joints in the three-dimensional space can effectively express the spatiotemporal motion characteristics of the gait. Therefore, our goal is to establish a 3D gait constraint model to express the motion relationship of gait elements in 3D space.

From the previous module, we obtained the 3D coordinate matrix of 14 joints. Here, the main function of the GCMG module is to build a 3D gait constraint model. The main workflow of the GCMG module is shown in Figure 2. Through the work of a temporal constraint extractor and temporal motion constraint computation, temporal motion

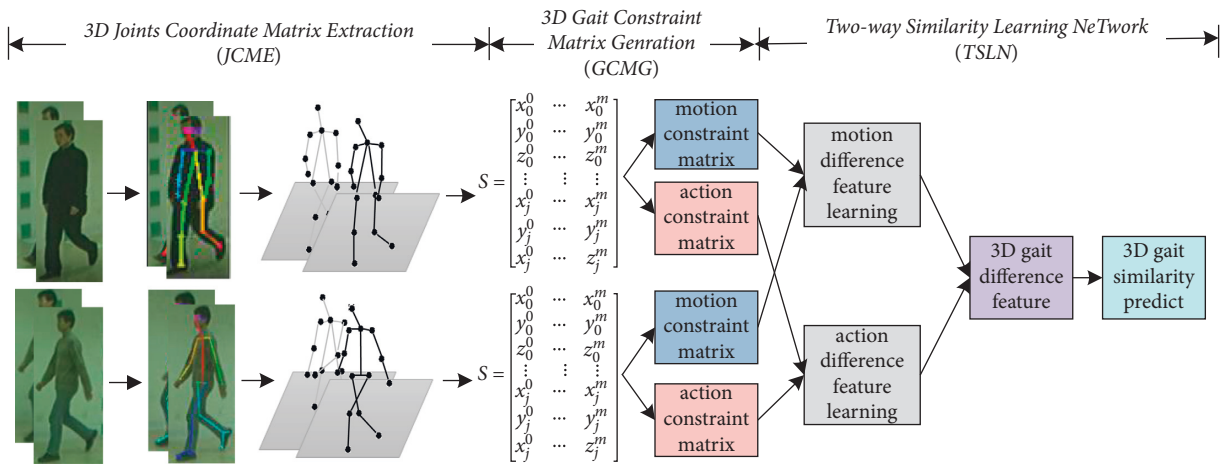


Figure 1: The framework of 3D gait similarity learning.

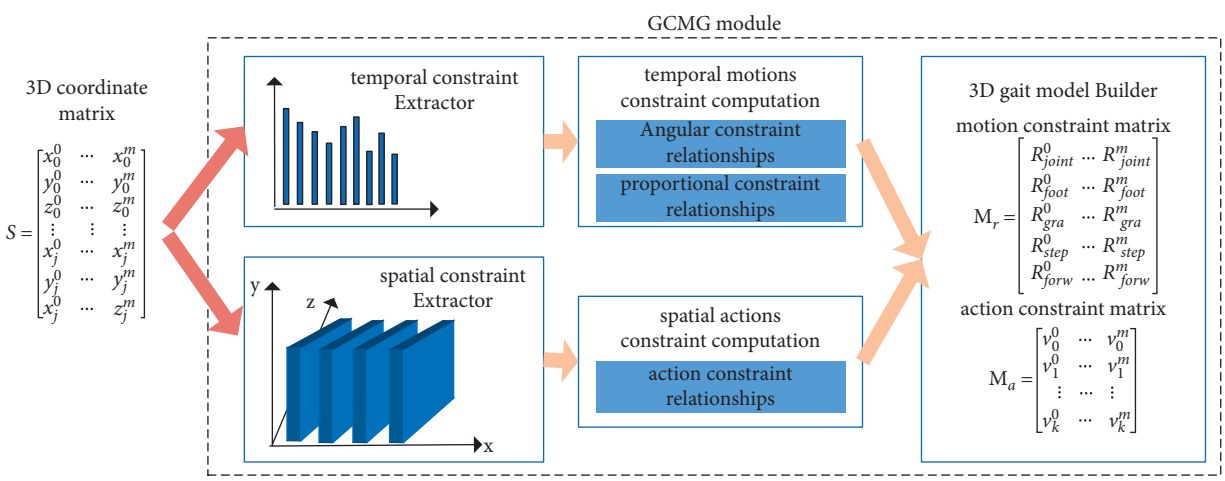


Figure 2: Workflow of the GCMG module.

constraint vectors are generated. In the same way, the spatial action constraint vectors are generated through the work of a spatial constraint extractor and a spatial actions constraint computation. Then, the 3D gait model builder constructs a 3D gait constraint matrix combining the two types of vectors obtained. The 3D gait constraint matrix consists of a motion constraint matrix and an action constraint matrix. The implementation details are described in the following section.

3.2.1. Motion Constraint Matrix Construction. The temporal motion constraint relationships concern constraint change relationships between joints in time series. These change relationships are divided into two types: angular constraint relationships and proportional constraint relationships. An angular constraint relationship is composed of an adjacent joint angle, a foot angle, and a body gravity deviation angle. The ratio of the height to the step length and the ratio of the joint vertical height compose a proportional constraint relationship.

Figure 3 shows the constraint relationships between joints. Figure 3(a) shows the 14 joints estimated by OpenPose3D. In Figure 3(b), α describes the angular constraint between adjacent joints, denoted by R^m_{joint}. In Figure 3(c), β describes the angular constraint of the foot direction, denoted by R^m_{foot}. γ describes the angular constraint of the gravity deviation, denoted by R^m_{foot}. These angular constraint relationships express the interactions between joints. In addition, movement information of the human body structure is critical for gait characterization. Feng et al. [33] used the changes in a joint heatmap to express gait characteristics while ignoring the movement changes in the human body structure. The dynamic proportional constraint of the human body structure expresses the overall movement characteristics of the human body. Two types of proportional constraint relationships are constructed here. The first is the constraint of the ratio between body height and step length. Figure 3(d) shows the vertical body height and step length (the proportion of z_0 to d_{10-13}), denoted by R^m_{step}. The second is the constraint of the ratio between the vertical height of the body and the joints, denoted by R^m_{forw}. The

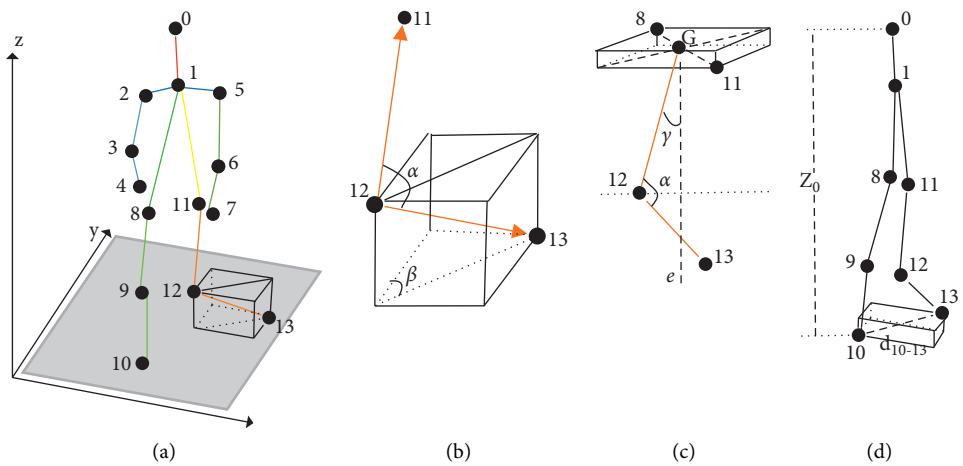


Figure 3: Joint motion constraint relationships: (a) 3D pose of human body joints, (b) restraint relationship of the left knee joint and left ankle joint, (c) center of gravity and joint constraints, and (d) schematic diagram of the height and step length constraints.

calculation processes of the above five constraint relationships are defined in formulas (1)–(5). $\overline{n}$ is the normal vector of the XOY plane in

$$\begin{cases} R_{\text{joint}}^{m} = \alpha = \cos^{-1}\left(\dfrac{\overline{N_iN_j}\cdot\overline{N_iN_j}}{|\overline{N_iN_j}||\overline{N_iN_j}|}\right), \\ R_{\text{foot}}^{m} = \beta = \cos^{-1}\left(\dfrac{\overline{N_iN_j}\cdot\overline{n}}{|\overline{N_iN_j}|}\right), \\ R_{gra}^{m} = \gamma = \cos^{-1}\left(\dfrac{\overline{N_GN_i}\cdot\overline{N_GN_j}}{|\overline{N_GN_i}||\overline{N_GN_j}|}\right), \\ R_{\text{step}}^{m} = \dfrac{z_0}{\sqrt{(x_{10}-x_{13})^2+(y_{10}-y_{13})^2}}, \\ R_{\text{forw}}^{m} = \dfrac{z_0}{\sqrt{(x_0-x_i)^2+(y_0-y_i)^2}}. \end{cases} \tag{1}$$

3.2.2. Action Constraint Matrix Construction. Because the joint action of each frame is also different, spatial action constraint relationships can be used to express alternative joint position relations of the gait elements such as step length, step width, and joint displacement in 3D space. As shown in Figure 4, the displacement vector v_k^m of the k-th joint in the m-th frame describes the action difference of the joints in the space dimension, which can be defined in formula (2), where Δt represents the time interval between two consecutive frames.

$$v_k^m = \frac{N_k^m - N_k^{m-1}}{\Delta t}. \tag{2}$$

According to formula (3), the 3D gait model builder of the GCMG module combines the five motion constraint vector types to generate a 3D motion constraint matrix named M_r, which describes the motion constraint features of joints in a time series. The sequence vector v_k^m consists of a 3D action constraint matrix named M_a, which describes the action constraint features of joints in space. Matrices M_r and M_a are called 3D gait constraint matrices, and their structures are shown in Figure 2. The established 3D gait constraint matrix not only expresses the relationships among the joint constraint changes in time and space dimensions but is also robust to view changes.

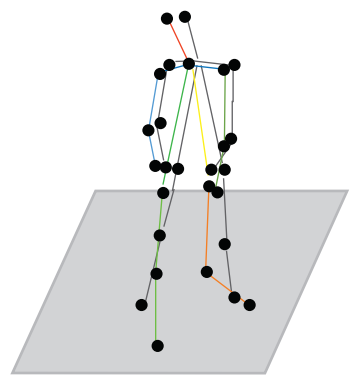


Figure 4: Schematic diagram of joint displacement.

$$\begin{cases} M_r = \left[R_{\text{joint}}^m, R_{\text{foot}}^m, R_{gra}^m, R_{\text{step}}^m, R_{\text{forw}}^m\right]^T, \\ M_a = \left[v_k^m\right]^T. \end{cases} \tag{3}$$

3.3. Two-Way Similarity Learning Network. As mentioned earlier, the gait verification task is efficient and valuable because it is often necessary to judge whether the identities of two people are the same when the identities are unknown in real scenarios. The 3D gait constraint matrix generated by the GCMG module expresses view-robust gait features from both temporal and spatial dimensions. Instead of obtaining the gait difference features of input pairs from GEI, we built a two-way similarity learning network named TSLN to

model nonlinear spatiotemporal difference features of 3D gait constraint matrix input pairs. In TSLN, the temporal motion constrained difference extractor learns the motion constrained difference features of M_r, and the spatial motion constrained difference extractor learns the motion constrained difference features of M_a. The spatiotemporal difference feature blocks fuse the two difference features obtained above and achieve the similarity prediction result of the input pair through two fully connected layers (FC-1 and FC-2) and a softmax layer. The structure of the TSLN module is shown in Figure 5.

3.3.1. Motion Constraint Difference Feature. This feature is obtained by the temporal motion constraint difference feature extractor called the LSTM branch. M_r expresses the motion constraint relationship between joints in time series, so the LSTM network is used for nonlinear modeling of the temporal difference features of the joint motion of a pair of M_r because of its excellent performance in learning the temporal dependencies of long data sequences. LSTM has memory units and a gate mechanism. Its special multiplicative units are called gates and are used to control the flow of information, and the memory units contain memory cells with self-connections that store the temporal cell states of the network. Paper [6] specified that one layer taking in two inputs can simulate the subtraction operation to compute the difference between a pair of features in the context of neural networks. Therefore, the input of the LSTM branch is the linear difference between the two inputs, that is, $M_{r1} - M_{r2}$.

There are three sigmoid gates to protect and control the cell state at each sequence index position of each LSTM layer: the forget gate, input gate, and output gate. Here, the input of each sequence index is a combination of row vectors containing two M_r at time t: that is, $m_{r1}^t - m_{r2}^t$. The output information of the forget gate, input gate, and output gate is described in formulas (4)–(6), where σ is a sigmoid activation function, and W_f, W_i, and W_o and b_f, b_i, and b_0 are the weights and biases of the forget gate, input gate, and output gate, respectively.

$$f_t = \sigma\left(W_f \cdot \left[h_{t-1}, m_{r1}^t - m_{r2}^t\right] + b_f\right), \tag{4}$$

$$i_t = \sigma\left(W_i \cdot \left[h_{t-1}, m_{r1}^t - m_{r2}^t\right] + b_i\right), \tag{5}$$

$$o_t = \sigma\left(W_o \cdot \left[h_{t-1}, m_{r1}^t - m_{r2}^t\right] + b_0\right). \tag{6}$$

This branch consists of two LSTM layers and a flattening layer. The input is an 88-dimensional vector; the output of the flattening layer is a 4,400-dimensional vector, which represents the motion constraint difference feature of $M_{r1} - M_{r2}$, denoted by D_r.

3.3.2. Action Constraint Difference Feature. This feature is achieved by the spatial action constraint difference feature extractor called the ResNet-50 branch. M_a expresses the spatial action constraint features of joints, so it is necessary to use a network that can express local features. ResNet-50 is suitable for this task. First, its properties such as convolution, nonlinear activation, and pooling are suitable for extracting local input-dependent features and generating high-level features. Second, the similarity between different subjects may be higher than that of the same subject due to similar body size and the influence of external factors. This can prevent the loss of low-level features in the process of high-level feature extraction in order to fully express the difference features of subjects. Third, the dimensionality of M_a is relatively low compared to the appearance model, and vanishing gradient problems may occur during convolution. ResNet-50 can avoid the problem of vanishing gradients. Thus, ResNet-50 is used to model nonlinear action difference features from the linear action difference of the M_a pairs, namely $M_{a1} - M_{a2}$.

ResNet-50 is formed by repeatedly superimposing a large number of residual blocks. It sends the output of the initial layer to the following layers in a leap and then accumulates it together with the outputs of these layers. The residual block structure is shown in Figure 6. $m_{a1} - m_{a2}$ is the input of the residual block, that is, the difference feature of the previous block. f is the learned residual mapping. The activation function is a rectified linear unit (ReLU), and the bias is omitted here. $m_{a1} - m_{a2}$ crosses two layers and is accumulated with the mapping feature $f(m_{a1} - m_{a2})$ as the output of the residual block. The output of this branch is the nonlinear weight difference feature D_a.

The two channels of ResNet-50 are fed to 45-dimensional vectors to simulate subtraction operations $M_{\mathrm{a}1} - M_{\mathrm{a}2}$. The two input parts are reweighted through the convolutional layer, and then new weights are added to obtain the simulated subtraction result. Finally, the flattening layer of this branch achieves a 4,096-dimensional action constraint difference feature, denoted by D_a.

3.3.3. Spatiotemporal Difference Features. Following two branches, the spatiotemporal difference features block first merges D_r and D_a, that is, it performs $\mathrm{Concat}(D_r, D_a)$ operations. Through the two fully connected layers, the dimensions of $D_r + D_a$ are reduced to 128-dimensional 3D gait difference feature vectors. The last layer of this extractor uses the softmax loss function to directly obtain the 3D gait similarity value V_{Sim}. The detailed parameters of the TSLN module are shown in Table 1.

4. Experiments and Results

Three experiments were conducted to evaluate the performance of the proposed method. The new method performance evaluation experiments were first conducted under cross-view and cross-walking conditions to determine the performances of the constructed 3D gait model and the similarity learning network. Next, two-branch evaluation experiments were conducted to determine the extraction capacity of the temporal and spatial difference features. Finally, comparative analysis experiments were conducted to compare the performance of our method and other popular state-of-the-art methods.

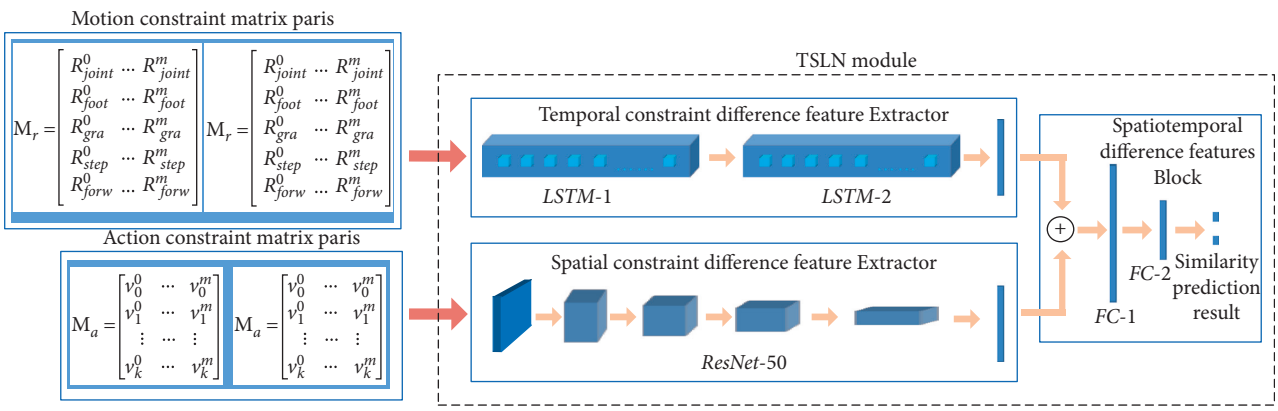


Figure 5: Structure of the TSLN module.

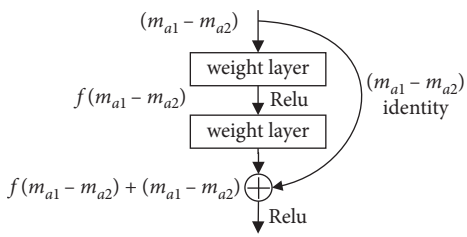


Figure 6: Residual block input mapping.

Table 1: Parameters of the TSLN module.

	Layer	Input	Output	Activation function
Temporal motion constraint difference extractor	LSTM-1	88×50	88×50	Tanh
	LSTM-2	88×50	88×50	Tanh
	Flatten		4,400	
Spatio action constraint difference extractor	Residual block	See ResNet-50 parameters		ReLU
	Flatten		4,096	
Spatiotemporal difference features Extractor	FC-1	8,496	1,024 (dropout)	
	FC-2	1,024	128 (dropout)	
	Softmax	128	2	

4.1. Data Sets. Currently, OU-MVLP and CASIA-B are popular public gait data sets widely used by the gait recognition research community. The OU-MVLP is the largest gait data set and is tailored for gait recognition methods that rely on silhouettes or GEIs. It does not provide RGB images. The proposed method obtains joint coordinates by body pose estimation from video sequences and then constructs a view-robust 3D motion constraint matrix model. Therefore, the proposed method cannot be evaluated on OU-MVLP due to its provided data. CASIA-B is a multiview gait data set provided by the Institute of Automation, Chinese Academy of Sciences, which provides pedestrian video sequences. Therefore, we use the CASIA-B gait database to evaluate the proposed method. The CASIA-B contains 124 pedestrians, where each pedestrian has 11 corresponding view angles (every two views are separated by 18°, i.e., 0°, 18°, . . ., 180°) and three kinds of walking conditions (walking with a bag (BG), walking with a coat (CL), and normal walking (NM)). The resolution of the video is 320×240, and the frame rate is 25 fps. To acquire the 3D gait constraint matrix of each person, we use OpenPose (developed by Carnegie Mellon University (CMU)) to obtain 14 joint coordinates of the human body and construct $120 \times 11 \times 10 \times 2$ gait constraint matrices with the GCMG module. Figure 7 shows some of the motion constraint change relationships of #001, #010, and #068. Figures 7(b)–7(f) show some of the motion constraint relationships over a period of time. The motion constraints of different subjects display some differences under the three walking conditions.

The training process of the TSLN module uses the motion constraint matrix containing all the walking conditions of #001–074, the gallery set uses NM#01–04 of #075–124, and the probe set uses NM#05–06, BG#01–02, and CL#01–02. The number of positive samples was $109 \times 110 \times 124$. To ensure a balanced distribution of positive and negative samples, we comprehensively consider the viewing angles and walking conditions to evenly match the negative samples of the two subjects. Ultimately, we generate approximately 1.5 million negative samples.

4.2. Method Performance Evaluation. We first conduct cross-view and cross-walking condition experiments to evaluate the performance of the new method. The average recognition rate of each viewing angle in the probe set is shown in Table 2. For the three conditions, the average recognition rates at 0° and 180° are the lowest. The reason for this may be that when subjects are facing the camera or facing away from it, the accuracy of 3D pose estimation is relatively low. This situation, in turn, leads to deviations in the 3D coordinates of the joints and further affects the accuracy of the 3D motion constraint matrix. The average recognition rate at 54°–108° is better than that at other viewing angles.

Furthermore, the recognition rates of BG and CL are lower than that of NM: that is, CL < BG < NM. This conclusion is consistent with the results in [6], but the influencing factors of the two conclusions are different. In [6], pedestrians walking with a bag and with a coat affected their gait silhouettes, which further impacted GEI generation. In

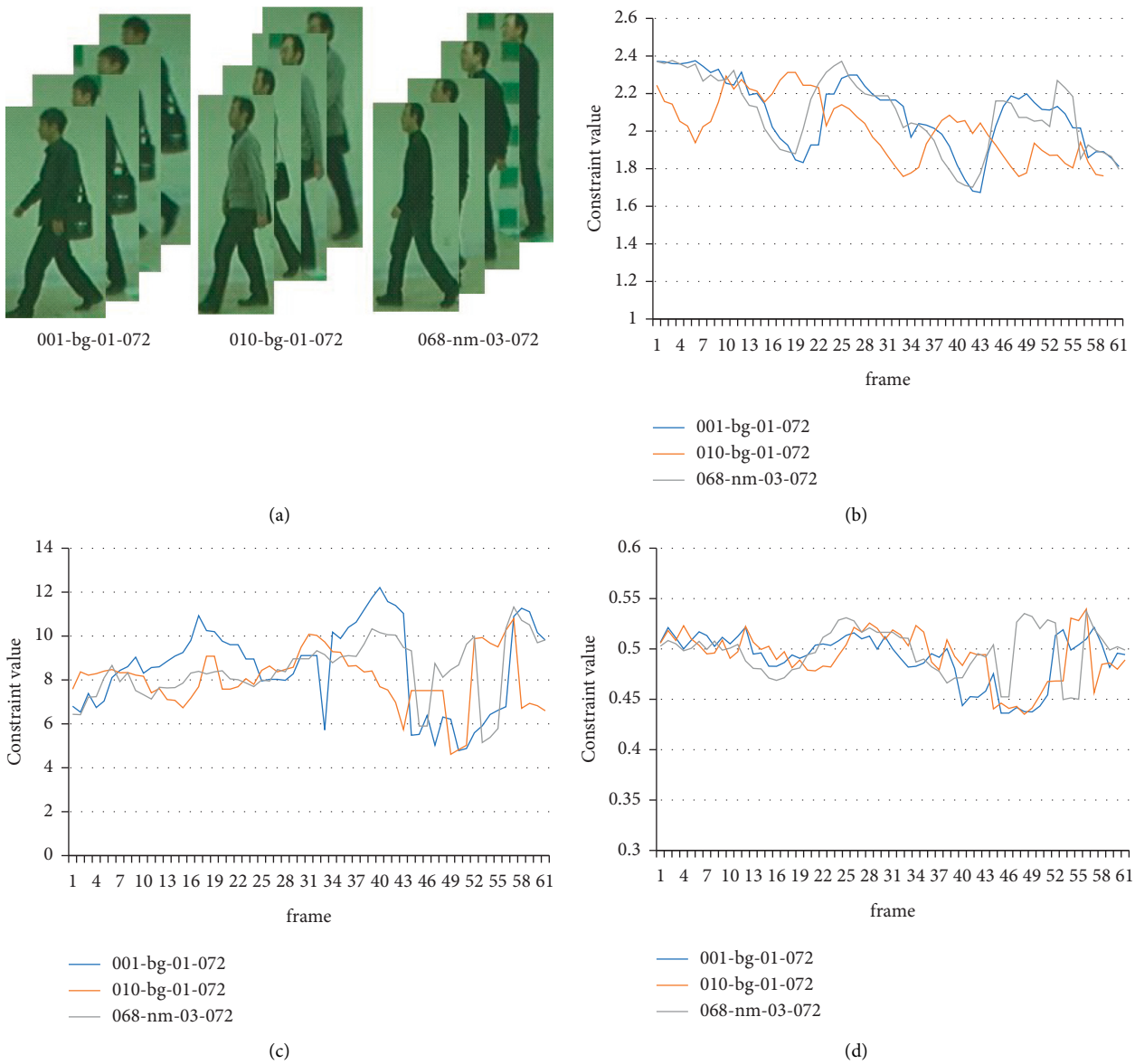


Figure 7: Continued.

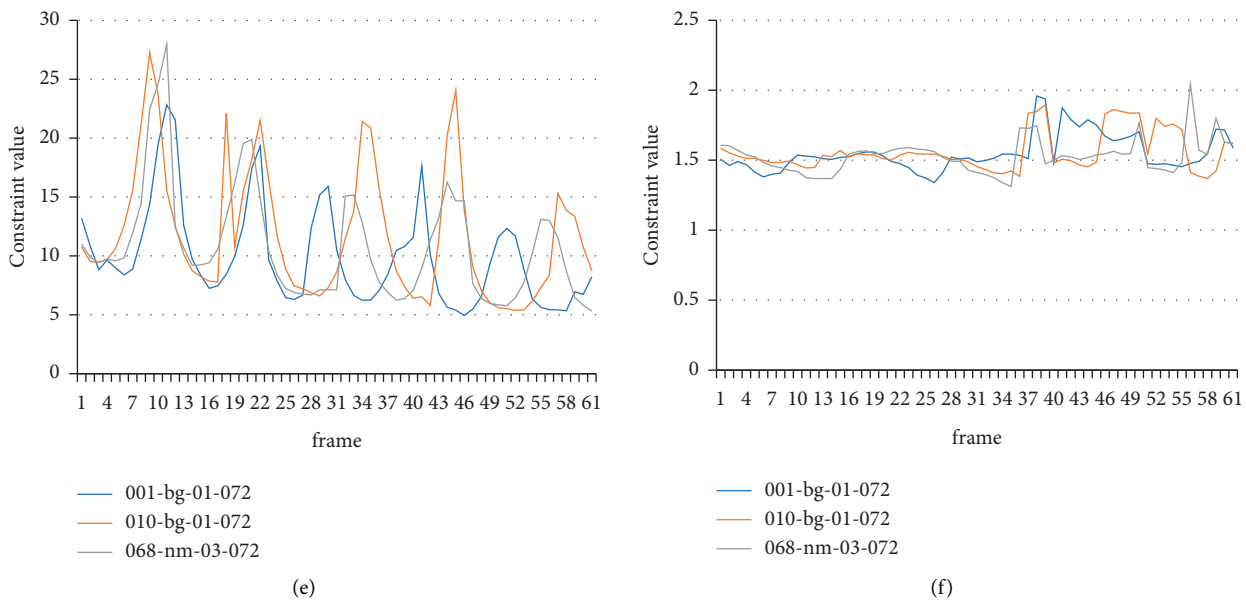


Figure 7: Motion constraint change relationships of #001, #010, and #068 at 72°: (a) video frame, (b) constraint changes of the left knee joint, (c) constraint changes in the neck and left hip joint, (d) constraint changes in the gravity center of the body and left hip joint, (e) the ratio constraint changes in the vertical height of the body to the step length, and (f) constraint changes in the right hip joint and right knee joint.

Table 2: Recognition rates (%) of cross-view and cross-walking states on CASIA-B.

Gallery set		Probe set (NM#05–06)										
		0°	18°	36°	54°	72°	90°	108°	126°	144°	162°	180°
NM#01-04	NM#05–06	55.7	68.5	72.5	75.5	70.9	69.4	71.4	73.8	76.3	68.3	55.8
	BG#01–02	37.3	44.7	50.8	47.6	51.8	42.2	47.6	51.8	52.4	39.2	35.9
	CL#01–02	19.6	23.4	41.5	36.6	36.1	31.7	39.2	37.0	36.3	32.8	24.8

Table 3: Recognition rates (%) under normal walking conditions on CASIA-B.

Gallery set (NM#01–04)	Probe set (NM#05–06)										
	0°	18°	36°	54°	72°	90°	108°	126°	144°	162°	180°
0°	92.1	84.3	66.1	54.2	28.3	30.2	38.3	34.2	47.9	56.5	64.2
18°	86.3	93.1	87.2	85.9	60.6	53.2	50.2	46.7	52.5	73.8	62.7
36°	72.4	84.8	95.7	91.1	83.8	75.5	78.6	66.8	65.4	68.2	52.1
54°	47.5	78.3	86.9	94.9	87.5	74.9	84.7	77.5	74.7	45.1	31.6
72°	35.3	56.4	83.8	86.3	96.5	92.1	83.5	85.1	88.2	39.3	25.4
90°	24.2	49.4	75.2	89.6	91.1	95.6	89.3	89.4	77.8	47.9	24.7
108°	37	67.7	79.5	83.7	91.9	90.9	94.6	94.7	86.6	68.2	43.8
126°	39.2	46.9	59.2	79.4	77.4	80.3	87.8	95.7	89.2	79.7	51.6
144°	52.7	58.3	63.7	67.8	82	76.6	85.2	88.2	95.4	88.6	75.8
162°	61.3	67.3	49.8	64.5	51.4	68.2	65.9	84.1	85.3	96.4	86.4
180°	66.8	66.5	50.9	32.9	29.3	25.4	27.3	49.6	76.1	87.1	95.3

the new method, walking with a coat affects the accuracy of 3D pose estimation and further impacts the 3D gait constraint matrix generation. Additionally, walking with a bag has less influence on the joint coordinate estimation than walking with a coat. Similar to those for the NM state, the average recognition rates for the viewing angle range of 36o-144° are 49.2% and 36.9%, respectively, which are higher than those obtained at 0° and 180°.

Table 3 lists the recognition rate of each viewing angle under normal walking conditions. The recognition rate of the same angle is higher than that of the cross-view, and the average recognition rate is 94.7%. At 36°–144°, the recognition rate reaches 95.3%. When the viewing angles difference of input pairs differ by approximately 90°, the recognition performance drops sharply. However, it is interesting that the recognition performance instead increases when the viewing angle difference between input pairs is close to 180°. For example, when the angle of the probe set sample is 18° and the angle of the gallery set sample is 126°, the recognition rate is 46.9%. When the angle of the gallery

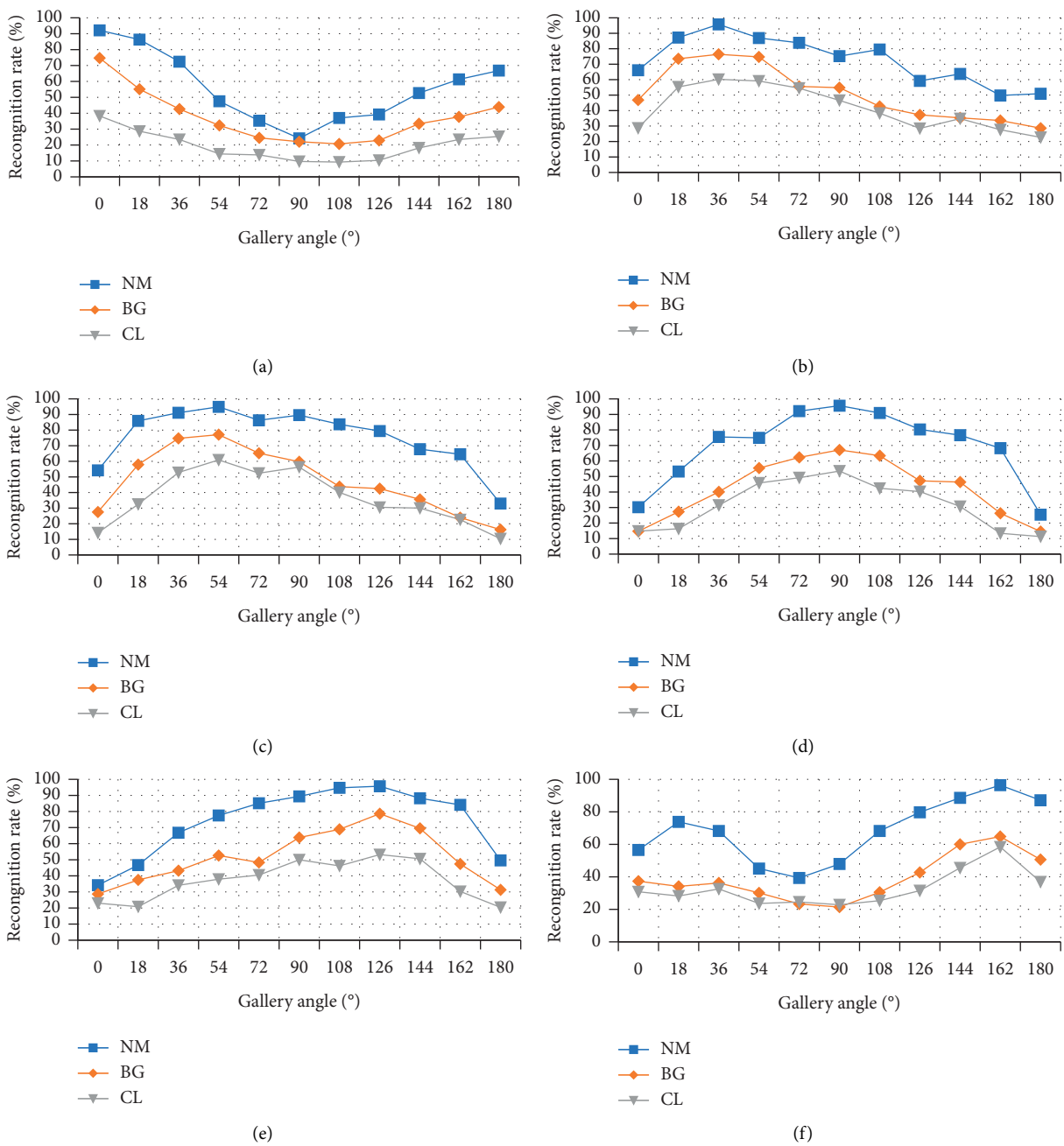


Figure 8: The average recognition rate of the three walking conditions at given view angles on CASIA-B: (a) probe angle at 0°, (b) probe angle at 36°, (c) probe angle at 54°, (d) probe angle at 90°, (e) probe angle at 126°, and (f) probe angle at 162°.

Table 4: Average recognition rates (%) of the LSTM branch, CNN branch, and proposed method on CASIA-B.

	Gallery set	Probe set			Mean
		NM#05–06	GB#01–02	CL#01–02	
LSTM branch	NM#01-04	62.4	42.7	31.7	45.6
ResNet-50 branch		54.1	38.9	27.5	40.2
Proposed method		67.9	47.8	33.9	49.9

set is 162°, the recognition rate rises to 67.3%. That is, the joint 3D coordinates of the two complementary angles are closer than those of the other angles.

Figure 8 shows the average recognition rate of the three walking states at the probe set view angles of 0°, 36°, 54°, 90°, 126°, and 162°.

4.3. Branch Performance Evaluation. Next, the effectiveness of the 3D gait model and proposed two-way network is verified, and the results are shown in Table 4. The LSTM branch is used alone to distinguish the difference features of

Table 5: Average recognition rate (%) comparison of the proposed method with DeepCNNs and PoseGait.

Gallery set	DeepCNNs [5]			PoseGait [36]			Ours		
				NM #01-04					
Probe set	NM#05–06	BG#01–02	CL#01–02	NM#05–06	BG#01–02	CL#01–02	NM#05–06	BG#01–02	CL#01–02
0°	88.7	64.2	37.7	55.3	35.3	24.3	55.7	37.4	26.8
18°	95.1	80.6	57.2	69.6	47.2	29.7	68.5	49.3	32.1
36°	98.2	82.7	66.6	73.9	52.4	41.3	72.5	50.1	43.6
54°	96.4	76.9	61.1	75.0	46.9	38.8	75.5	51.5	39.3
72°	94.1	64.8	55.2	68.0	45.5	38.2	70.9	46.7	41.2
90°	91.5	63.1	54.6	68.2	43.9	38.5	69.4	43.0	39.6
108°	93.9	68.0	55.2	71.1	46.1	41.6	71.4	47.5	39.9
126°	97.5	76.9	59.1	72.9	48.1	44.9	73.8	50.9	43.7
144°	98.4	82.2	58.9	76.1	49.4	42.2	76.3	49.3	42.5
162°	95.8	75.4	48.8	70.4	43.6	33.4	68.3	45.4	35.7
180°	85.6	61.3	39.4	55.4	31.1	22.5	55.8	33.8	24.2
Mean	94.1	72.4	54.0	68.7	44.5	35.9	68.9	45.9	37.2
Total mean	73.5	49.7	50.7						
Variances	14.9	61.2	70.8	46.0	34.7	51.1	44.7	29.8	40.9
Average variances	49.0	43.9	38.5						

a pair of joint motion constraint matrices, and the recognition rate is 45.6%. The CNN branch is used alone to distinguish the similar features of a pair of joint action constraint matrices, and the recognition rate is 40.2%. The average recognition rate of the dual-branch network is 49.9%. These results show that the constructed 3D motion constraint matrix can fully express the spatiotemporal nature of gaits.

4.4. Comparison with State-of-the-Art Methods. Finally, we perform a comparative performance analysis of our approach with respect to other popular state-of-the-art gait recognition techniques, including the appearance-based methods DeepCNNs [5] and the model-based method PoseGait [36]. Table 5 shows the comparison results of the three methods. The proposed method and the LB network of DeepCNNs directly mix two comparative samples to simulate the subtraction operation and learn the difference features between the input pairs to obtain similarity values. DeepCNNs use a CNN-based similarity learning network to learn the difference features of a pair of GEIs, and the average recognition rate is 73.5%. This result is higher than those of our method and PoseGait. As stated in the literature [5], DeepCNNs use the higher-dimensional features of the input GEI, while PoseGait uses 14 human joints as gait features, and our method uses 14 human joints and the center of gravity of the body as the gait features. Another reason for the results above is that the 3D joint coordinates are estimated twice (the two-dimensional joint coordinates are estimated from the video sequence first, and then the three-dimensional coordinates are estimated from the two-dimensional coordinates), not directly from the image sequence (such as for GEIs), which induces more significant errors. Compared with PoseGait, our method uses LSTM and ResNet-50 to learn the spatiotemporal difference features of joint constraints, and the average recognition rate is improved to a certain extent. In addition, the number of training samples is greater than 2 million in our method. The recognition results show that it is feasible to measure the similarities among input pairs separately from the time and space dimensions.

To evaluate the stability of their recognition performance, the variances in the recognition rates of all viewing angles were quantified. The variances of the three walking conditions for the three methods are listed in the second to last row of Table 5. Our variance is the smallest at BG and CL, 29.8 and 40.9, respectively. For the three walking states, the variance of the recognition rate of our method is 38.5, which is lower than the values of the other two methods. The average recognition rate of the proposed method is slightly higher than that of PoseGait. However, our method is much more stable than PoseGait in cross-view recognition. This illustrates that the proposed 3D gait constraint matrix is robust to viewing angle changes.

The sequence-based GaitPart method achieves excellent recognition performance: the recognition rate is 78.7% under CL walking conditions on CASIA-B. Although the recognition rate of our method is not as good as that of GaitPart, in the case of a large viewing angle span, our method using low-dimensional features still achieved good recognition stability (variance statistics: GaitPart – 39.3 and our method – 40.9).

5. Conclusions

This work proposes a two-way similarity learning method based on 3D poses to realize end-to-end gait recognition. First, considering gait invariance in 3D space and the periodic variation in joint motion constraints, we propose a 3D gait motion constraint matrix to express the time sequence constraint changes and spatial constraint changes in the 3D space gait. Then, the LSTM branch is used to learn the temporal constraint difference feature, and the ResNet-50 branch learns the spatial motion difference feature of the input pair. The method was evaluated with respect to the

CASIA-B database, and the results show that the recognition rate significantly improved in the case of large changes in the view angle. The proposed 3D gait model has good robustness for view angle and interfering factors such as clothing variation, but the feature dimension is low. Future work will consider combining apparent-based methods to study gait recognition methods based on multimodal gait models and further improve the accuracy of cross-view gait recognition.

Data Availability

The processed data used to support the findings of this study have not been made available because the data also form part of an ongoing study.

Conflicts of Interest

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

Acknowledgments

This work was supported by The Key Research and Development Program in Social Development Domain of Shaanxi Province (Grant no. 2022SF-424) and the Natural Science Project of Xi'an University of Architecture and Technology (Grant no. ZR19048).

References

[1] D. S. Kim, G. A. Tung, U. Akbar, and J. H. Friedman, "The evaluation of the swallow tail sign in patients with parkinsonism and gait disorders," *Journal of the Neurological Sciences*, vol. 428, Article ID 117581, 2021.

[2] D. Karabulut, Y. Z. Arslan, F. Salami, S. I. Wolf, and M. Götze, "Biomechanical assessment of patellar tendon advancement in patients with cerebral palsy and crouch gait," *The Knee*, vol. 32, pp. 46–55, 2021.

[3] S. Niyogi and E. Adelson, "Analyzing and recognizing walking figures in XYT," in *Proceedings of the Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition*, pp. 469–474, IEEE, Seattle, WA, USA, June 1994.

[4] C. Zhang, W. Liu, H. D. Ma, and H. Fu, "Siamese neural network based gait recognition for human identification," in *Proceedings of the Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, pp. 2832–2836, IEEE, Shanghai, China, March 2016.

[5] Z. Wu, Y. Huang, L. Wang, X. Wang, and T. Tan, "A comprehensive study on cross-view gait based human identification with deep CNNs," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 2, pp. 209–226, 2017.

[6] Y. Makihara, R. Sagawa, Y. Mukaigawa, T. Echigo, and Y. Yagi, "Gait recognition using a view transformation model in the frequency domain," in *Proceedings of the Proc. European Conference Computer Vision*, pp. 151–163, Graz, Austria, May 2006.

[7] W. Kusakunniran, Q. Wu, J. Zhang, and H. Li, "Gait recognition under various viewing angles based on correlated motion regression," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 6, pp. 966–980, 2012.

[8] W. Kusakunniran, Q. Wu, J. Zhang, and Y. Li, "Support vector regression for multi-view gait recognition based on local motion feature selection," in *Proceedings of the IEEE International Conferene Computer Vision and Pattern Recognition*, pp. 13–18, San Francisco, CA, USA, June 2010.

[9] M. H. Khan, M. Shahid Farid, M. S. Farid, and M. Grzegorzek, "A non-linear view transformations model for cross-view gait recognition," *Neurocomputing*, vol. 402, pp. 100–111, 2020.

[10] K. Bashir, T. Xiang, and S. G. Gong, "Cross-view gait recognition using correlation strength," in *Proceedings of the Conference British Machine Vision*, pp. 1–11, Wales, UK, September 2010.

[11] X. Xing, K. Wang, T. Yan, and Z. Lv, "Complete canonical correlation analysis with application to multi-view gait recognition," *Pattern Recognition*, vol. 50, pp. 107–117, 2016.

[12] W. Kusakunniran, Q. Wu, J. Zhang, H. Li, and L. Wang, "Recognizing gaits across views through correlated motion Co-clustering," *IEEE Transactions on Image Processing*, vol. 23, no. 2, pp. 696–709, 2014.

[13] W. Xu, C. Zhu, and Z. Wang, "Multiview max-margin subspace learning for cross-view gait recognition," *Pattern Recognition Letters*, vol. 107, pp. 75–82, 2018.

[14] T. Connie, K. O. M. Goh, and A. B. J. Teoh, "Multi-view gait recognition using a doubly-kernel approach on the Grassmann manifold," *Neurocomputing*, vol. 216, no. 5, pp. 534–542, 2016.

[15] Y. He and J. Zhang, "Deep learning for gait recognition: a survey," *Pattern Recognition and Artificial Intelligence*, vol. 31, no. 5, pp. 442–452, 2018.

[16] W. Sheng and X. Li, "Siamese denoising autoencoders for joints trajectories reconstruction and robust gait recognition," *Neurocomputing*, vol. 395, pp. 86–94, 2020.

[17] S. Yu, H. Chen, Q. Wang, L. Shen, and Y. Huang, "Invariant feature extraction for gait recognition using only one uniform model," *Neurocomputing*, vol. 239, pp. 81–93, 2017.

[18] S. Gul, M. I. Malik, G. M. Khan, and F. Shafait, "Multi-view gait recognition system using spatio-temporal features and deep learning," *Expert Systems with Applications*, vol. 179, 2021.

[19] X. Li, Y. Makihara, C. Xu, Y. Yagi, and M. Ren, "Gait recognition invariant to carried objects using alpha blending generative adversarial networks," *Pattern Recognition*, vol. 105, 2020.

[20] X. Ding, K. Wang, C. Wang, T. Lan, and L. Liu, "Sequential convolutional network for behavioral pattern extraction in gait recognition," *Neurocomputing*, vol. 463, pp. 411–421, 2021.

[21] H. Wu, J. Weng, X. Chen, and W. Lu, "Feedback weight convolutional neural network for gait recognition," *Journal of Visual Communication and Image Representation*, vol. 55, pp. 424–432, 2018.

[22] A. Zhao, J. Li, and M. Ahmed, "SpiderNet: a spiderweb graph neural network for multi-view gait recognition," *Knowledge-Based Systems*, vol. 206, 2020.

[23] C. Yan, B. L. Zhang, and F. Coenen, "Multi-attributes gait identification by convolutional neural networks," in *Proceedings of the Proc. IEEE Int. Conf. Image and Signal Processing*, pp. 14–16, Shenyang, China, October, 2015.

[24] S. Li, P. Sun, and Y. B. Lang, "Research on gait recognition method in surveillance video based on LSTM," *Journal of People's Public Security University of China (Science and Technology)*, vol. 26, no. 2, pp. 23–28, 2020.

[25] T. Wolf, M. Babaee, and G. Rigoll, "Multi-view gait recognition using 3D convolutional neural networks," in *Proceedings of the Proc. IEEE Int. Conf. Image Processing*, pp. 25–28, Phoenix, AZ, USA, September2016.

[26] D. Thapar, G. Jaswal, A. Nigam, and C. Arora, "Gait metric learning siamese network exploiting dual of spatio-temporal 3D-CNN intra and LSTM based inter gait-cycle-segment features," *Pattern Recognition Letters*, vol. 125, pp. 646–653, 2019.

[27] H. Chao, Y. He, J. Zhang, and J. Feng, "Gaitset: regarding gait as a set for cross-view gait recognition," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, no. 1, pp. 8126–8133, 2019.

[28] C. Fan, Y. Peng, C. Cao et al., "GaitPart: temporal part-based model for gait recognition," in *Proceedings of the Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition (CVPR)*, pp. 13–19, Seattle, WA, USA, June 2020.

[29] S. Hou, C. Cao, X. Liu, and Y. Huang, "Gait lateral network: learning discriminative and compact representations for gait recognition," in *Proceedings of the 16th European Conference on Computer Vision (ECCV 2020)*, pp. 382–398, Glasgow, UK, November 2020.

[30] D. Kastaniotis, I. Theodorakopoulos, C. Theoharatos, G. Economou, and S. Fotopoulos, "A framework for gait-based recognition using Kinect," *Pattern Recognition Letters*, vol. 68, pp. 327–335, 2015.

[31] J. Tang, J. Luo, T. Tjahjadi, and Y. Gao, "2.5D multi-view gait recognition based on point cloud registration," *Sensors*, vol. 14, no. 4, pp. 6124–6143, 2014.

[32] J. Tang, J. Luo, T. Tjahjadi, and F. Guo, "Robust arbitrary-view gait recognition based on 3D partial similarity matching," *IEEE Transactions on Image Processing*, vol. 26, no. 1, pp. 7–22, 2017.

[33] T.-N. Nguyen, H.-H. Huynh, and J. Meunier, "Human gait symmetry assessment using a depth camera and mirrors," *Computers in Biology and Medicine*, vol. 101, pp. 174–183, 2018.

[34] R. Liao, C. Cao, E. B. Garcia, S. Yu, and Y. Huang, "Pose-Based Temporal-Spatial Network(PTSN) for Gait Recognition with Carrying and Clothing Variations," in *Proceedings of the Conference BiometricRecognition*, pp. 474–483, Berlin, Germany, October 2017.

[35] Y. Feng, Y. Li, and J. Luo, "Learning effective gait features using LSTM," in *Proceedings of the Proc. IEEE Int. Conf. pattern recognition (ICPR)*, pp. 4–8, Beijing, China, December 2016.

[36] R. J. Liao, S. Q. Yu, and W. Z. An, "A model-based gait recognition method with body pose and human prior knowledge," *Pattern Recognition*, vol. 98, 2020.

[37] P. Zhang, X. J. Wu, and J. Yang, "A multi-view method for gait recognition based on the length of body sparts," *Acta Automatica Sinica*, vol. 33, no. 2, pp. 210–213, 2007.