Mobile Information Systems

Advanced Human Life and Applications in Digital Twin 2021

Lead Guest Editor: Hoon Ko
Guest Editors: Chun-Wei Tsai and Jongmyung Choi

Advanced Human Life and Applications in Digital Twin 2021

Mobile Information Systems

Advanced Human Life and Applications in Digital Twin 2021

Lead Guest Editor: Hoon Ko
Guest Editors: Chun-Wei Tsai and Jongmyung Choi

Contents

Study on the Application of Digital Information Technology in Music Teaching
Hao Cheng, Hao-ze Zhong, and Ke-cheng Ben
Research Article (9 pages), Article ID 2816848, Volume 2022 (2022)

A Study on the Construction and Utilization of Korean Prehistoric Remains Database
Heesoo Choi, Jong Youl Hong, Sangheon Kim, and Kwangcheol Rim
Research Article (12 pages), Article ID 7081286, Volume 2022 (2022)

Hindawi
Mobile Information Systems
Volume 2023, Article ID 3704217, 12 pages
https://doi.org/10.1155/2023/3704217

Research Article

Deep Learning-Based Business Recommendation System in Intelligent Vehicles

Kyungho Yu,[1] Kiho Lim ,[2] and Pankoo Kim [1]

[1]*Department of Computer Engineering, Chosun University, Gwangju 61452, Republic of Korea*
[2]*Department of Computer Science, William Paterson University of New Jersey, Wayne, NJ 07470, USA*

Correspondence should be addressed to Kiho Lim; limk2@wpunj.edu

Received 17 January 2022; Revised 7 September 2022; Accepted 7 October 2022; Published 14 April 2023

Academic Editor: Floriano Scioscia

The advancements in intelligent vehicle technologies are facilitating the growth of information technology (IT) platforms, unlike conventional automobiles. In-vehicle-infotainment (IVI) is becoming an appealing element in intelligent vehicles as it offers various experiences to users; however, it requires personalized services to provide even more sophisticated user experiences. It is supposed that passengers search for businesses that provide products or services they found interesting in videos played via IVIs while the vehicle is driving autonomously. In that case, it could be more effective to use images that can express the user's preference as a query for the search than to utilize texts such as product names. Accordingly, this study proposes a recommendation system that informs users of businesses near an intelligent vehicle when a passenger inputs an image of a product or service into an IVI system. The proposed recommendation system involves training deep learning-based image classification models with the user's interest images to classify the category, measure the similarity with the business category using Word2vec, and finally provide the locations of the businesses with a high degree of similarity via IVI, using a smartphone. The experimental results indicated that the user's interest image exhibited 85% accuracy for category classification via the EfficientNet B0 model, while the similarity between the image and business categories using Word2vec was particularly high in the business category similar to the actual image category.

1. Introduction

Having the advancements in 5G network technology and artificial intelligence, significant progress has also been made in the Internet-of-things (IoT) technology connecting objects with other objects, including intelligent vehicles, which can provide various services to drivers by connecting them to the Internet. Intelligent vehicles are always connected to the Internet, which facilitates vehicle-to-vehicle and vehicle-to-infrastructure communications. Intelligent vehicles are growing as one of the information technology (IT) platforms on which various content and services can be enjoyed via in-vehicle-infotainment (IVI) systems rather than simple transportation means such as conventional automobiles [1–3].

For intelligent vehicles to appeal to users based on necessity, services conveying convenience and comfort can be provided via IVI systems [2, 3]. Moreover, IVI enables driving and traveling to be more enjoyable in intelligent vehicles with entertainment activities such as listening to music or watching movies. Consequently, IVI has become a critical marketing element for automakers when consumers purchase intelligent vehicles [2]. To offer a variety of experiences to consumers, automakers provide developers with software development kits (SDKs) appropriate for the operating systems (OS) of their vehicles which infotainment-using smartphones are being developed by software companies such as Apple and Google. Studies on IVI services have been conducted on accident prevention based on the driver's heart rate [4], video streaming [5], music recommendation [6], and vehicle maintenance [7]. In addition, automobile and software companies, including academia, have been attempting to provide a wide range of experiences and comfort to users via the IVI systems in intelligent vehicles. In the future, IVIs must focus on

personalization and context awareness to enhance user experience, going beyond providing various experiences to users in intelligent vehicles [3].

In intelligent vehicles capable of level 3 or higher automation driving, passengers can perform secondary tasks using IVIs, rather than driving [8]. Such vehicles can autonomously drive on expressways or parks, thus allowing their drivers or passengers to watch videos, read books, or go on social media with IVIs. While watching videos in an intelligent vehicle that can drive autonomously, passengers may decide to purchase furniture or an item they see in the video, but they may not know the exact name to search as a keyword when searching for nearby stores. Furthermore, while surfing the Internet or going through social media, passengers may find certain meals or hairstyles they like, but they may not know the exact search word for an appropriate restaurant or hair salon. Accordingly, infotainment services are required to provide assistance when passengers are unaware of the exact keyword to use when searching for nearby stores for a service or an item they discover while using IVIs.

The easiest way to find restaurants or stores to purchase an item in an intelligent vehicle is to adopt IVI's navigation system. The navigation system in intelligent vehicles converts the voice into text based on automatic speech recognition, and then the converted text undergoes natural language processing to recommend destination points by identifying the destination and point-of-interest from the user's input text [9]. However, the navigation system of IVIs using text cannot provide sufficient information on a product or a service to a user; hence, it is difficult to navigate to the user's desired destination. As aforementioned, conventional IVI navigation systems cannot be used if passengers do not know the exact name of a product or a service. If the text cannot be used as a query when using the IVI's navigation system, an image-based location search can recommend better destinations than a text-based search because an image can better represent user preferences, which cannot be reflected in texts when an image is used as a query [10]. Therefore, a system is required to recommend nearby businesses by using the image of a product or service as a query when the passenger of an intelligent vehicle decides to utilize the navigation system. Existing deep learning-based recommendation systems have exhibited high performance across several recommendation systems, including product, place, and movie recommendations. However, these systems have certain limitations, requiring multiview information such as user preferences (ratings and clicks) and various attributes (image, description, and reviews). The multiview information integration is a compute-intensive task. In addition, the main challenge in such integration is to retain information relevant to prediction and minimize other irrelevant information, which can reduce a large amount of data, time, and cost requirements when making recommendations. Given the wide availability of image data and the state-of-the-art performance of DL models for classification, we proposed a business location recommendation system that takes only image data of a product or service category desired by a passenger in an intelligent vehicle. The proposed method uses a DL model to identify a product or service category and recommends nearby businesses that fit better into the classified category. Therefore, using image data only to express a user's preference as input and posing the recommendation problem as an image classification task mitigates the limitations mentioned above of existing recommendation systems.

The goal of this study is to propose a recommendation system that provides the user with a peripheral business location suitable for the product category based on the image entered into the IVI system. The process of the proposed IVI business location recommendation system is shown in Figure 1. The proposed recommendation system consists of two components: the user's interest in image classification and the similarity measurement between an input image and business categories. In an Intelligent vehicle, the passenger enters a picture of a store or service into the IVI system that they want to find. The input image goes through a deep learning-based image classification model to predict the input image category. Then, the similarity is measured between the word embedding of the category and the business category of the image. Once the similarity is measured, nearby stores with high similarity are navigated via IVI, using the smartphone.

The remainder of this paper is organized as follows. Related studies on deep-learning based recommendation systems and infotainment for intelligent vehicles are reviewed in Section 2. Then, the proposed image-based business recommendation system is presented in Section 3. Subsequently, the results obtained from experimenting with the collected images and relevant analysis are provided in Section 4. Finally, the conclusion of this study and possible future research directions are provided in Section 5.

2. Related Works

2.1. Infotainment for Intelligent Vehicles. Infotainment is a compound word for information and entertainment. IVI is a significant factor for consumers to purchase intelligent vehicles [2]. Existing car infotainment is divided into methods that either use the service built into the car embedded OS or connect a smartphone to the vehicle. In order to provide more services to users through IVI, an automaker that makes intelligent vehicles is putting a lot of effort into creating applications by providing SDK (soft development kit) suitable for the company's OS to developers. Infotainment using smartphones is led by software companies such as Apple-Apple CarPlay [11] and Google-Android Auto [12].

Existing research on IVI can be grouped into research on convenient operation and infotainment applications in using infotainment in intelligent vehicles. As an example of a convenient process, the authors in [13] have proposed a text-based infotainment system that takes voice command as input and uses Kaldi and voice recognition tools to change the command into text for running a service. In [14, 15], the authors have proposed an alternative infotainment system in which a driver can control the system by using gestures. The system is powered by a computer vision-based gesture recognition system. In research on infotainment applications, there is a study on using a wearable device to check

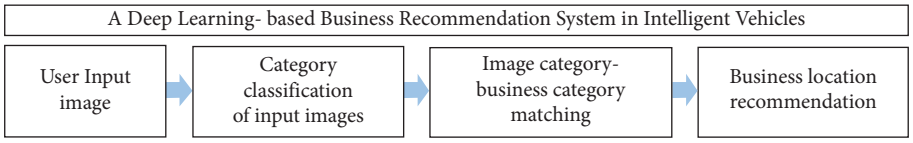


Figure 1: Process of the proposed business-location recommendation system.

heart rate and monitoring through infotainment to prevent a vehicle crash due to a stroke or heart attack while driving [4], and smooth multimedia streaming in an intelligent vehicle in operation. A study [5] protects the appropriate data rate and frames for this purpose. And, there is a study that recommends music by inputting the driver's profile, current situation, and personal preference as inputs [6]. The vehicle maintenance system "e-talk" can check the vehicle's condition through infotainment using the IoT sensors attached to the vehicle [7].

2.2. Recommendation System Based on Deep Learning. Recently, research is actively being conducted on applying deep learning (DL) techniques to conventional recommendation systems such as collaborative or content-based filtering [16]. For DL-based collaborative filtering, studies have been conducted on giving recommendations to a user by combining two different latent vectors after inputting information (including the descriptions, reviews, and ratings of a product or a place for embedding) of another user in the network who has similar preferences as that user [17, 18]. Information on a product or place is expressed as latent factors of the same dimension through a trained embedding layer and training in a neural network. The main advantage of such representation is that more complicated interactions can be captured in the categorical data. Similarly, in another study, product information such as images, descriptions, and review text were input into an auto-encoder to extract its features, and then the extracted features were integrated with the last layer to make recommendations to users [19]. The advantages of a DL-based recommendation system are that features can be extracted using the user's preferences (review, rating, and clicks) and various attributes (image, description, and review) from a large amount of data. In addition, detailed recommendations can be provided via the complicated interactions between the extracted features [20].

Existing deep learning-based recommendation systems have exhibited high performance across several recommendation systems, including product, place, and movie recommendations; however, these systems are posed with certain limitations as they require multiview information. The main challenge is to retain information relevant to prediction and minimize other irrelevant information. The existing systems require a large amount of data, time, and cost to make detailed recommendations [21]. Therefore, unlike existing recommendation systems that employ large-scale data and complex deep learning models, the proposed business location recommendation system classifies the image category of a product or service desired by a passenger in an intelligent vehicle using a deep learning model and recommends nearby businesses that fit into the classified category.

3. Methods

3.1. Businesses Recommendation Systems in Intelligent Vehicles. This section describes the overall process of the system for recommending businesses near a vehicle when a passenger inputs an image of interest for a product or service into IVI within intelligent vehicles, as illustrated in Figure 2. First, a user's interest category is selected to speculate the product or service image the user may be interested in. The range of individuals' interest products or services is substantially wide; hence, local business data provided by Yelp were adopted to limit the range of the user's interest categories. Each store in Yelp's local business data comprises multiple categories, and among them, a category is selected as the user's interest category, provided the frequency of appearance is high. The dataset is constructed by collecting images with the selected category as the web search keyword. Subsequently, the category classification of interest images is performed by training a deep learning-based image classification model with the user's interest image dataset. To recommend appropriate businesses to users, the image category and business data need to be matched. The similarity between image and business categories is mapped in the same dimension using Word2vec and then measured accordingly. Once the similarity is measured between the two categories, the location of a business near the intelligent vehicle boarded by the passenger is provided to users via IVI, using a smartphone.

In Sections 3.2 and 3.3, image classification via deep learning and text similarity measurement using Word2vec, which are crucial parts of the proposed recommendation system, is explained.

3.2. Image Classification Model for Business Category. Deep learning models were adopted in this study to extract visual information from the image input by a user into the infotainment system of intelligent vehicles and to classify business categories. Using the ImageNet dataset, deep learning models for image classification were trained by applying transfer learning of the collected user's interest category data to pretrained ResNet [21], Inception v3 [22], and EfficientNet B0 [23]. The architecture of each model is illustrated in Figure 3. Figure 3(a) presents the architecture of ResNet. Generally, it is expected that the performance of a deep learning model improves as the number of hidden layers in the model increases. However, deeper DL models suffer from vanishing gradient problems, and the use of pooling layers puts restrictions on the depth of the model. ResNet adopts a skip connection where the layer input is directly connected to the layer output to address such problems (Figure 3(a)). Using skip connection address the

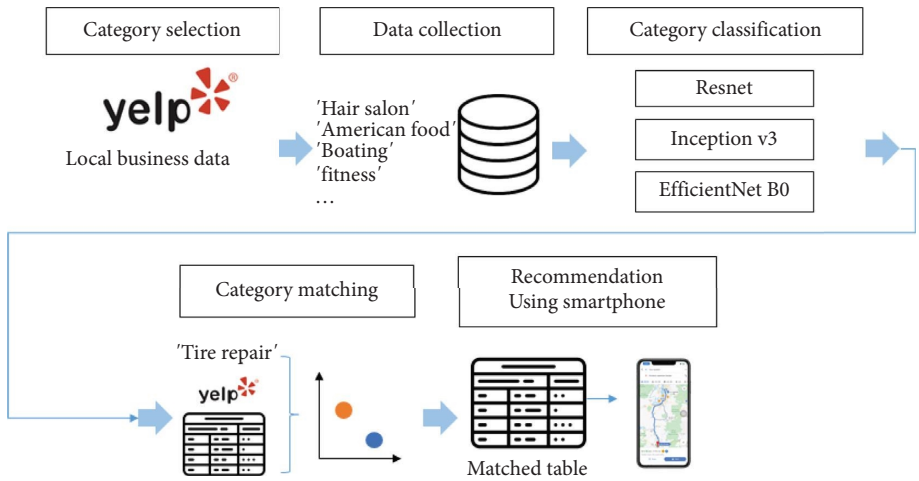


Figure 2: Business location recommendation process.

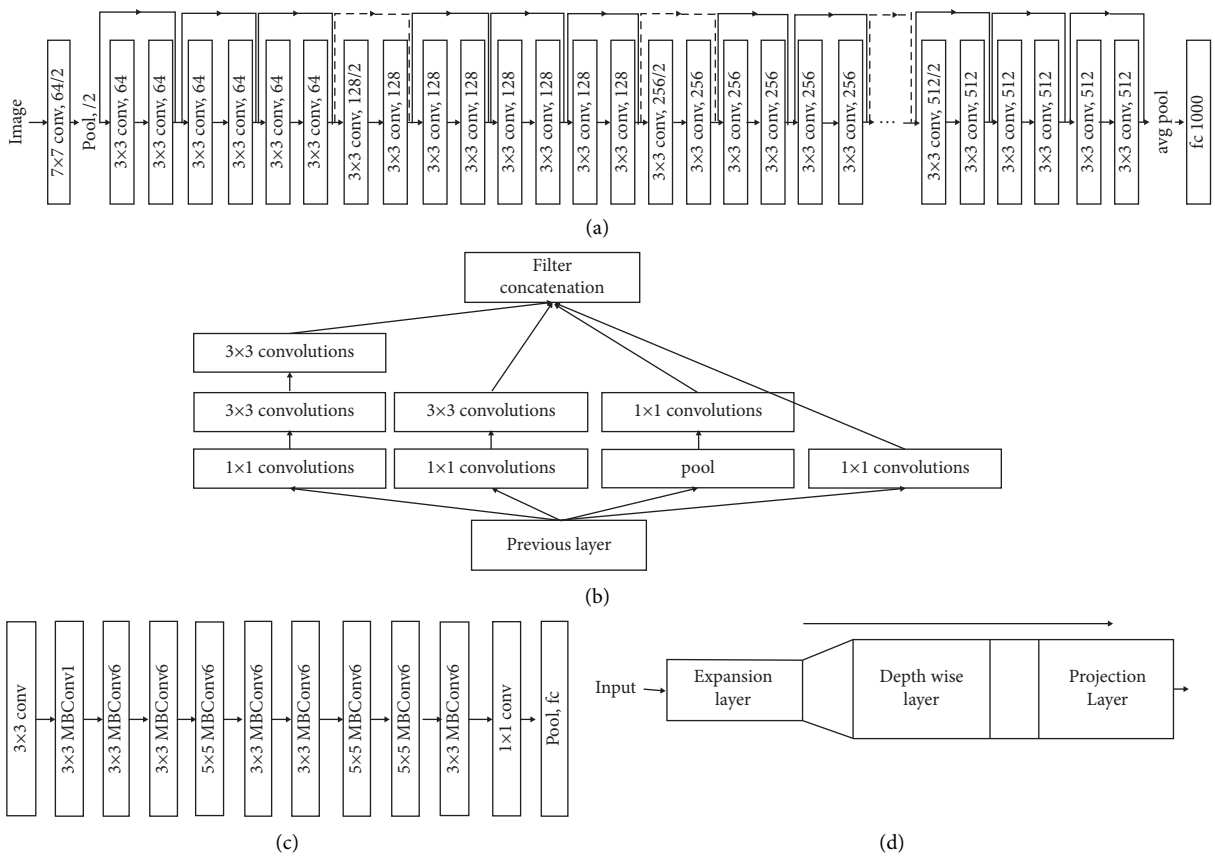


Figure 3: Architecture of (a) ResNet, (b) Inception v3 module, (c) EfficientNet B0, and (d) MBConv.

performance degradation challenges even when layers get deeper as the neural network learns the difference between input and output values. Figure 3(b) illustrates the module of the Inception v3 model. Inception v3 is a convolution neural network comprising 48 layers, where the model is configured by stacking multiple inception modules with the inception model presented in Figure 3(b) as the base. One module is configured using convolution filters of different sizes, unlike AlexNet or VGGNet models, which utilize the filter of the same size in one layer for convolution. More diverse features can be extracted from the input image because the filters with different convolution sizes are adopted, and the computation amount can be reduced in the inception model by performing 1×1 convolution to reduce the number of parameters to be delivered to the convolution filter of a large size. For EfficientNet, a scale-up method required for improving the performance of the existing ConvNet, was investigated. Experiments were conducted on three scale-up

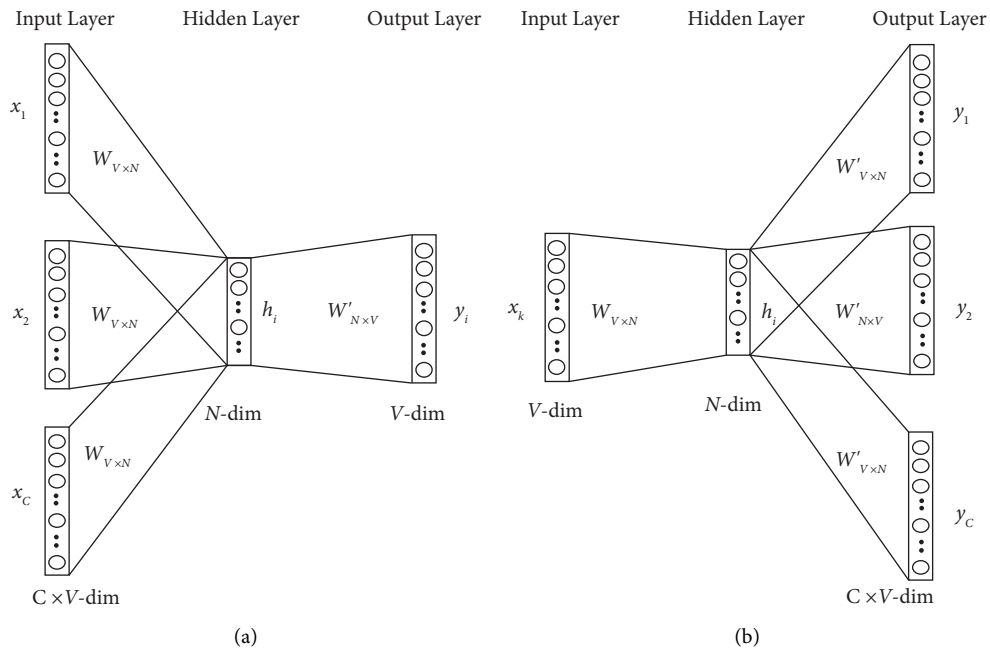


Figure 4: Architecture of Word2vec. (a) CBOW. (b) Skip-gram.

methods, which include increasing the depth of a neural network, increasing the channel width, and increasing the resolution of an image for training. In addition, the EfficientNet B0 model was created using the MBConv (Figure 3(d)) of MobileNet. MBconv is a mobile inverted bottleneck convolution adopted in MobileNet. This block expands the low-dimensional input to a 1×1 layer, expands it again into 3×3 convolution, and the outputs using the 1×1 convolution layer. This configuration is adopted because more features can be extracted by converting from the low- to high-dimensional data and using the projection layer while reducing the computation amount. EfficientNet exhibited high performance in image classification by adopting scale-up methods and MBconv. In this study, the user's interest data collected via web crawling were used to train ResNet, Inception v3, and EfficientNet B0 models to identify an appropriate model for the collected data.

3.3. Similarity Measurement Using Word2vec. Word embedding refers to converting text into numbers via distributed representation to enable a computer to understand human language. In other words, words comprising the text are mapped to real number vectors. Word2vec can convert text into vectors by identifying the relationship between words in a sentence and can perform operations such as addition or subtraction between words using the converted vectors. Figure 4 presents the architecture of Word2vec, which inputs the text into a thin neural network for training. The training methods of Word2vec include a continuous bag of words (CBOW), which predicts the center word based on the surrounding words, and the Skip-gram, which predicts surrounding words based on the center word. When compared, the Skip-gram model exhibited higher semantic and syntactic accuracies than the CBOW model. The similarity was measured using the word embedded vectors of image and business categories because categories other than the business categories are ignored if the image category simply determines whether it is included in the business category when matching the categories. For example, if the image input by the user is classified as "American food," the categories of restaurant A are "wine bars, restaurants, food, beer, seafood, and steak houses." The similarity between these two has no correlation if the text is simply matched with the words "American food;" however, the degree of similarity can be identified if Word2vec, which learns the relationship between words, is adopted to measure the similarity. As the business category includes multiple categories, the mean value of the categories is adopted to measure the similarity by comparing it with the embedding value of the image category.

4. Experiment

4.1. Dataset. As described in the previous section, images of user interest were collected using the frequency of business categories in Yelp business data by selecting the category that appears most frequently and points to the details and using it as a keyword of a web search engine. Figure 5 presents the frequency of business categories in the Yelp data. The categories with the highest frequencies are restaurants, food, and shopping. However, these categories were not selected as the user's interest category because the range of products or services is wide; instead, categories such as "American food," "boating," and "fitness" were selected as the user's interest category and adopted as web crawling keywords to build the user's interest image dataset. The number of categories in the dataset is 10, and

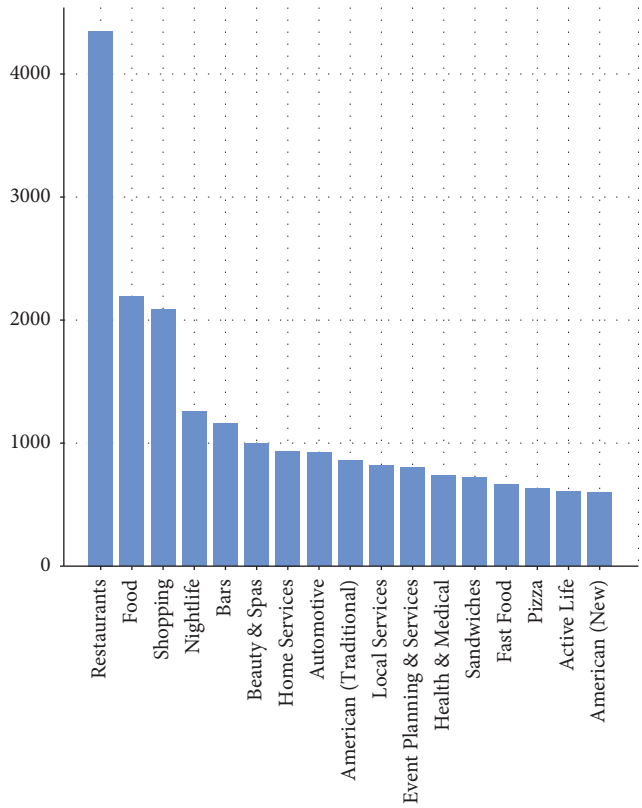


Figure 5: Frequency of business categories in Yelp data.

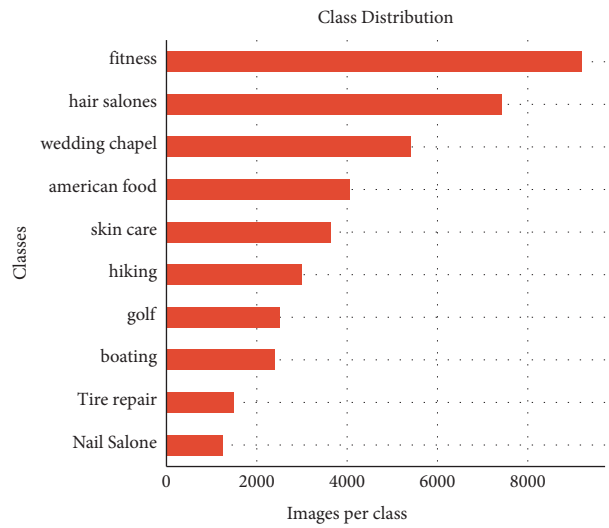


Figure 6: Data distribution.

the number of images in each category is presented in Figure 6. The original images were augmented twice to create 40,540 images in the dataset, in which the images were divided at a ratio of 6:3:1 for train, validation, and test data to train the deep learning models.

Table 1 presents Yelp business data, which is the business information data. The Yelp business dataset adopted in this study is a dataset comprising store categories and user reviews, and the names and categories of the businesses, as well as location information such as state, latitude, and longitude were used in this study.

5. Results

5.1. Training Accuracy and Loss of Image Classification. The computer used for the experiment was equipped with Intel i7-9750H and NVIDIA GeForce RTX 2080 Ti model. In total, 20 epochs were applied for the experiment of all three models, 50 layers were fixed and trained for the ResNet and Inception v3 models, while the entire layers were all trained for the EfficientNet B0 model. Figure 7 presents the results for training ResNet, Inception v3, and EfficientNet B0 with 10 category images. As illustrated in Figures 7(a) and 7(c), the ResNet and Inception v3 models exhibited improved accuracies in the training data; however, the accuracy in the validation data did not improve after 70%, as the number of training increased. Figure 7(e) presents the accuracy of the EfficientNet B0 model. Unlike the ResNet and Inception v3 models, the EfficientNet B0 model gradually improved the accuracy in both training and validation data as training proceeded. Therefore, the EfficientNet B0 model can be considered appropriate for image classification in this study. Figure 8 illustrates the comparison among accuracies of the three models when trained for 20 epochs. At 20 epochs, the ResNet model exhibited 96.71% and 75.74% accuracies in the training and validation data, respectively. The Inception v3 model exhibited 95.32% and 74.3% accuracies in the training and validation data, respectively. In addition, the EfficientNet B0 model exhibited 88.44% and 85.31% accuracies in the training and validation data, respectively.

5.2. Confusion Matrix of Image Classification. A confusion matrix is a measure used to evaluate the performance of deep learning classification. In binary classification, a confusion matrix represents the instances of true positive, true negative, false positive, and false negative. In general, higher true positive and true negative values and lower false positive and false negative values indicate good performance. A confusion matrix in multiclass classification is an excellent performance indicator, as the prediction of each class can be identified. Figure 9 presents the confusion matrix created based on the prediction results of the ResNet, Inception v3, and EfficientNet B0 models after training, where the maximum value is 1. The confusion matrix in Figure 10 demonstrates that the value in the diagonal direction is true positive, where the accuracy is higher as the value is closer to 1. When the test data were applied to each model, the ResNet, Inception v3, and EfficientNet B0 models exhibited 76%, 74%, and 85% accuracies, respectively. According to the confusion matrix, the prediction accuracy was low in the category related to nail salons, which can be improved if more data are collected and refined.

5.3. Category Matching Using Word2vec. Although it is important to accurately classify a user's input image into the proper category, it is also crucial to select candidates for business recommendations for users based on the classification

Table 1: Example of Yelp dataset.

Names	Categories	State	City	Address	Latitude	Longitude
Middle East Deli	Food, restaurants, grocery	NC	Charlotte	4508E Independence Blvd	35.194894	−80.767442
LA fitness	Yoga, fitness and instruction, gyms	AZ	Phoenix	7640W Thomas Rd	33.481551	−112.223225
Bank of America	Banks, financial services	NC	Charlotte	9715 Callabridge Ct	35.3275006	−80.9440573
Starbucks	Food, Coffee, and tea	AZ	Phoenix	1610N. 75Th Avenue, #100	33.466005	−112.221288

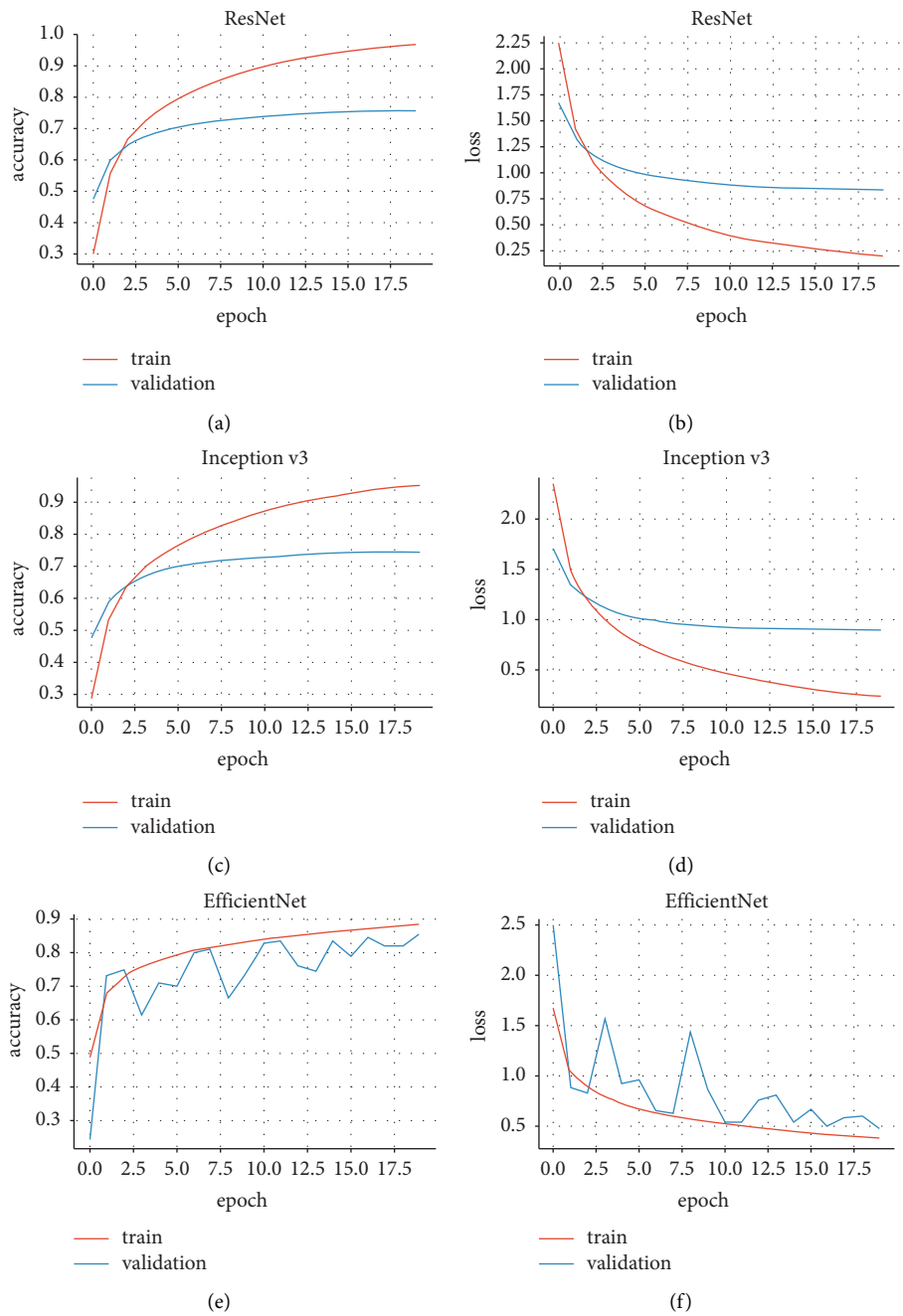


Figure 7: Training accuracy and loss of each model. (a) ResNet train accuracy. (b) ResNet train loss. (c) Inception v3 train accuracy. (d) Inception v3train loss. (e) EfficientNet B0 train accuracy. (f) EfficientNet B0 train loss.

results. Therefore, in this section, multiple business categories were applied using Word2vec when measuring the similarity between image and business categories.

Figure 10 presents the method for measuring the similarity between image and business categories using Word2vec. Word2vec learns the relationship between

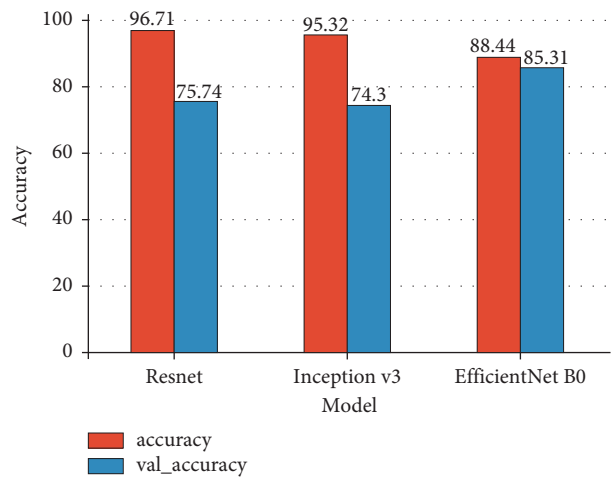


Figure 8: Accuracy of each model.

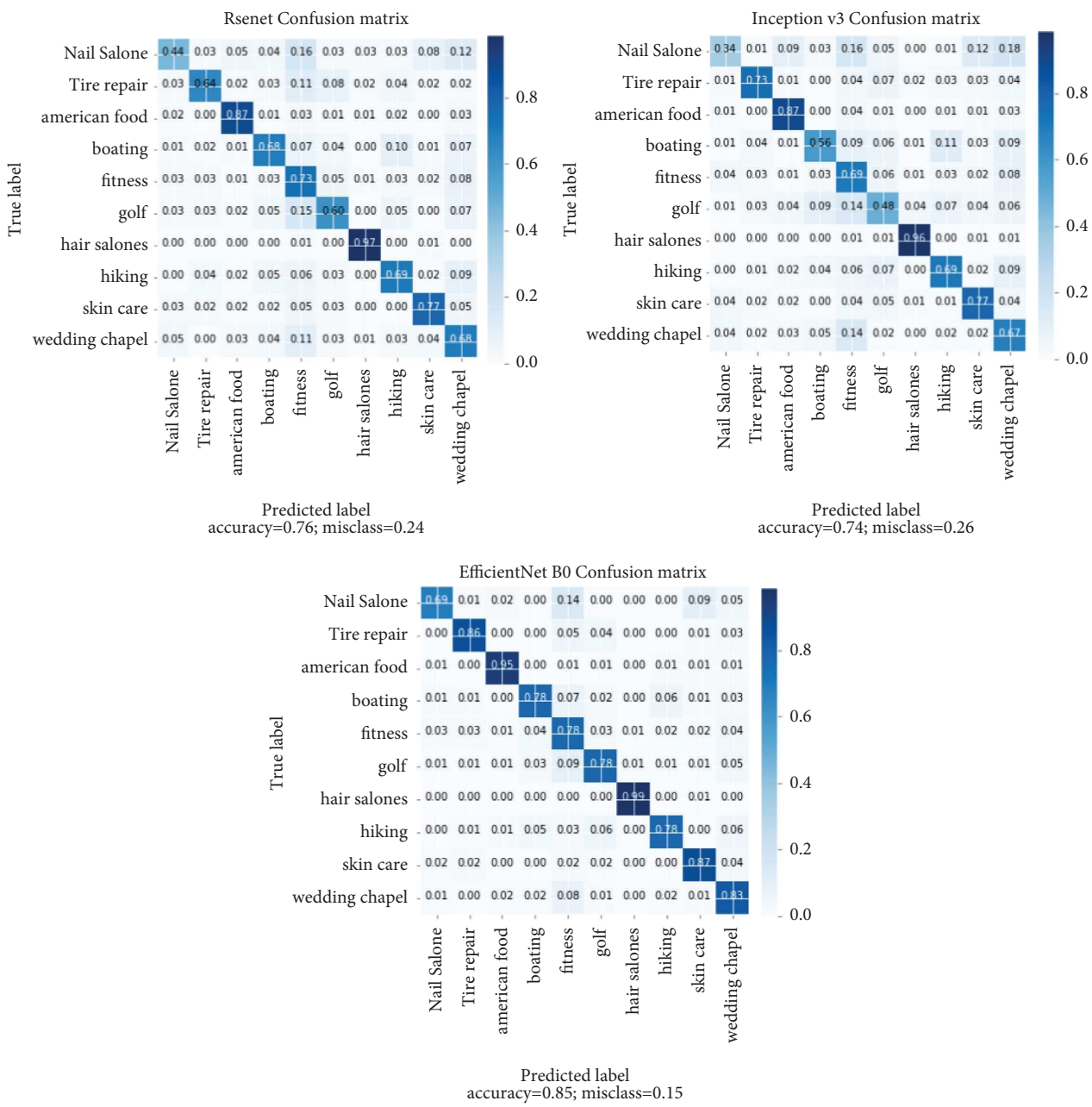


Figure 9: Confusion matrix of each model.

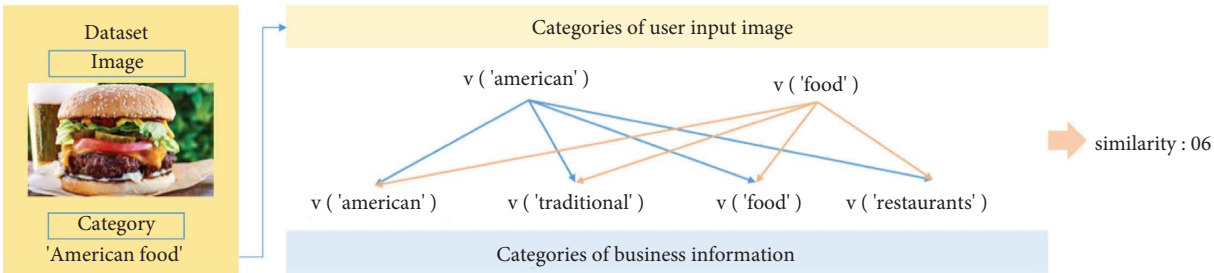


Figure 10: Measurement similarity between images and business categories using Word2vec.

Table 2: Measurement similarity between "American food" and other business categories.

Image categories	Business categories	Similarity
	"American," "traditional," "food," "Cajun," "creole," "fast," "food," "restaurants," "chicken," "wings"	0.6
	"Paddleboarding," "boating," "rafting," "kayaking," "jet," "skis," "active," "life"	0.14
	"Active," "life," "trainers," "gyms," "fitness," "instruction"	0.15
	"Professional," "services," "local," "services," "carpet," "cleaning"	0.24
	"Parks," "active," "life," "hiking"	0.14
"American," "food"	"Eyelash," "service," "beauty," "spas," "nail,, "salons"	0.17
	"Eyebrow,, "services," "hair,, "salons," "beauty," "spas"	0.18
	"Automotive," "tires,, "oil," "change," "stations," "transmission," "repair," "auto," "repair"	0.19
	"Wedding," "planning," "Italian," "caterers," "party," "event," "planning," "restaurants," "venues," "event," "spaces," "event,, "planning," "services"	0.19
	"Country," "clubs," "arts," "entertainment," "golf," "active," "life"	0.25

Table 3: Measurement similarity between "hair salon" and other business categories.

Image categories	Business categories	Similarity
	"American," "traditional," "food,","Cajun," "creole," "fast," "food," "restaurants," "chicken," "wings"	0.17
	"Paddleboarding," "boating," "rafting," "kayaking," "jet,", "skis," "active," "life"	0.13
	"Active," "life," "trainers," "gyms," "fitness," "instruction"	0.13
	"Professional," "services," "local," "services," "carpet," "cleaning"	0.19
	"Parks," "active," "life," "hiking"	0.06
"Hair," "salon"	"Eyelash," "service," "beauty," "spas," "nail," "salons"	0.35
	"Eyebrow," "services," "hair," "salons," "beauty," "spas"	0.48
	"Automotive," "tires," "oil," "change," "stations," "transmission," "repair," "auto," "repair"	0.07
	"Wedding," "planning," "Italian," "caterers," "party," "event," "planning," "restaurants," "venues," "event," "spaces," "event," "planning," "services"	0.16
	"Country," "clubs," "arts," "entertainment," "golf," "active," "life"	0.13

words and converts words into numbers. Therefore, the "tire repair" category was separated into "tire" and "repair" via preprocessing, while the business category was also divided into individual words for word embedding. The similarity comparison between image and business categories is the N : N comparison when categories are divided into words; therefore, the average cosine similarity between the image and business-category words was adopted as the similarity between the respective image and category. A model pretrained with Google news data were adopted as the Word2vec model, and the data of the state of Ohio in the Yelp local business data were utilized. Tables 2 and 3 present a part of the similarity measurement results.

Tables 2 and 3 present the results obtained from measuring the similarity between image classification categories "American food" and "hair salon" with other business categories, respectively. In Table 2, the businesses in the first row are most similar to "American food," while low scores were recorded in other business types. In Table 3, the businesses in the seventh row are the most similar business category to "hair salon." The business categories in the sixth row exhibited the second highest similarity, which is also illustrated in Figure 11 where "hair," "nail," and "skin" have similar values after word embedding, possible owing to the presence of a similar word "beauty" in the category. In the results obtained from measuring the similarity between image and business categories using Word2vec, the business

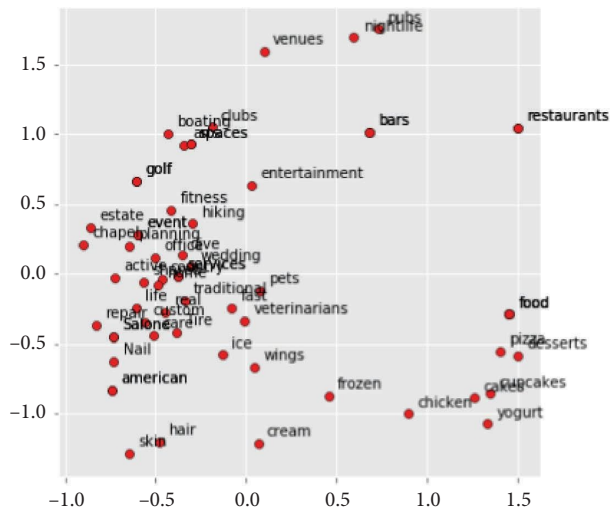


Figure 11: Word embedding of categories using Word2vec.

Figure 12: A mobile app prototype of the business location recommendation system.

categories with visibly high similarity also exhibited high similarity degrees in the results obtained from using Word2vec; hence, businesses other than the user's input image can be recommended because the similarity of each business category can be identified.

5.4. Mobile App Prototype. Figure 12 presents a mobile app prototype of the proposed system for recommending nearby business locations of a product or service found in videos, Internet, or social media when a passenger uses IVI for leisure with a smartphone. The business recommendation system proceeds from left to right in Figure 12. When an image of the food desired by a user is input, the candidates of the category for the image are output below the image. When the user selects a certain category, nearby business locations are displayed on the screen based on the user's location information. In addition, when the user selects a certain business, the closest location is displayed on the screen. If the technology related to intelligent vehicles is further advanced, it can be applied to the infotainment system, which adopts the OS of intelligent vehicles, or a business recommendation system via a head-up display to which applied augmented reality can be provided.

6. Conclusions

In this paper, we proposed a business recommendation system based on image data as opposed to the existing methods that make use of multiview information integration. The proposed recommendation system consists of two stages: the user's interest image classification and the measurement of similarity between an input image and business categories. To classify the user's interest image,

relevant images were collected via web crawling. Then, in the first stage, different deep learning-based image classification models were trained and tested on the acquired data. The results indicated that the EfficientNet B0 model exhibited better performance on the test data than its counterparts. In the second stage, the similarity between image and business categories was measured using Word2vec and converted to vectors; then, the cosine similarity was measured between each word to adopt the sum of the average values as the degree of similarity. Consequently, the business category that was visibly similar to the image category also exhibited a high degree of similarity based on Word2vec. In addition, the proposed recommendation system can be used for different scenarios, such as, when providing recommendation services to passengers and drivers searching for nearby businesses using the intelligent vehicle infotainment system. In the future, we plan to improve the accuracy of the classifier in the early steps as it may improve the overall performance of the proposed recommendation system and compare it with other classification models.

Data Availability

YELP dataset used to support the findings of this study are available at https://www.yelp.com/dataset.

Conflicts of Interest

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

Acknowledgments

This research was supported by the MSIT (Ministry of Science, ICT), Korea, under the High-Potential Individuals Global Training Program (Grant no. 2020-0-01578) supervised by the IITP (Institute for Information & Communications Technology Planning & Evaluation). This work was also supported by the National Research Foundation of Korea (NRF) grant funded by the Korea Government (MSIT) (Grantn no. NRF-2020R1A2C2007091).

References

[1] J. E. Siegel, D. C. Erb, and S. E. Sarma, "A survey of the connected vehicle landscape—architectures, enabling technologies, applications, and development areas," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 8, pp. 2391–2406, 2018.

[2] R. Coppola and M. Morisio, "Connected car: technologies, issues, future trends," *ACM Computing Surveys*, vol. 49, no. 3, pp. 1–36, 2016.

[3] J. Guo, B. Song, Y. He, F. R. Yu, and M. Sookhak, "A survey on compressed sensing in vehicular infotainment systems," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 4, pp. 2662–2680, 2017.

[4] G. Sinnapolu and S. M. Alawneh, "Intelligent wearable heart rate sensor implementation for in-vehicle infotainment and assistance," *Internet of Things*, vol. 12, Article ID 100277, 2020.

[5] L. Wang, H. Yang, X. Qi, J. Xu, and K. Wu, "ICast: fine-grained wireless video streaming over Internet of intelligent vehicles," *IEEE Internet of Things Journal*, vol. 6, no. 1, pp. 111–123, 2019.

[6] G. J. Dimitrakopoulos and I. E. Panagiotopoulos, "In-vehicle infotainment systems: using bayesian networks to model cognitive selection of music genres," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 11, pp. 6900–6909, 2021.

[7] S. Hussain, U. Mahmud, and S. Yang, "Car e-talk: an IoT-enabled cloud-assisted smart fleet maintenance system," *IEEE Internet of Things Journal*, vol. 8, no. 12, pp. 9484–9494, 2021.

[8] A. L. Kun, S. Boll, and A. Schmidt, "Shifting gears: user interfaces in the age of autonomous driving," *IEEE Pervasive Computing*, vol. 15, no. 1, pp. 32–38, 2016.

[9] Y. Zheng, Y. Liu, and J. H. Hansen, "Intent detection and semantic parsing for navigation dialogue language processing," in *Proceedings of the 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, pp. 1–6, Yokohama, Japan, October, 2017.

[10] C. Xu, Z. Guan, W. Zhao et al., "Recommendation by users' multi-modal preferences for smart city applications," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 6, pp. 4197–4205, 2021.

[11] Apple, "Apple CarPlay," http://www.apple.com/ios/carplay (accessed on 2021).12. Google, "Android AU", https://www.android.com/%20auto(accessed on 2021).

[12] Google, "Android AU," https://www.android.com/%20auto/.

[13] T. Ranzenberger, C. Hacker, and F. Gallwitz, "Integration of a Kaldi speech recognizer into a speech dialog system for automotive infotainment applications," *Studientexte zur Sprachkommunikation: Elektronische Sprachsignalverarbeitung*, vol. 2018, pp. 293–300, 2018.

[14] F. Parada-Loira, E. González-Agulla, and J. L. Alba-Castro, "Hand gestures to control infotainment equipment in cars," in *Proceedings of the 2014 IEEE Intelligent Vehicles Symposium*, pp. 1–6, Dearborn, MI, USA, June, 2014.

[15] B. Verma and A. Choudhary, "Framework for dynamic hand gesture recognition using Grassmann manifold for intelligent vehicles," *IET Intelligent Transport Systems*, vol. 12, no. 7, pp. 721–729, 2018.

[16] J. Bobadilla, S. Alonso, and A. Hernando, "Deep learning architecture for collaborative filtering recommender systems," *Applied Sciences*, vol. 10, no. 7, p. 2441, 2020.

[17] B. Zhang, M. Zhu, M. Yu, D. Pu, and G. Feng, "Extreme residual connected convolution-based collaborative filtering for document context-aware rating prediction," *IEEE Access*, vol. 8, pp. 53604–53613, 2020.

[18] Y. Guan, Q. Wei, and G. Chen, "Deep learning based personalized recommendation with multi-view information integration," *Decision Support Systems*, vol. 118, pp. 58–69, 2019.

[19] G. Gupta and R. Katarya, "Research on understanding the effect of deep learning on user preferences," *Arabian Journal for Science and Engineering*, vol. 46, no. 4, pp. 3247–3286, 2020.

[20] S. Zhang, L. Yao, A. Sun, and Y. Tay, "Deep learning based recommender system: a survey and new perspectives," *ACM Computing Surveys*, vol. 52, no. 1, pp. 1–38, 2019.

[21] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, Las Vegas, NV, USA, June, 2016.

[22] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2818–2826, Las Vegas, NV, USA, June, 2016.

[23] M. Tan and Q. Le, "Efficientnet: rethinking model scaling for convolutional neural networks," in *Proceedings of the International Conference on Machine Learning PMLR*, Baltimore, MD, USA, July, 2019.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 7025597, 11 pages
https://doi.org/10.1155/2022/7025597

Research Article

Real-Time Controllable Optimization Algorithm for Correlated Big Data in Cloud Computing Environment

Rutao Li[1] and Zaiyi Pu[2]

[1]*School of Electronic Information Engineering, Ningbo Polytechnic, Ningbo 315800, Zhejiang, China*
[2]*Education and Information Technology Center, China West Normal University, Nanchong 637009, Sichuan, China*

Correspondence should be addressed to Zaiyi Pu; zaiyipu@cwnu.edu.cn

Received 10 February 2022; Accepted 28 April 2022; Published 31 May 2022

Academic Editor: Jong M. Choi

The era of big data (BD) has arrived. How to train models to find correlations in data and help people make decisions has become a major research topic and direction. As an elastic and scalable distributed computing mode (refers to a whole consisting of multiple interconnected computers that cooperate to perform a common or different task in a set of system software environments, with minimal reliance on centralized control processes, data, and hardware), cloud computing can provide powerful computing and storage capabilities and has been widely used in BD query and difficult processing. This paper aims to study the algorithm in the environment of cloud computing. Different from the traditional research algorithms, the relational BD algorithm is controllable in real time. Moreover, it has been optimized and upgraded on the previous real-time controllable algorithm. Also, it performs serial and parallel simulation tests on the algorithm. When the optimal situation of the parallel algorithm is obtained, the test results show that the relevant mining time of the optimized algorithm is significantly shorter than the traditional data mining time under the same dataset. The traditional mining time is about 3.5 times the data mining time of this paper, and the running power consumption of the optimization algorithm is reduced to 20 W.

1. Introduction

Computers and the Internet have accelerated the change and dissemination of information. The advent of the era of information explosion has also promoted the transformation of storage technology from optical disks, chips, and card storage to disk arrays and even today's large-scale network disk array storage. The rapid development of data storage and data processing methods has accelerated the arrival of BD. Generally speaking, BD are massive data with complex structure, huge quantity, diverse types, and low value density. Their scale is too large to be understood. From a computer science perspective, BD refer to a collection of various structured and semistructured data. From a popular point of view, BD refer to a massive collection of data. Unlike traditional data, which tend to discover why things happen, BD have the advantage of predicting what will happen. Finding causal relationships is not helpful for users to make decisions, and finding correlations between things is more valuable. For example, people place more importance on when the airfare is cheaper than why the price of the plane changes. The relationship implied by BD is more worthy of social exploration.

BD's relational data topic mining refers to mining the implicit relationship of each data item in the database. Mining frequent itemsets (the most basic pattern is an itemset, which refers to a collection of several items; frequent patterns refer to itemsets, sequences, or substructures that frequently appear in a dataset; frequent itemsets refer to sets with support greater than or equal to the minimum support) of transaction sets in database is an important part and main goal of association rules. The relationship of each data item in the database can be clearly expressed by mining association rules, which is easy for people to understand. Using cloud computing to mine the relevance of BD is an efficient and low-cost method. Static algorithms can only process

some conventional data. For some places that need emergency processing and high risks, such as emergency fire monitoring, cabin gas monitoring, and so on, static algorithms are obviously not enough to meet the requirements. Although the current dynamic algorithms can control the data in real time, these algorithms need to be further optimized and upgraded in order to calculate the data more accurately and even with the smallest error.

The innovation of this paper lies in the use of dynamic real-time controllable data algorithms. The algorithm has stronger applicability and wider range of use, and the data can be seen at any time. In addition, this paper optimizes on the basis of the previous real-time controllable algorithm, which speeds up the data detection speed, makes the data update faster, and makes the calculation results more accurate.

2. Related Work

With the research and development of BD such as the Internet of Things, the Internet, and medical care, more and more scholars have entered the study of BD. Among them, Puig et al. applied data analysis to aid in fraud detection and maintenance demand forecasting. Puig et al. introduced new algorithms and methods [1]. Qian et al. collected real-time observations of hourly mean wave height, temperature, and pressure at the Maidao station in Qingdao, China. Using eight quality control methods, they explored data quality and identified the most effective methods for the Wheat Road station. After using eight quality control methods, the average wave height, temperature, and pressure data passed the test in percentages of 89.6%, 88.3%, and 98.6%, respectively [2]. Puthal et al. called this an online security verification problem. To solve this problem, Puthal et al. proposed a dynamic key length based security framework (DLSeF). The theoretical analysis and experimental results of the DLSeF framework show that the efficiency of processing streaming data can be significantly improved [3]. To address these challenges, Ren et al. proposed a workshop material delivery framework. They studied the key technologies of the proposed framework. To demonstrate the implementation of the proposed framework, proof-of-concept scenarios were designed to demonstrate the implementation of the framework [4]. Bao et al. proposed a novel gamma control method to tune the vertical growth rate, a method for estimating plasma vertical instability. The experimental results show that the time evolution of the real-time vertical growth rate conforms to the target value, and the real-time vertical growth rate can be adjusted by gamma control [5]. Janzen and Mann introduced a feedback control method that automatically adjusts multiple exposure settings for compositing to increase the dynamic range of sensory processes. It is synthesized to capture an extremely high dynamic range with minimal uncertainty [6]. The research of the above scholars can promote the development of big data in certain aspects, but the research on the correlation and real-time control of big data is not in-depth and perfect and needs to be further optimized.

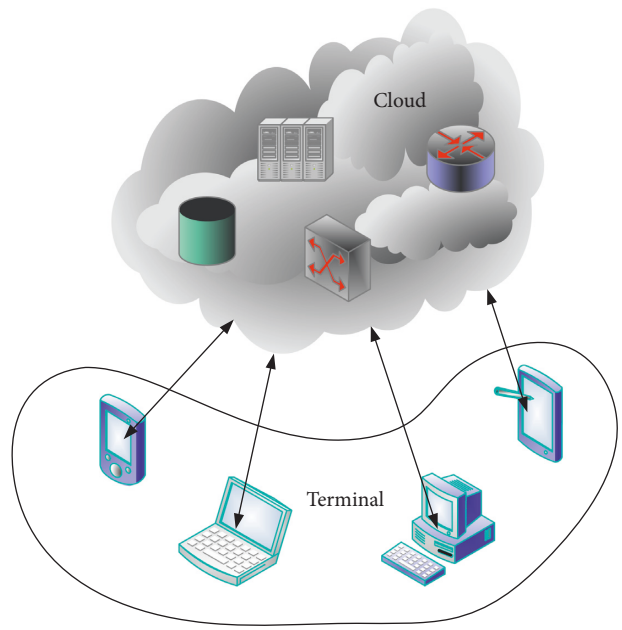


Figure 1: The principle of cloud computing.

3. Data Controllable Algorithm Based on Cloud Computing

3.1. Cloud Computing

(1) *Origin*. In August 2006, Google CEO first proposed the concept of "cloud computing" at the SES San Jose 2006 conference. After this concept was formally proposed in the history of cloud computing, it caused a great sensation [7].

(2) *Definition*. The definition of cloud computing is not yet unified. According to NIST's improvement, cloud computing is a model that can obtain the required resources from a shared pool of configurable computing resources anytime, anywhere, conveniently, and on demand, and resources can be quickly provided and released [8, 9]. It minimizes the effort of managing resources or interaction with service providers [10, 11]. The principle of cloud computing is shown in Figure 1.

(3) *Features*. The main features of cloud computing are as follows. First, without manpower, users can utilize the existing computing resources of any service provider. It includes processing power, storage space, or applications without human interaction. Second, it is convenient and fast, and the resources in the cloud computing system can be accessed and used anytime, anywhere. The third is information sharing. The computing resources of service providers can be combined to provide services, and these combined resources may be distributed in multiple data centers around the world. At the same time, the provider's computing resources can be shared and used by multiple users. The fourth is resource payment. Users can use these computing resources flexibly. A user can

Table 1: Types of cloud computing services.

Genre	Concept	The proportion of applications (%)
IaaS	It provides virtualized computing resources to individuals or organizations of cloud computing providers	6.03
PaaS	Provides a platform for developers to build applications and services over the global Internet.	25.57
SaaS	Providing pay-as-you-go software applications over the Internet; cloud computing providers host and manage software applications and allow their users to connect to and access applications over the global Internet.	67.4

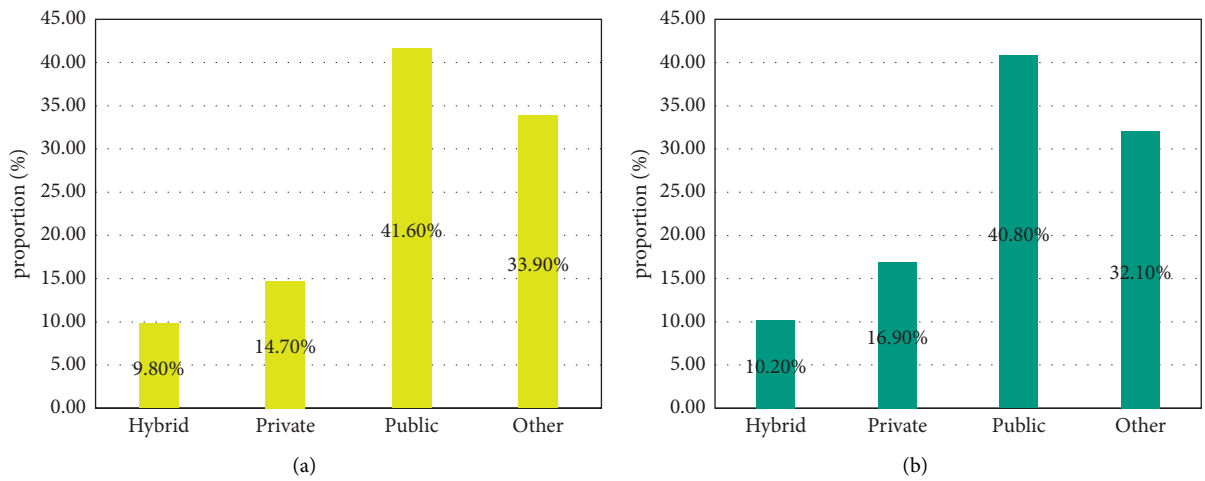


Figure 2: Domestic cloud service usage in the past two years. (a) 2019. (b) 2020.

apply for more resources when they are needed and can release them when they are no longer needed. From the user's point of view, these resources are unlimited, and users only pay for the resources they use. The fifth is strong adaptability. The cloud computing system is an adaptive system, which can automatically balance the load and optimize the utilization of resources and can update resources adaptively with the changes of the data age. Users can also monitor resource usage [12–14].

(4) *Classification*. According to the type, there are software (SaaS), platform (PaaS), and infrastructure (IaaS). There are private type, public type, mixed type, and community type according to the method. The specific content of the three service types is shown in Table 1, and the application proportion of the four service methods is shown in Figure 2 [15].

(5) *Core Technology*. The core technologies of cloud computing mainly include programming technology and information security technology. Among them, the first six technologies are the most important [16].

(6) *Application Areas*. Cloud computing technology is commonly used in today's Internet services. The most common are online mailboxes (such as Google and Baidu) and online search engines. Users can search for the resources they need at any time on their mobile devices and share data resources through the network cloud. The same goes for online mailboxes, where sending and receiving e-mails used to be a tedious and time-consuming process. With the popularization of cloud computing and network technology, e-mail has become a part of social life. As long as it is in the network environment, real-time delivery of e-mails can be realized. At present, the main application areas can be divided into storage cloud (such as Microsoft, Google, and other large networks), financial cloud (such as Alibaba's Alibaba

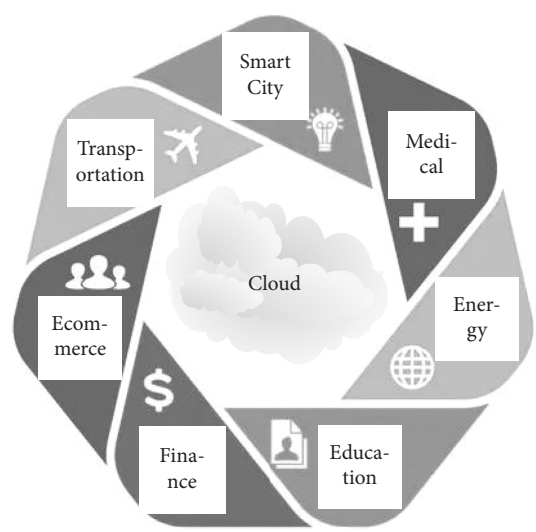


Figure 3: Application areas of cloud computing.

Figure 4: BD usage.

Cloud, Tencent Financial Cloud, and so on), education cloud, medical cloud, and so on. The details are shown in Figure 3.

(7) *Problems Faced*. At present, several problems faced by cloud computing mainly include serious information leakage, users have no right to access, data systems are not comprehensive, and there is no sound legal protection. To solve these problems, we must first improve the legal system and publicize the awareness of legal security. Secondly, the access rights are used correctly, and the data system is comprehensively improved [17].

3.2. The Correlation of BD

(1) *The Definition of BD*. Generally speaking, under different requirements, the required time processing range is different [18]. Figure 4 shows some important uses of BD.

(2) BD are characterized by complex structure, huge quantity, diverse types, and low value density. At present, BD are changing from calculating some data to analyzing all data, from calculating microscopic results to discovering macro trends, and from exploring cause and effect to exploring information correlation.

(3) *The Core Technology of BD*. The core technologies of BD include BD collection technology, storage technology, mining and analysis technology, and visualization technology. The details are shown in Figure 5.

(4) The main application fields of BD include e-commerce industry, financial industry, biotechnology, smart government, education industry, transportation industry, medical industry, etc. The usage ratios of these main application areas are shown in Table 2. The data on the proportion of applications here are compiled from public information.

(5) *The Association Rules of BD*. When this paper studies the correlation of BD, it first analyzes the attributes of BD. BD refer to a massive collection of unfiltered data. They can rationally express objective things in many aspects, such as words, symbols, letters, shapes, and so on, which are all data expressions in different forms. Data exist in data values and data structures, and large amounts of data form a complex network. In the network, the data value is its information entity, and the data structure can be regarded as the relationship between entities. The association between data can include time and space association, entity and virtual association, network level association, and so on [19–22].

The relationship between time and space refers to the attribute of describing data using space and time, which is beneficial to people's data mining. Time correlation can usually be divided into time points and time periods, and the expression of development and change of things needs to be represented by time attributes. Spatial correlation is mostly used in the visualization of geographic data, and it is more helpful for people to understand when the background is selected as a familiar map in the visualization based on

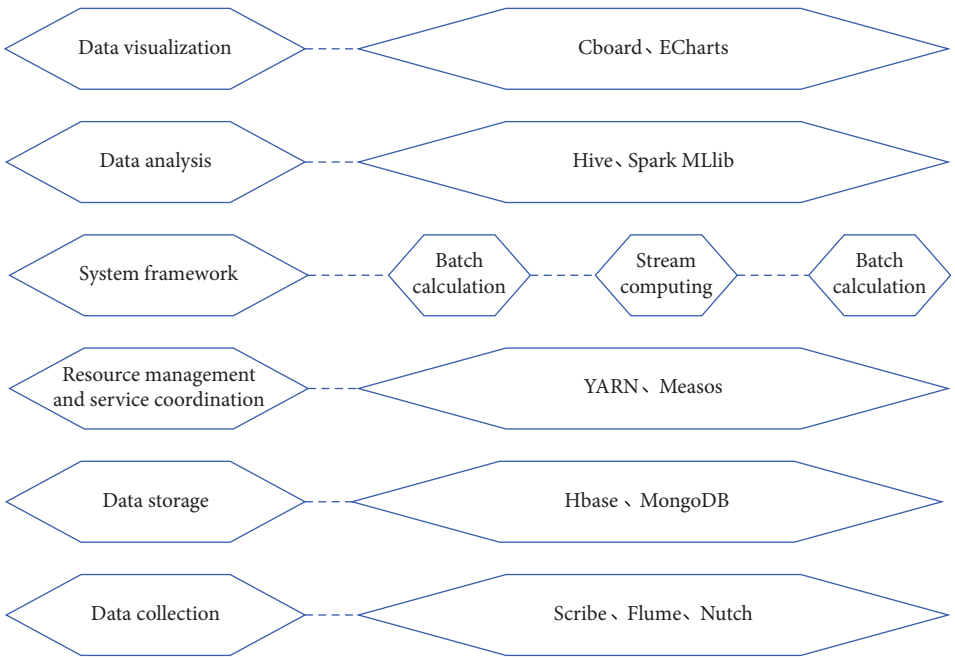


Figure 5: Core technologies of BD.

Table 2: The proportion of BD applications.

Type of enterprise	Proportion (%)
Government affairs	13
Internetwork	8
Medical and health	14
Education	8
Traffic	6
Finance	30
Electronic commerce	4
Others	17

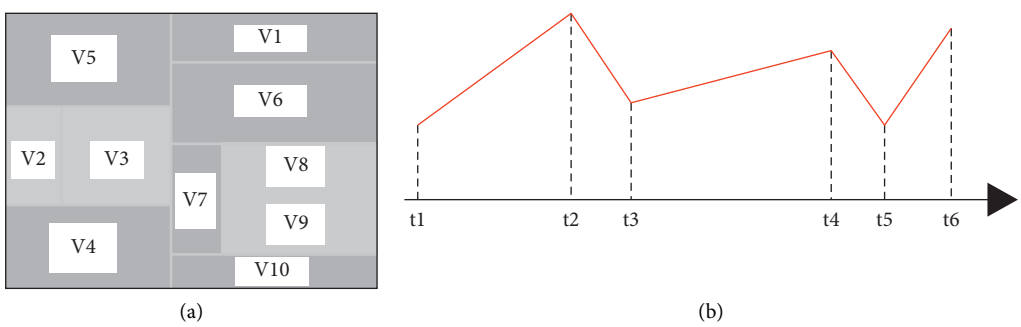


Figure 6: Spatiotemporal correlation of data. (a) Spatial correlation. (b) Time correlation.

geographic information. As shown in Figure 6, there is no BD correlation in time and space [23, 24].

Entity associations are often used in visualization to represent entities with different visual representations. Entity attributes in BD can usually be changes and combinations of three types of entity attributes: category attributes, interval attributes, and numerical attributes. The combination of virtual and real BD can more conveniently detect and control entities, as shown in Figure 7.

3.3. BD Real-Time Controllable Algorithm. The so-called real-time controllability of BD is to monitor and control these data anytime and anywhere at the terminal.

Taking the K-means algorithm as an example and assuming that the center of the data cluster is W, then

$$W_i = \left(\frac{1}{n_i}\sum i_{i1}, \frac{1}{n_i}\sum i_{i2}, \frac{1}{n_i}\sum i_{i2}, \ldots, \frac{1}{n_i}\sum i_{ip}\right), \quad (1)$$

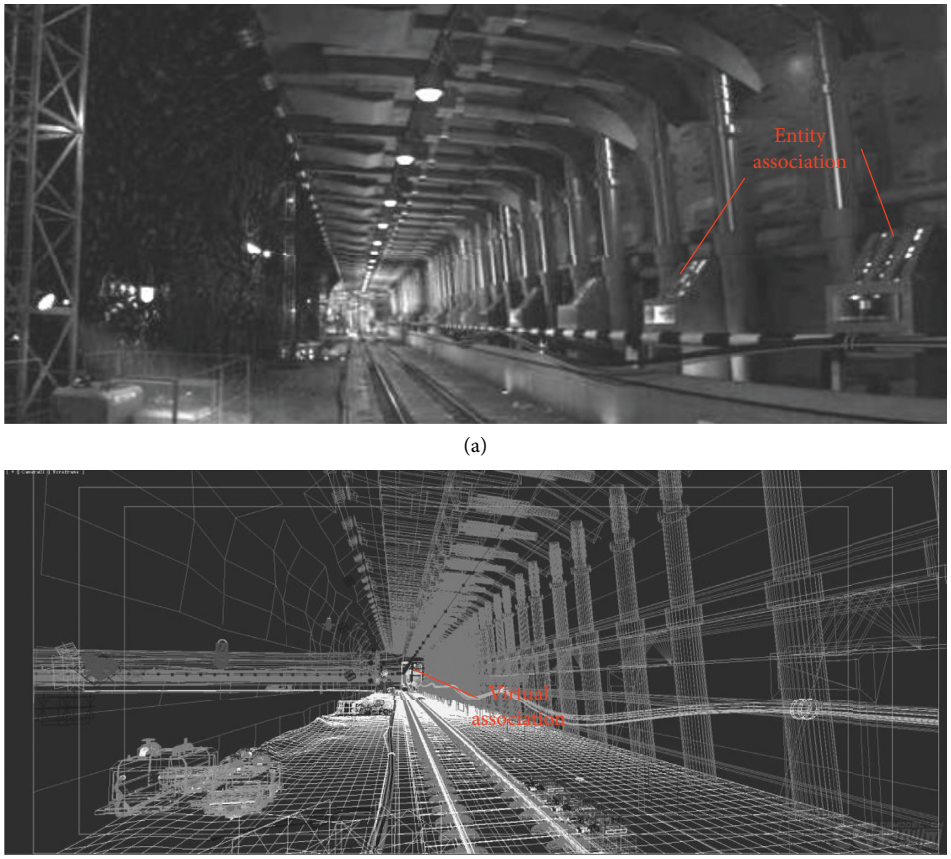


Figure 7: Virtual-real association of data. (a) Entity association. (b) Virtual association.

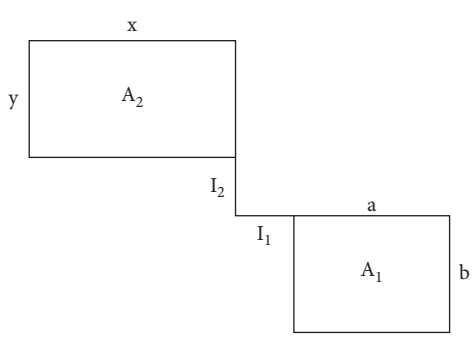


Figure 8: Side length and distance relationship of two minimum circumscribed matrices.

where n_i is the number of data in clusters and i_{ip} is the object of p-dimensional data.

The Euclidean distance of the two data is

$$d(I_i, I_j) = \sqrt{|i_{i1} - i_{j1}|^2 + |i_{i2} - i_{j2}|^2 + \cdots + |i_{ip} - i_{jp}|^2}, \quad (2)$$

where i_{jp} is another object of P-dimensional data.

The average distance from all data points in the cluster to the center point is called intracluster similarity, which can be expressed as

Table 3: Serial-parallel algorithm simulation parameters.

Dataset size (kB)	Number of association rules
1220	14975
2484	14810
4960	14804
9960	14831
19921	14546
39866	14523

$$\text{inner} = \frac{1}{n} \sum_{I \in C_i} \sqrt{|i_1 - W_{i1}|^2 + |i_2 - W_{i2}|^2 + \cdots + |i_p - W_{ip}|^2}, \quad (3)$$

where C_i represents the cluster, and the smaller the inner, the higher the similarity.

The minimum distance between clusters and cluster centers is called intercluster similarity, which can be expressed as

$$\begin{aligned} \text{ext} &= \min W_i - W_j \\ &= \min \sqrt{|W_{i1} - W_{j1}|^2 + |W_{i2} - W_{j2}|^2 + \cdots + |W_{ip} - W_{jp}|^2}. \end{aligned} \quad (4)$$

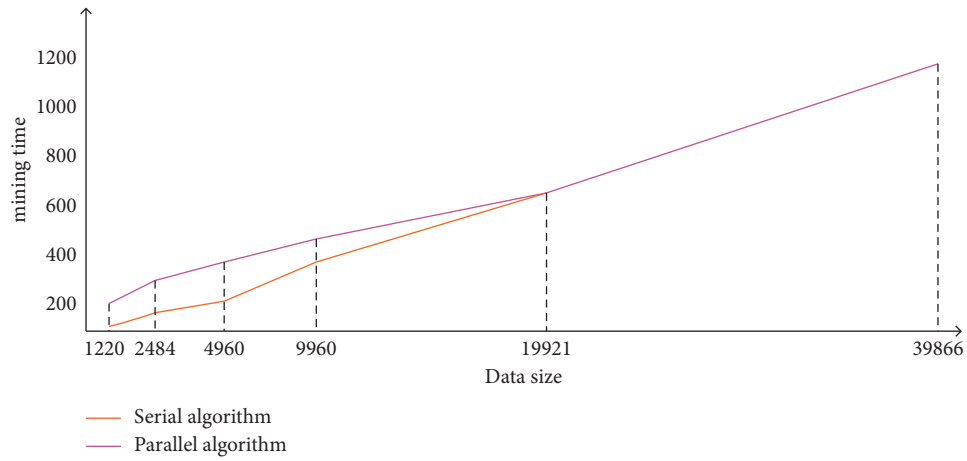


Figure 9: Comparison of serial and parallel algorithms.

The smaller the ext, the higher the similarity. On the contrary, it is smaller.

Taking the nearest neighbor algorithm as an example, suppose a dataset is U, x_i represents the data in it, and the m attributes of the data x are x_{im}.

Then, there are

$$U = \{x_{ij}\}_{n*m}. \tag{5}$$

Normalize to get

$$R = \{r_{ij}\}_{n*m}. \tag{6}$$

According to the information entropy theory (information entropy is a measure of the uncertainty of information; the greater the uncertainty of the information, the greater the information entropy and the greater the value of the information), the expression of each attribute weight is

$$\begin{aligned} E_j &= -\frac{1}{\ln n}\sum_{i=1}^{n} r_{ij}^X \ln r_{ij}^X, \\ r_{ij}^X &= \frac{r_{ij}}{\sum_{i=1}^{n} r_{ij}}, \end{aligned} \tag{7}$$

where r_{ij}^X represents the matrix weight; then, the weight of the jth attribute is

$$w_j = \frac{1 - E_j}{\sum_{i=1}^{m} E_i}. \tag{8}$$

The weighted distance for any two data in the dataset can be expressed as

$$D(x_i, x_j) = \sqrt{\sum_{p=1}^{m} w_p (x_{ip} - x_{jp})^2}. \tag{9}$$

Then, the sum of the distances of all the data in the set is

$$D_u = \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} D(x_i, x_j). \tag{10}$$

Table 4: Simulation parameter settings.

Dataset size (kB)	Number of blocks
3800	50
12500	145
25000	300
33000	405
40000	536

Set the distance standard of adjacent data to ε; then,

$$\varepsilon = D_u \times \frac{n(n-1)}{2}. \tag{11}$$

Set the neighborhood of the data to δ_i; then,

$$\delta_i = \{x \in U(x_i, x_j) < \varepsilon\}. \tag{12}$$

In order to determine whether the data in a dataset or function are in the neighborhood content, define

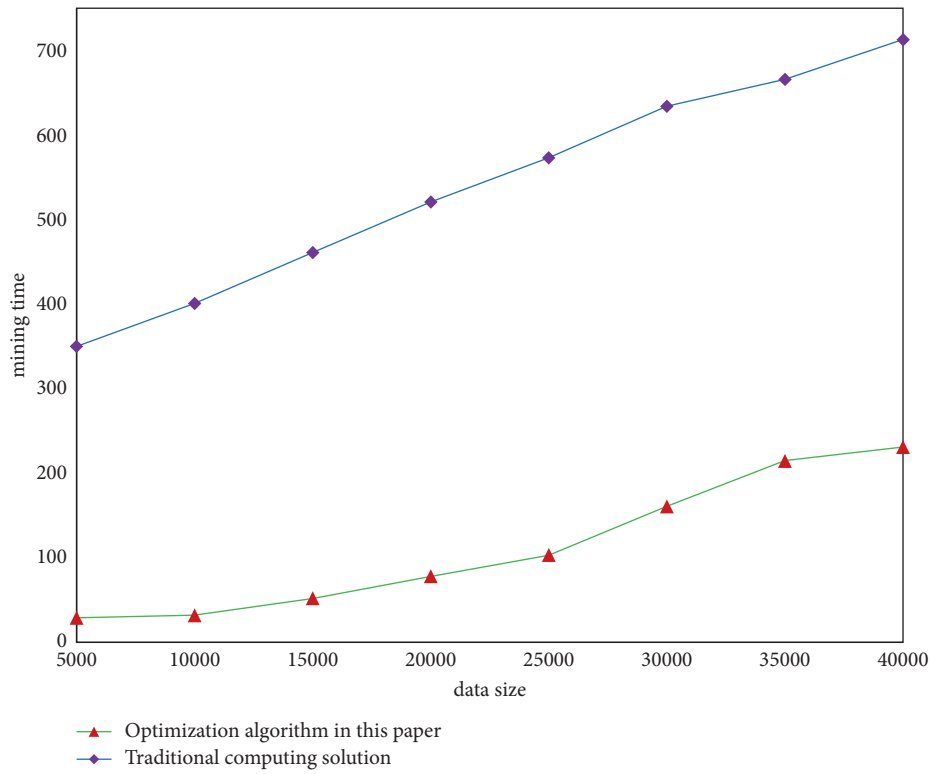
$$IS(x_i, x_j) = \begin{cases} 0, & x_i \notin \delta(x_j), \\ 1, & x_i \in \delta(x_j). \end{cases} \tag{13}$$

Assuming that the probability of data x_j appearing in the neighborhood space of the rest of the data is P_C, then

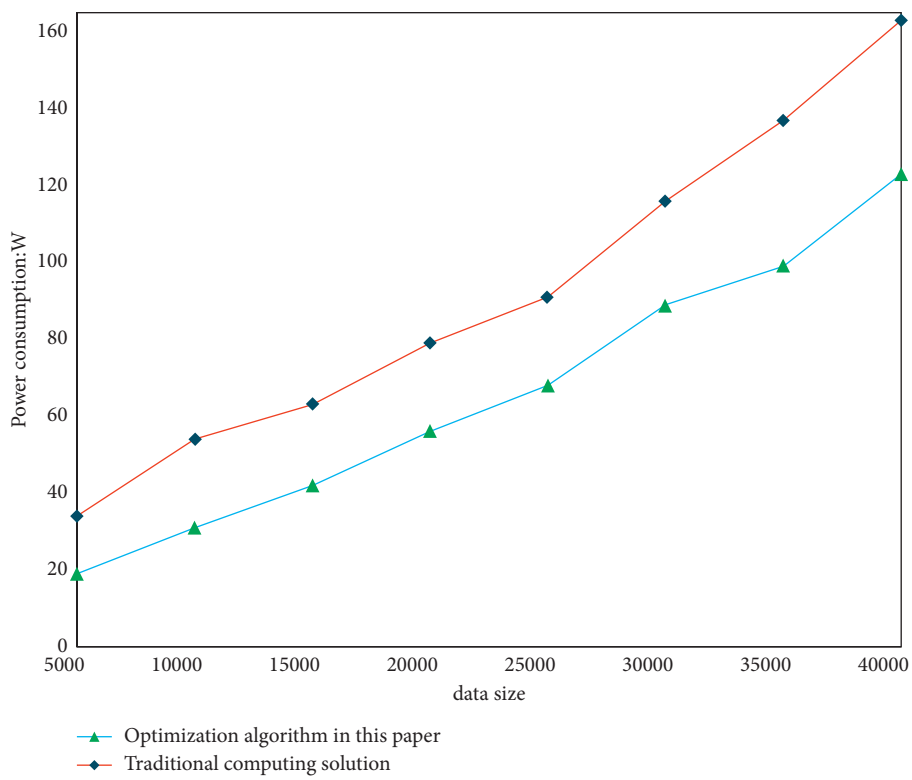
$$P_C = \frac{\sum_{j=1}^{n} IS(x_i, x_j)}{n}. \tag{14}$$

4. Optimization Algorithm Test of Dynamic *K* Value

4.1. Real-Time Controllable Optimization Algorithm for Correlated BD. This paper combines the proximity algorithm to determine the dynamic K value of the K-means algorithm, so as to construct the network model of the correlation data. The construction steps are as follows: first, the entire dataset is to be clustered, and the algorithm is applied to cluster it. The obtained clustering results are adjusted. Then, using each of the clusters as a dataset,



(a)



(b)

Figure 10: Comparison between optimization algorithms and traditional cloud computing. (a) Comparison of mining time. (b) Power consumption comparison.

perform corresponding clustering and adjustment. This step is performed iteratively on each resulting subcluster until the data network is constructed [25–27].

Assuming that the network space composed of one or more data has the two minimum circumscribed matrices A_1 and A_2, then the relationship between the side lengths and distances between the two can be shown in Figure 8.

Then, the area of the new space minimum circumscribed matrix obtained by the two is

$$S = (a + l_2 + x) \times (b + l_1 + y). \tag{15}$$

When the center points of A_1 and A_2 remain unchanged, the formula can be expanded to get

$$S = xy + (b + l_1)x + (a + l_2)y + (a + l_2)(b + l_1). \tag{16}$$

It can be seen from this that the area of the newly constructed minimum circumscribed rectangle is related to the area and perimeter of the smallest circumscribed rectangle that composes it.

According to the previous proximity algorithm and K-means algorithm, the distance between the two circumscribed matrices can be obtained:

$$D(A1, A2) = e^{(S/1+S)(C/1+C)} \sqrt{\sum_{J=1}^{M} W_j \left(c_{1j} - c_{2j}\right)^2}, \tag{17}$$

where S is the area and C is the perimeter.

In this paper, the parameter λ is introduced on the basis of formula (11) for data optimization, namely:

$$\varepsilon = \lambda * D_u \times \frac{n(n-1)}{2}. \tag{18}$$

In order to optimize the clustering effect between each data, the value of the number of clusters of any child node data of the intermediate node is

$$k' = \left| \frac{\sum_{i=1}^{n} k_i}{n} \right|. \tag{19}$$

4.2. Optimization Algorithm Simulation Test. Firstly, the serial algorithm is simulated, and the introduced improved adjacent K-means algorithm is run under Win7, and the specific performance of the Oracle database (Oracle database system is a popular relational database management system in the world; the system has good portability, convenient use and powerful functions, and is suitable for all kinds of large, medium, and small computer environments; it is an efficient, reliable, and high-throughput database solution) for mining data is recorded. The serial and parallel algorithms mine the same dataset and keep increasing the size of the dataset, and both algorithms set the same threshold, different datasets must be tested, a total of 6 times. The simulation parameters are shown in Table 3, and the experimental results are shown in Figure 9.

From Figure 9, the experimental results show that as the scale of data processed by the serial-parallel algorithm gradually increases, the memory consumed by the serial algorithm gradually increases. When the data size reaches about 39 M, the serial algorithm will report insufficient memory and cannot complete association rule mining, while the improved parallel algorithm can complete the task. But since the Hadoop platform runs in pseudo-parallel mode, the performance of a single node is the same as that of a serial algorithm. Also, the parallel algorithm requires interaction between tasks, so when the data size is small, the parallel algorithm takes much more time than the serial algorithm. As the data size increases, this gap gradually decreases. It can be seen that parallel algorithms are more advantageous when dealing with large datasets.

Then, the optimization algorithm and the traditional algorithm are simulated and tested in the fully distributed mode by using the parallel method. The same thresholds are also set for the two algorithms: $\delta = 0.2$, $\delta = 0.6$, and $\delta = 0.4$, and different datasets are tested for 6 times. The simulation parameters are shown in Table 4, and the results are shown in Figure 10.

It can be seen from Figure 10 that compared with the traditional cloud computing, the optimization algorithm in this paper has significantly lower mining time for data correlation. When the amount of data reaches 40 MB, traditional cloud computing takes about 700 s, while the optimized algorithm in this paper only takes about 200 s, and the traditional cloud computing time is about 3.5 times that of this paper. In addition, in terms of power consumption, the power consumption of the algorithm optimized in this paper is always lower than that of traditional cloud computing, and the difference between the two is about 20 W.

5. Discussion

The work done in this paper is purely aimed at the algorithm, but the research of the algorithm is to maximize its effect. If only the dynamic algorithm is used to replace the static algorithm, with the emergence of high-speed and multi-core processors, this research may not be of great significance. Therefore, if the research in this paper can be applied to the protocol, it may have greater practicality.

When the paper improves the algorithm simulation, because the dataset is still small, it is expected that the dataset will be enlarged in the follow-up work and then simulated to show the advantages of the improved algorithm. At present, most of the data operations in the system rely on manual arrangement of data information, which cannot really be better used by the majority of users. It cannot design and implement fully automated operations, provide users with a good interface, and cannot directly obtain data mining results.

With the development of society, BD has increasingly penetrated into human's daily work, life, and various safety-critical application environments. In particular, the rapid development of mobile electronic products and mobile Internet applications has brought new research topics for energy saving and reliability optimization of real-time systems. To this end, continuous and in-depth research and practical work are required to adapt and optimize real-time controllable algorithms.

6. Conclusion

In the research on the real-time controllable optimization algorithm of BD, this paper firstly explains the background meaning of the era of BD in the abstract part. Then, this paper explains the purpose of this research and the theoretical algorithm used. Then, this paper explains the research value of BD background and cloud computing in the reference part. This article exemplifies the related research of many scholars in the real-time controllable optimization technology of BD and analyzes their research results and the shortcomings of their research.

In the theoretical research part, this paper first introduces cloud computing. It includes its source, definition, characteristics, classification, core technology, application fields, and the challenges cloud computing is facing. Then, this paper introduces the concept, characteristics, core technology, main sources, and association rules of BD and explains with the help of charts.

Finally, in the algorithm design of optimization and upgrading, this paper proposes a combination of proximity algorithm and K-means algorithm. It is optimized and upgraded through parallel computing. After several simulation tests of performance parameters, the results obtained are compared with traditional cloud computing, and the advantages of the scheme proposed in this paper are obtained.

Data Availability

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

Conflicts of Interest

The authors declare that they have no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Acknowledgments

This study was supported by the Zhejiang Provincial Natural Science Foundation of China under grant no. LGG20F020013 (Rutao Li). This study was also supported by the Excellent Talent Foundation of China West Normal University (no. 17YC497) (Zaiyi Pu).

References

[1] V. Puig, C. Ocampo-Martínez, R. Pérez, G. Cembrano, J. Quevedo, and T. Escobet, "Advances in industrial control] real-time monitoring and operational control of drinking-water systems," *BD Analytics and Knowledge Discovery Applied to Automatic Meter Readers*, Springer, Berlin, Germany, 2017.

[2] C. Qian, A. Liu, R. Huang et al., "Quality control of marine BD——a case study of real-time observation station data in Qingdao," *Journal of Oceanology and Limnology*, vol. 37, no. 06, pp. 195–205, 2019.

[3] D. Puthal, S. Nepal, R. Ranjan, and J. Chen, "DLSeF: a dynamic key-length-based efficient real-time security verification model for bd stream," *ACM Transactions on Embedded Computing Systems*, vol. 16, no. 2, pp. 1–24, 2017.

[4] S. Ren, X. Zhao, B. Huang, Z. Wang, and X. Song, "A framework for shopfloor material delivery based on real-time manufacturing big data," *Journal of Ambient Intelligence and Humanized Computing*, vol. 10, no. 3, pp. 1093–1108, 2019.

[5] N. N. Bao, Y. Huang, B. J. Xiao et al., "Design of real-time feedback control of vertical growth rate on EAST," *Nuclear Science and Techniques*, vol. 32, no. 7, pp. 1–9, 2021.

[6] R. Janzen and S. Mann, "Extreme-dynamic-range sensing: real-time adaptation to extreme signals," *IEEE multimedia*, vol. 24, no. 2, pp. 30–42, 2017.

[7] Y. Wang, J. Li, and H. H. Wang, "Cluster and cloud computing framework for scientific metrology in flow control," *Cluster Computing*, vol. 22, no. 1, pp. 1–10, 2017.

[8] L. Ogiela, M. R. Ogiela, and H. Ko, "Intelligent data management and security in cloud computing," *Sensors*, vol. 20, no. 12, p. 3458, 2020.

[9] S. Namasudra and P. Roy, "Ppbac: popularity based access control model for cloud computing," *Journal of Organizational and End User Computing*, vol. 30, no. 4, pp. 14–31, 2018.

[10] Z. Cao, L. Jin, C. Wan, Y. Song, Y. Zhang, and X. Wang, "Optimal cloud computing resource allocation for demand side management," *IEEE Transactions on Smart Grid*, vol. 8, no. 4, pp. 1943–1955, 2017.

[11] J. Du, L. Zhao, J. Feng, and X. Chu, "Computation offloading and resource allocation in mixed fog/cloud computing systems with min-max fairness guarantee," *IEEE Transactions on Communications*, vol. 66, no. 4, pp. 1594–1608, 2018.

[12] T. Hirai, H. Masuyyama, S. Kasahara, and Y. Takahashi, "Performance analysis of large-scale parallel-distributed processing with backup tasks for cloud computing," *Journal of Industrial and Management Optimization*, vol. 10, no. 1, pp. 113–129, 2017.

[13] X. Liu, X. Zhang, W. Li, and X. Zhang, "Swarm optimization algorithms applied to multi-resource fair allocation in heterogeneous cloud computing systems," *Computing*, vol. 99, no. 12, pp. 1–25, 2017.

[14] Y. Zhang, X. Xiao, L. X. Yang, Y. Xiang, and S. Zhong, "Secure and efficient outsourcing of PCA-based face recognition," *IEEE Transactions on Information Forensics and Security*, vol. 15, 2019.

[15] W. Wei, X. Fan, H. Song, X. Fan, and J. Yang, "Imperfect information dynamic stackelberg game based resource allocation using hidden markov for cloud computing," *IEEE Transactions on Services Computing*, vol. 11, no. 99, pp. 78–89, 2018.

[16] S. Wang, K. Liang, J. K. Liu, J. Chen, J. Yu, and W. Xie, "Attribute-based data sharing scheme revisited in cloud computing," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 8, pp. 1661–1673, 2017.

[17] A. Barsoum and M. A. Hasan, "Provable multicopy dynamic data possession in cloud computing systems," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 3, pp. 485–497, 2017.

[18] J. L. Tsai and N. W. Lo, "A privacy-aware authentication scheme for distributed mobile cloud computing services," *IEEE Systems Journal*, vol. 9, no. 3, pp. 805–815, 2017.

[19] L. Ping, J. Li, Z. Huang et al., "Multi-key privacy-preserving deep learning in cloud computing," *Future Generation Computer Systems*, vol. 74, pp. 76–85, 2017.

[20] W. Tang and W. Feng, "Parallel map projection of vector-based big spatial data: c," *Computers, Environment and Urban Systems*, vol. 61, no. Part B, pp. 187–197, 2017.

[21] Z. Lv, S. Lv, and J. Li, "Clinical characteristics and analysis of risk factors for disease progression of COVID-19: a retrospective Cohort study," *International Journal of Biological Sciences*, vol. 17, no. 1, pp. 1–7, 2021.

[22] M. Lee, L. Mesicek, and K. Bae, "AI advisor platform for disaster response based on big data," *Concurrency and Computation: Practice and Experience*, vol. 12, Article ID e6215, 2021.

[23] T. G. Rodrigues, K. Suto, H. Nishiyama, and N. Kato, "Hybrid method for minimizing service delay in edge cloud computing through VM migration and transmission power control," *IEEE Transactions on Computers*, vol. 66, no. 5, pp. 810–819, 2017.

[24] M. Abdel-Basset, M. Mohamed, and V. Chang, "NMCDA: a framework for evaluating cloud computing services," *Future Generation Computer Systems*, vol. 86, pp. 12–29, 2018.

[25] H. Cai, B. Xu, L. Jiang, and A. V. Vasilakos, "IoT-based big data storage systems in cloud computing: perspectives and challenges," *IEEE Internet of Things Journal*, vol. 4, no. 1, pp. 75–87, 2017.

[26] Z. Lv, "The security of Internet of drones," *Computer Communications*, vol. 148, pp. 208–214, 2019.

[27] Y. Guo, H. Mi, and M. Habibi, "Electromechanical energy absorption, resonance frequency, and low-velocity impact analysis of the piezoelectric doubly curved system," *Mechanical Systems and Signal Processing*, vol. 157, Article ID 107723, 2021.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 2057198, 9 pages
https://doi.org/10.1155/2022/2057198

Research Article

Multilabel Emotion Recognition Technique considering the Characteristics of Unstructured Conversation Data

Myungjin Lim,[1] Moungho Yi,[2] Pankoo Kim,[1] and Juhyun Shin[3]

[1]*Department of Computer Engineering, Chosun University, Gwangju 61452, Republic of Korea*
[2]*Department of Electronic Engineering, Chosun University, Gwangju 61452, Republic of Korea*
[3]*Department of new Industry Convergence, Chosun University, Gwangju 61452, Republic of Korea*

Correspondence should be addressed to Juhyun Shin; jhshinkr@chosun.ac.kr

Received 4 February 2022; Revised 5 April 2022; Accepted 25 April 2022; Published 9 May 2022

Academic Editor: Chun-Wei Tsai

Due to the recent increase in non-face-to-face services due to COVID-19, the number of users communicating through messengers or SNS (social networking service) is increasing. As a large amount of data is generated by users, research on recognizing emotions by analyzing user information or opinions is being actively conducted. Conversation data such as SNS is freely created by users, so there is no set format. Due to these characteristics, it is difficult to analyze using AI (artificial intelligence), which leads to a decrease in the performance of the emotion recognition technique. Therefore, a processing method suitable for the characteristics of unstructured data is required. Among the unstructured data, most emotion recognition in Korean conversation recognizes a single emotion by analyzing emotion keywords or vocabulary. However, since multiple emotions exist complexly in a single sentence, research on multilabel emotion recognition is needed. Therefore, in this paper, the characteristics of unstructured conversation data are considered and processed for more accurate emotion recognition. In addition, we propose a multilabel emotion recognition technique that understands the meaning of dialogue and recognizes inherent and complex emotions. A deep learning model was compared and tested as a method to verify the usefulness of the proposed technique. As a result, performance was improved when it was processed in consideration of the characteristics of unstructured conversation data. Also, when the attention model was used, accuracy showed the best performance with 65.9%. The proposed technique can contribute to improving the accuracy and performance of conversational emotion recognition.

1. Introduction

With the development of the Internet and the increase of non-face-to-face services due to COVID-19, the number of cases of communication through messengers or SNS (social networking service) is increasing. As a large amount of data is generated by users, research to analyze user information and opinions is being actively conducted [1, 2]. Among them, emotion recognition is a research field that recognizes emotions by analyzing personal opinions about social issues or products [3, 4]. However, due to the nature of non-face-to-face communication, misunderstandings may occur unexpectedly as they are interpreted differently between the sender and the receiver. And sometimes the contextual context intended by the sender is not communicated to the receiver. Transparent emotional exchange is an important factor for dialogue without misunderstanding between the parties involved.

Text data is the most developed emotional expression means that can more precisely express the various emotions that humans experience. Recognizing the emotions conveyed by textual expressions is essential for machines to read human emotions. In order for AI (artificial intelligence) to accurately grasp the emotional responses contained in text expressions, it is essential to have a rich and sophisticated database of human emotional expression languages. And since the emotions that humans experience in social situations are much more diverse than the basic emotions, it is necessary to classify as many emotions as possible [5]. Recognizing the user's emotions from text data allows us to

analyze the user's intentions in various ways [6]. It also helps to understand the user's point of view.

There is no set format for conversations used on the Internet. Among these unstructured conversational data, if we look at the characteristics of Korean, many new words and profanities are generated and used frequently [7]. In addition, the use of abbreviations that freely transformed Korean is increasing. The use of such unstructured data sometimes causes communication problems because users other than the main user group have difficulties in understanding words and meanings. In addition, since it does not have a dictionary meaning, performance degradation and limitations of the emotion recognition algorithm occur [8]. Therefore, it is necessary to identify and process the characteristics of unstructured Korean, which includes words such as neologisms, profanities, and abbreviations. Another characteristic of unstructured Korean is that it is difficult to analyze because it is different from English. Therefore, it is important to understand the order of words and the meaning of the entire sentence. If only the word is used as a feature, data sparseness may occur depending on the characteristics of Korean, where morphological transformations occur in various ways. Therefore, a better understanding of the characteristics of unstructured Korean conversation data and a specialized processing method are needed.

Text emotion classification studies mainly recognize emotions through vocabulary. Alternatively, emotions in word or sentence units are analyzed using an emotion dictionary or a large number of uni-grams. However, in this classification, it is important to balance the quality of training data, the amount of training data, and the data [9]. In addition, as a study of emotion classification using models, attempts to improve emotion classification performance using CNN models, LSTM models, and CNN-LSTMs continued [10–12]. However, research that improves quality by improving data rather than improving the model is more effective in terms of performance [13].

Multilabel emotion recognition in conversations is an important task in providing the ability to understand the user's various emotions. As a related study, using MEISD (Multimodal EmotionLines Dataset), it was proved that the results of multilabel emotion recognition of text are more accurate than other inputs based on DialogueRNN [14]. In addition, using SemEval-2018, multiple emotions of sentences were recognized based on AttnConvnet [15]. And we applied a content-based method to tweets to classify multilabel emotions [16]. Attempts to recognize multiple emotions in various ways are continuing, but most of them were studies using English data sets. Most of the studies using the Korean conversation data set have a limitation in that they are classified as a single emotion. Therefore, considering the characteristics of the Korean language, it is necessary to understand the meaning of dialogue and recognize the underlying emotions. And we need a way to recognize the various emotions included in the sentence.

Therefore, in this paper, the following technique is proposed to solve the processing problem suitable for unstructured Korean conversation data and the single emotion recognition problem. It is a data processing method that considers the characteristics of unstructured conversation data for more accurate emotion recognition and a multilabel emotion recognition method applying it. It contributes to improving the accuracy and performance of emotion recognition by processing unstructured data and recognizing various emotions inherent in sentences. The structure of this paper is as follows. The related study describes emotion recognition of unstructured texts and emotion analysis in Korean, and the proposed method describes a multilabel emotion recognition technique that considers the characteristics of unstructured conversation data. In the experimental results, the proposed technique is tested and evaluated, and finally, conclusions and future research are described and finalized.

2. Related Work

In text emotion recognition, when a machine recognizes human emotions, it classifies emotions on current input data. In addition, beyond this level, it is necessary to enable more accurate emotional recognition intelligently according to past memories, emotional subjects, personality, or inclinations [17]. In the past, most emotions were judged by extracting emotion keywords. Since this method loses various syntactic or semantic information contained in natural language sentences, there is a limit to recognizing complex human thoughts such as emotions.

Unstructured text emotion classification studies mainly try to recognize emotions through vocabulary. In the study of [18], an emotion word dictionary and an emotion emoticon dictionary corresponding to Ekman's six emotion categories were constructed. Based on this, we tried to recognize emotions in the blog text by using the SVM classifier. In the study of [19], the data from Twitter, a representative social network service, were analyzed as the subject of analysis. Instead of constructing and utilizing an emotion dictionary for emotion classification, a large number of uni-grams were used. However, in the classification of the supervised learning method using dictionary or vocabulary features, the problem of insufficient data is the most problematic. Various emotion classes and vocabulary used for each class appear in various ways [20]. Therefore, it is very difficult to construct sufficient learning data to enable learning at an appropriate level using these vocabularies. Therefore, it is important to balance the quality of training data, the amount of training data, and the data.

In the study of [21], as a study on the analysis of Korean emotions, emotions were analyzed using a dictionary of emotional words centered on Korean words like English. In this case, there is a disadvantage that one word can fall into the problem of ambiguity in which it is interpreted as having more than one meaning. Considering this, in the study of [22], emotion was learned by focusing on the order of words without using part-of-speech tagging or emotion word dictionary. We present an algorithm with high accuracy in Korean that compares and analyzes the pattern of word order and input sentences. However, in the case of having a clear emotional word rather than an ambiguous expression, the accuracy was rather low. Therefore, not only the order of words but also the use of the emotional word dictionary is important in the analysis of Korean emotions.

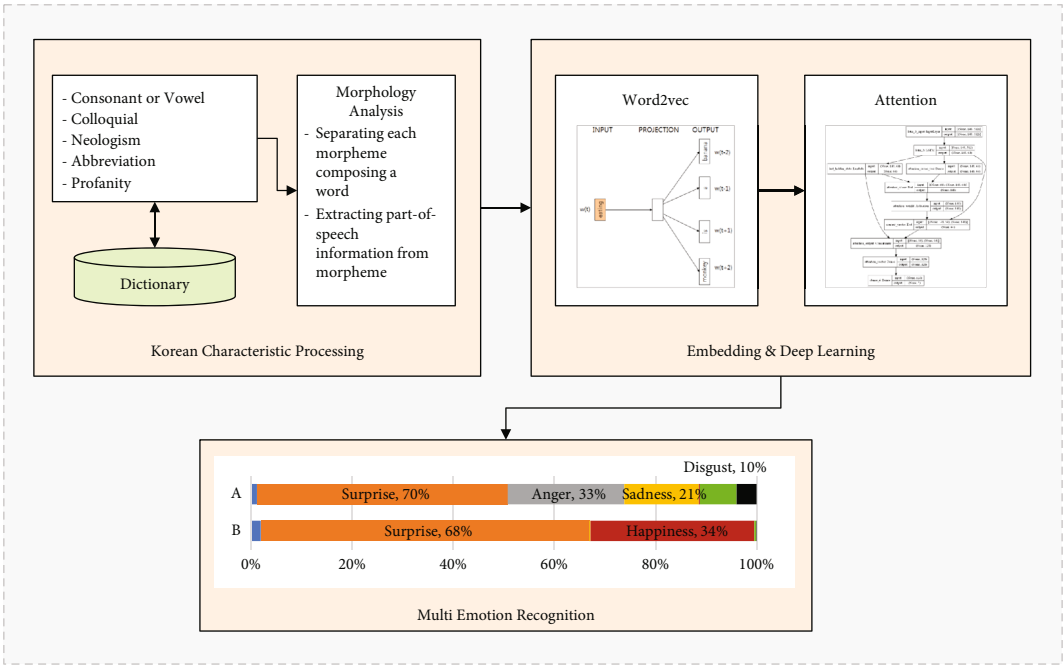


Figure 1: System architecture.

Table 1: Rules for processing unstructured words.

Division	Original data → processing data
Consonant or vowel	k → keu, h → heu, d → da, u → ngu, yu → ngyu
Colloquial	~dang → ~ da, ~ngyos → ~ ngyo, ~neong → ~ neo
Neologism	romgok → tear, TMI → too much information
Abbreviation	masang → pain in the heart
Profanity	kkomsi → the technique is vicious and cunning

Input: OriginalDataList
Output: ProcessDataList
Definition:*Dictionary* – standard word definition for unstructured words
[consonant or vowel, colloquial, abbreviation, neologism, profanity]

```
READ OriginalDataList;
for each sentence in OriginalDataList do
    for each word in sentence do
        if word in Dictionary then
            word ← standardword;
        end if
    end for
end for
ProcessDataList ← OriginalDataList
return ProcessDataList;
```

Algorithm 1: Korean characteristic processing.

And in order to solve the ambiguity problem, the entire sentence should be used for emotion analysis.

As a study of emotion classification using deep learning, in the study of [10], the CNN model showed good performance in sentence and document classification with simple convolution and pooling. In addition, in the study of [11], the LSTM model was trained considering the order of continuous input data. In terms of generation, it shows good performance not only in machine translation but also in various problems. In the study of [12], context was reflected within the classification model. In addition, the CNN-LSTM model was used to classify speech emotions in

```
Input: d:dataset
Output: Matrix W(512,113) of one-hot vectors for each possible byte value (0-255)
let f be a list of tuples (byte_value, frequency);
for i :=0 to 511 do
    freq ←0;
    for each item j in d do
        freq ← freq + frequencyOfOccurence(i,j);
    end for
    append (i,freq) tuple to f;
end for
f ← sort f based on frequencies;
W ← word2vec(f,113);
return W;
```

Algorithm 2: Word2vec embedding.

```
Input: d:dataset, l:dataset true lables, W:word2vec matrix
Output:ME of Attention trained model on test dataset
Definition:ME – multiemotion
let f be the featureset 3d matrix;
for i in dataset do
    let f_i be the featureset matrix of sample i;
    for j in i do
        v_j ← vectorize(j,w)
        append v_j to f_i
        append f_i to f
    end for
end for
    f_train, f_test, l_train, l_test ← split feature set and lables into train subset and test subset
    M ← Attention (f_train, l_train)
    ME ← evaluate (f_test, I_test, M)
return ME
```

Algorithm 3: Attention model.

Table 2: Number of data by emotion class.

Data set	Neutral	Surprise	Anger	Sadness	Happiness	Disgust	Fear	Total
Single	4,830	5,898	5,665	5,267	6,037	5,429	5,468	38,594
Continuous	43,786	4,866	3,628	1,972	1,030	220	98	55,600
Merge	48,616	10,764	9,293	7,239	7,067	5,649	5,566	94,194
(%)	(51.6%)	(11.4%)	(9.9%)	(7.7%)	(7.5%)	(6.0%)	(5.9%)	(100%)

conversations to enable automatic learning by extracting a large amount of vocabulary. All emotion classification performance was improved, but there was a limitation in that they were classified as a single emotion.

As a study on multilabel emotion recognition, [14] used MEISD (Multimodal EmotionLines Dataset) to recognize multilabel emotion based on DialogueRNN. It was proved that the results of multilabel emotion recognition of text are more accurate than that of audio and video. In the study of [15], SemEval-2018 was used to recognize multiple emotions of a sentence based on AttnConvnet, which combines attention and convolution. The two-step procedure used by humans to analyze a sentence was first to understand the meaning of the sentence and to classify the emotions. In the study of [16], multilabel sentiment was classified by applying a content-based method (word and character n-gram) to tweets. We showed that the content-based word unigram outperformed other methods. However, most of the studies were conducted using English data. Most of the studies using the Korean conversation data set were classified as a single emotion. Therefore, it is necessary to understand the meaning of conversation in consideration of the characteristics of the Korean language and to recognize the various emotions included in the sentence by recognizing the emotions embedded in it.

Most of the AI research that automatically recognizes and extracts emotions from text has been model-centric.

Table 3: Comparison of emotion recognition performance by data set.

Data set	Original data	Process data	Performance comparison
Single	43.2%	56.4%	13.2%
Continuous	40.7%	51.6%	10.9%
Merge	47.1%	65.9%	18.8%

But beyond research, the most important thing in AI development in practice was data, not models. How the data is preprocessed, how it is collected, what the size is, how good the quality is, and how the training/evaluation set is divided greatly affect the development of AI systems. In reality, it is the improvement of the data, not the improvement of the code, that increases the performance of the AI system [13]. Therefore, the method to improve the model is important, but the study that improves the quality by improving the data is more effective in terms of performance. In this paper, as a method to improve data, it is processed in consideration of the data imbalance resolution and the characteristics of unstructured conversation data. And, as a way to improve the model, we propose a multilabel emotion recognition technique in sentences using LSTM with attention.

3. Proposed Method

3.1. System Architecture. This section describes the data processing method and multilabel emotion recognition method considering the characteristics of the proposed unstructured conversation data. Figure 1 is a system configuration diagram of the proposed method. As a data set processing process, the characteristics of unstructured conversation data are considered and processed. It converts ellipsis, colloquial words, neologisms, abbreviations, and slang words into standard words and analyzes morphemes. The word vectorization is carried out by learning the morpheme-analyzed sentence using the Skip-gram method of the Word2vec model. It trains a deep learning model with learned word vectors and emotion classes and recognizes multiple emotions in conversational sentences through prediction.

3.2. Feature Processing of Unstructured Conversation Data. This section describes a data processing method considering the characteristics of unstructured conversation data. The unstructured Korean data set has language destructive features such as "consonant or vowel," "colloquial," "neologisms," "abbreviations," and "profanity." As a result, the morpheme analyzer that separates sentences by part-of-speech does not work properly [19]. Therefore, it is necessary to process the data considering the characteristics of unstructured data. In this paper, "consonant or vowel," "colloquial," "neologisms," "abbreviations," and "profanity" are defined as unstructured words. Table 1 shows an example of a processing rule for an unstructured word.

As shown in Table 1, "consonant or vowel" using only consonant or vowel such as "k" and "u" is converted to "keu" and "ngu," which are forms of "consonant + vowel" according to the characteristics of Korean. It converts "colloquial" words that are frequently used online, such as "~neong," into standard words such as "~neo." A "neologism" is converted to a standard language by referring to the dictionary of neologism [23] provided by the Naver Open Dictionary PRO and the neologism list [24] provided by Wikipedia. The "abbreviation" is converted to standard language by referring to the list of abbreviations on Namu Wiki [25]. "Profanity" is converted to standard language by referring to the Standard Korean Dictionary of the National Institute of the Korean Language [26]. According to these processing rules, a standard word suitable for an unstructured word is matched and stored in the unstructured word dictionary. Next, morphological analysis is performed. There were many sentences in which spaces were ignored in the data set. For this reason, the Okt (Open Korean Text) morpheme analyzer was used, which is good for analyzing sentences without spaces in Korean data.

Algorithm 1 shows the unstructured Korean processing rule algorithm used in this paper. Extract a sentence with an original data set (OriginalDataList) as input, and extract a word from the sentence. Next, it is checked whether the word is included in the dictionary (dictionary: defines standard words for unstructured words). If it exists in the dictionary, it is converted to a defined standard word. The original data set converted to the standard language is stored in the process data set (ProcessDataList).

3.3. Word2vec Embedding. This section describes the process of learning using the Word2vec model on data processed considering the characteristics of unstructured Korean. As unstructured Korean is highly dependent on words, it is important to understand the order of words and the meaning of the entire sentence. Therefore, as the embedding method, the Word2vec model, which can broadly grasp the meaning by considering the context before and after, was used. In order to understand the meaning according to the order of the words, the entire sentence of the data is trained with the Word2vec model to vectorize the words.

Algorithm 2 shows the algorithm of the Word2vec embedding model used in this paper. The learning method used in the Word2vec model used Skip-Gram, which predicts the context-word from the center-word. As hyperparameters, set the Learning Rate to 0.05, Dimension (Vector space) to 512, Window Size to 2, and Min Count to 5. Through embedding, it is possible to extract semantically similar words from data. And by using vector values as input values, deep learning becomes possible.

3.4. Multilabel Sentiment Recognition. In this section, we describe the process of learning vector and sentence emotion classes learned with the Word2vec model for multilabel emotion recognition using deep learning. Attention is used as a deep learning model to predict the complex emotions of a given sentence. Attention is a model that re-references the encoder's input sequence every time the decoder predicts the output word. At this time, an attention value is additionally required because the input sequence is viewed more

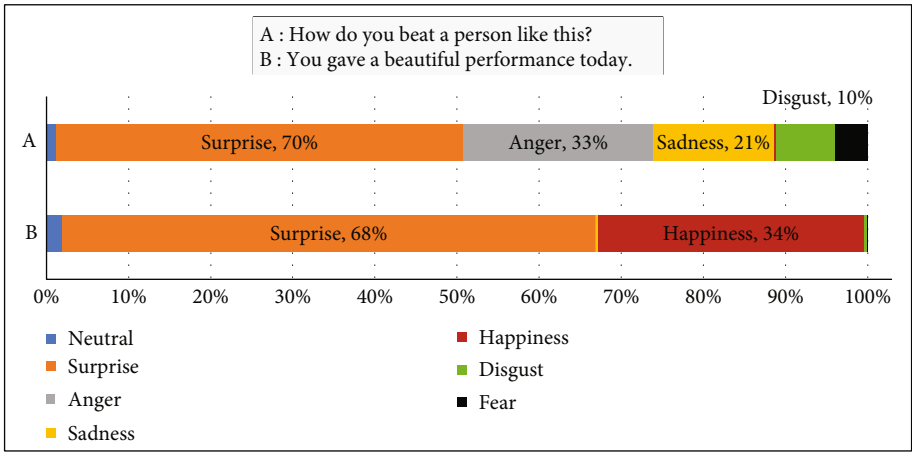


Figure 2: Results of multiemotion recognition.

Table 4: Comparative evaluation table for each emotion.

Emotion	LSTM			CNN-LSTM			Attention		
	Precision	Recall	F1-score	Precision	Recall	F1-score	Precision	Recall	F1-score
Neutral	0.387	0.387	0.387	0.387	0.444	0.410	0.459	0.470	0.459
Surprise	0.442	0.475	0.453	0.479	0.559	0.513	0.565	0.588	0.577
Anger	0.475	0.486	0.486	0.536	0.502	0.513	0.600	0.624	0.612
Sadness	0.464	0.585	0.519	0.559	0.582	0.571	0.636	0.671	0.659
Happiness	0.695	0.706	0.695	0.755	0.732	0.743	0.824	0.801	0.813
Disgust	0.618	0.486	0.541	0.686	0.536	0.605	0.765	0.683	0.718
Fear	0.640	0.629	0.640	0.640	0.686	0.663	0.765	0.754	0.754
Average	0.532	0.536	0.532	0.577	0.577	0.574	0.659	0.656	0.656
Accuracy		0.533			0.577			0.659	

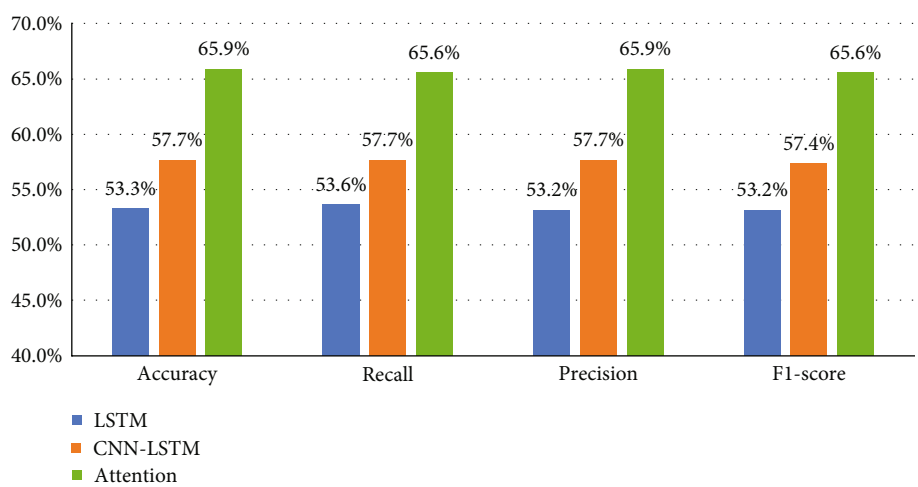


Figure 3: Comparison graph of the evaluation methods.

closely as the input word related to the predicted word. Equations (1) to (3) represent the equations for obtaining the attention value used in this paper:

$$score(s_t, h_i) = s_t^T h_i, \tag{1}$$

$$\alpha^t = softmax(e^t), \tag{2}$$

$$a_t = \sum_{i=1}^{N} \alpha_i^t h_i. \tag{3}$$

Equation (1) shows a method to obtain the attention score of the hidden state s_t at time step t of the decoder and the hidden state h_i at time step i of the encoder. Equation (2) shows a method of obtaining the attention distribution, a probability distribution in which the sum of all values becomes 1 by applying the softmax function to the attention score obtained in Equation (1). Attention distribution obtains a vector in the form of [0.1, 0.4, 0.1, 0.3], and each value is called attention weight. Equation (3) shows the method of obtaining the attention value through the attention weight obtained in Equation (2) and each hidden state. Finally, the attention value a_t is connected to the hidden state at time t of the decoder.

Algorithm 3 shows the algorithm of the attention model used in this paper. The designed model is composed of attention to solve the vanishing gradient problem, which is a disadvantage of LSTM. The input data is the number of collected sentences and 113, which is the maximum number of words per sentence. It is a 512-dimensional embedded Word2vec model that learns words and consists of 3 dimensions (number of sentences, word length, and embedding vector). Attention is located in the output part of LSTM so that it can be predicted well even if the length of the sentence is long, and 128 units are configured. In the output layer, to predict multiple emotions, the output of Dence was set according to the number of emotions, and softmax was used as the activation function.

Using the attention model, you can understand the meaning of a sentence, so you can better understand the dependencies between words. Through the semantic information extraction, it becomes possible to recognize the emotions inherent in sentences and to recognize complex emotions [27]. Even with data classified as a single emotion, using the proposed model, multiple emotions in a sentence can be recognized. Therefore, if the multilabel emotion recognition method is used, it is possible to understand the meaning of the conversation and recognize various emotions inherent in it so that more accurate emotion recognition and performance improvement are possible.

4. Experimental Results

In this chapter, we experiment with the data processing method and the multilabel emotion recognition method considering the characteristics of the proposed unstructured conversation data. In order to prove the usefulness of the data set and multilabel emotion recognition technique used in the experiment, the experimental results and performance evaluation are described.

4.1. Data Set. This section describes the data set used to test and evaluate the performance of the proposed technique. The data used were the "single conversation data set including Korean emotion information" and the "continuous conversation data set" provided by AI-HUB [28]. The two data sets were constructed by labelling the sentences selected by the web crawling on SNS, comments, and conversations, which are unstructured data, as emotion classes. The single conversation data set is divided into "sentence" and "emotion" and consists of a total of 38,594 sentences. Each sentence is classified into seven emotion classes: fear, surprise, anger, sadness, neutral, happiness, and disgust. The continuous dialogue data set is divided into "dialog," "utterance," and "emotion." It consists of a total of 55,600 sentences and is classified into 7 emotion classes. Table 2 shows the number of data for each emotion class.

As shown in Table 2, it can be seen that the data imbalance of the continuous conversation data set is more severe than that of the single conversation data set. In this paper, two data sets are merged and used to construct training data sufficient to learn emotion recognition deep learning. The merged data set (merge) consists of data with an imbalance in the number of data for each emotion class. The most emotions are the neutral class, with 48,616, accounting for 51.6%. The lowest emotion is the fear class, with 5,566, accounting for 5.9%. In this paper, to solve the data imbalance problem, the number of data per emotion class is balanced with the smallest number of 5,500. A total of 38,500 data, 5,500 of each of the seven emotions, were used in the experiment.

4.2. Performance Evaluation. In this section, we describe the performance evaluation results of the data processing method and the multilabel emotion recognition method considering the characteristics of the unstructured conversation data proposed in this paper. For the comparison with multiple data sets, a single conversation data set, a continuous conversation data set, and a merged data set (merge) were used as the original data. Then, unstructured words were converted into standard words, morphologically analyzed, and used as processing data. Next, the words were vectorized by learning the sentences with the Word2vec model. Multiple emotions were recognized through deep learning with the generated vector and emotion class. Based on 5,500 per emotion class, the emotion class that is less than 5,500 was used as it is. In the train, the test data were used in a ratio of 8 : 2. For each data set, the results of learning with the original data set and the results of learning with the processed data set were compared. Table 3 shows the results of comparing the emotion recognition performance of the original data and the processed data of the three data sets.

As shown in Table 3, it can be seen that the performance of the model trained with the processed data is more improved than the original data for all three data sets. And it can be seen that the performance of the data set that merges the two data sets is the best. In addition, it can be seen that the performance of the single conversation data set is better because the data imbalance of the continuous dialogue data set is more severe. Therefore, for more accurate emotion recognition and performance improvement, a method for balancing data and a processing method considering the characteristics of unstructured conversation data are needed.

We also analyze the results of multilabel emotion recognition. Emotions are often expressed together with other emotions rather than being expressed alone. And most emotions can be classified into positive and negative. "Happiness" is classified as positive and "anger, sadness, disgust, and fear" as

negative. However, "neutral" is an emotion that cannot be classified, and "surprise" is an emotion that is ambiguous to analyze as positive or negative. There are "positive surprise," and there are "negative surprise." Figure 2 shows the multi label emotion recognition results of two sentences classified as "surprise." In the data set, two sentences were classified with the same emotion of "surprise," but when analyzed semantically, sentence A (How do you beat a person like this?) is negative, and sentence B (You gave a beautiful performance today) is positive. Contains the meaning of as a result of multilabel emotion recognition of two sentences, sentence A contains negative multiple emotions of surprise 70%, anger 33%, sadness 21%, and disgust 10%. On the other hand, it can be seen that sentence B contains multiple positive emotions of surprise 68% and happiness 34%. Also, if the multilabel emotion recognition is applied to sentences classified as "neutral," even a small amount of emotion can be recognized, making emotion classification possible. Therefore, if the multilabel emotion recognition method is used, it is possible to understand the meaning of the conversation and recognize the inherent emotions and complex emotions so that more accurate emotion recognition and performance improvement are possible.

4.3. Comparative Evaluation. In this section, we describe the results of comparative experiments after learning three deep learning models using the processed data of the data set that merged the two data sets. In the total data, training data and validation data were composed in a ratio of 8 : 2. Learning through training data and evaluation through validation data were conducted. The deep learning model used LSTM, CNN-LSTM, and attention, which are widely used in text emotion classification studies. As for the evaluation method, accuracy, precision, recall, and F1-score were used focusing on representative emotions. Table 4 shows the comparison evaluation table for emotion recognition, and Figure 3 shows the comparison graph of the four evaluation methods. Among the proposed three deep learning models, the attention model showed the best performance with 65.9%. In the comparative evaluation by emotion class, the classification performance of the "happiness" emotion class was the best than other emotions. This is interpreted because when the seven emotions are classified as positive and negative, the "happiness" class is classified as positive, and other emotions are classified as negative. In addition, the "surprise" class included both positive and negative emotions, so the performance was classified as lower than other emotions. The "neutral" class with ambiguous classification had the lowest performance. Therefore, the data is processed by solving the data imbalance and considering the characteristics of the unstructured conversation data. Then, by classifying the sentences into multilabel emotion classes and recognizing multiple emotions, it is possible to understand the meaning of conversations and to recognize more accurate emotions embedded in the sentences and improve performance.

5. Conclusions

In this paper, for more accurate emotion recognition, a multilabel emotion recognition technique that considers the characteristics of unstructured conversation data is proposed. The data sets "singlec information" and "continuous conversation data set" provided by the AI-HUB were used. The number of data for each emotion class was composed of unbalanced data, so the number of data was balanced to 5,500. In a data processing process that considers the characteristics of unstructured Korean, unstructured words such as ellipsis, colloquial words, neologisms, abbreviations, and profane words were converted into standard words. After analyzing the morphemes using Okt, the entire sentence of the data set was vectorized by learning it with the Word2vec model. Multiple emotions were recognized by using a deep learning model with vectors and emotion classes learned with the Word2vec model. The data set is classified as a single emotion per sentence. However, using the multilabel emotion recognition method, it is possible to recognize various emotions inherent in sentences, thus showing that more accurate emotion recognition and performance improvement are possible. As a result of performance evaluation of the proposed technique, it was shown that there is a difference in emotion recognition performance depending on whether unstructured data is included and the imbalance of data. And it was seen that the accuracy of the deep learning model trained with processed data was improved by 18.8% compared to the original data. In addition, as a result of comparing accuracy, precision, recall, and Fl-score of three deep learning models, attention was 65.9% higher than CNN and CNN-LSTM, showing the highest accuracy. Therefore, the proposed technique resolves the data imbalance and recognizes multiple emotions in a sentence by applying a data processing method that considers the characteristics of unstructured conversation data. As a result, it was proved that it is possible to improve the emotion recognition performance by being able to see the change of emotion embedded in the sentence. By using the multilabel emotion recognition method that considers the characteristics of unstructured conversation data, it is possible to understand the meaning of a conversation and to recognize detailed emotions other than the representative emotions in a sentence so that more accurate emotion recognition is possible. Through these subtle emotional changes, we can recognize the flow of emotions in conversation. This study can contribute to improving the accuracy and performance of conversational emotion recognition. In addition, AI that accurately recognizes emotions can be applied to robots that interact directly with humans and can be used in various fields such as counseling therapy, emotional engineering, emotional marketing, and emotional education. For future research, we plan to build a deep learning model that recognizes multiple emotions in a continuous conversation by specifically reflecting multilabel emotions and build a system to predict continuous emotions.

Data Availability

The data used to support the findings of this study are available at https://aihub.or.kr/opendata/keti-data/recognition-laguage/KETI-02-009 and https://aihub.or.kr/opendata/keti-data/recognition-laguage/KETI-02-010.

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 Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (Nos. 2019R1F1A1057325 and NRF-2020R1A2C2007091).

References

[1] G.-M. Yoon, "Performance improvement of movie recommendation system using genetic algorithm and adjusting artificial neural network parameters," *The Journal of KING Computing*, vol. 10, no. 5, pp. 56–64, 2014.

[2] J. Seo and J. Park, "Data filtering and redistribution for improving performance f collaborative filtering," *The Journal of KING Computing*, vol. 17, no. 4, pp. 13–22, 2021.

[3] G. Son, W. Lee, K. Han, and S. Kyeong, "The study of feature vector generation and emotion recognition using EEG signals," *The Journal of KING Computing*, vol. 16, no. 2, pp. 72–79, 2020.

[4] M. J. Lim and J. H. Shin, "Continuous emotion recognition method applying emotion dimension," in *2021 Spring Conference of KISM*, vol. 10no. 1, pp. 173–175, 2021.

[5] P. Ji-Eun, S. Sun-Ju, and S. Jin-Hun, "Classification and intensity assessment of Korean emotion expressing idioms for human emotion recognition," *Journal of the Ergonomics Society of Korea*, vol. 31, no. 5, pp. 617–627, 2012.

[6] S. Kang, H. Park, and J. Seo, "Emotion classification of user's utterance for a dialogue system," *Korean Journal of Cognitive Science*, vol. 21, no. 4, pp. 459–480, 2010.

[7] S. J. Park, *Sensitivity Identification Method for New Words of Social Media Using the naïve Bayes Classification*, vol. 8, Chosun University, 2018.

[8] M. H. Yi, M. J. Lim, H. Ko, and J. H. Shin, "Method of profanity detection using word embedding and LSTM," *Mobile Information Systems*, vol. 2021, 9 pages, 2021.

[9] M. Purver and S. Battersby, "Experimenting with distant supervision for emotion classification," in *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics*, pp. 482–491, 2012.

[10] Y. Kim, "Convolutional neural networks for sentence classification," in *Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 1746–1751, 2014.

[11] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[12] D. Shin, *Emotion Classification in Dialogue Utterances Using CNN-LSTM Complex Model*, vol. 2, Korea University, 2017.

[13] *From model-centric to data-centric AI development*, 2021, November 2021, https://jiho-ml.com/weekly-nlp-35/.

[14] M. Firdaus, H. Chauhan, A. Ekbal, and P. Bhattacharyya, "MEISD: a multimodal multi-label emotion, intensity and sentiment dialogue dataset for emotion recognition and sentiment analysis in conversations," in *Proceedings of the 28th International Conference on Computational Linguistics*, pp. 4441–4453, 2020.

[15] K. I. M. Yanghoon, L. E. E. Hwanhee, and J. U. N. G. Kyomin, "AttnConvnet at SemEval-2018 task 1: attention-based convolutional neural networks for multi-label emotion classification," 2018, https://arxiv.org/abs/1804.00831.

[16] I. Ameer, N. Ashraf, G. Sidorov, and H. Gómez Adorno, "Multi-label emotion classification using content-based features in twitter," *Computación y Sistemas*, vol. 24, no. 3, pp. 1159–1164, 2020.

[17] X. Zeng, Q. Chen, S. Chen, and J. Zuo, "Emotion label enhancement via emotion wheel and lexicon," *Mathematical Problems in Engineering*, vol. 2021, 11 pages, 2021.

[18] S. Aman and S. Szpakowicz, "Identifying expressions of emotion in text," in *International Conference on Text, Speech and Dialogue. Lecture Notes in Computer Science*, vol. 4629, pp. 196–205, 2007.

[19] K. Lee, J. Baik, and S. Lee, "Estimating a pleasure-displeasure index of word based on word similarity in SNS," *Journal of Computing Science and Engineering*, vol. 20, no. 3, pp. 159–164, 2014.

[20] J. Kim, T. Hong, and P. Kim, "Replacing out-of-vocabulary words with an appropriate synonym based on Word2VnCR," *Mobile Information Systems*, vol. 2021, 7 pages, 2021.

[21] S.-Y. Kim and H. Yu, "Korean multiple sensibility analysis technique of SNS unstructured data," in *2018 Summer Conference of KACE*, vol. 22no. 2, pp. 147–149, 2018.

[22] M. Kang, *Sentiment analysis in text mining using ensemble-HMM*, Hanyang University, 2016.

[23] "Open dictionary pro," *Naver open dictionary PRO*, 2021, September 2021, https://open-pro.dict.naver.com/.

[24] "List of New Internet Terms in Korea," *Wikipedia*, 2021, September 2021, https://ko.wikipedia.org/wiki/%EB%8C%80%ED%95%9C%EB%AF%BC%EA%B5%AD%EC%9D%98_%EC%9D%B8%ED%84%B0%EB%84%B7_%EC%8B%A0%EC%A1%B0%EC%96%B4_%EB%AA%A9%EB%A1%9D.

[25] "Abbreviation," *namu wiki*, 2021, November 2021, https://namu.wiki/w/%EC%A4%84%EC%9E%84%EB%A7%90?from=%EC%A4%80%EB%A7%90.

[26] "Standard Korean dictionary," *National Institute of Korean Language*, 2021, October 2021, https://stdict.korean.go.kr/main/main.do.

[27] M. Lim, M. Yi, and J. Shin, "Performance improvement method of deep learning-based emotion recognition," *The Journal of KING Computing*, vol. 17, no. 5, pp. 88–95, 2021.

[28] "Cognitive Technology-Language Intelligence," *AI Hub*, 2021, February 2021, https://aihub.or.kr/keti_data_board/language_intelligence.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 5593147, 17 pages
https://doi.org/10.1155/2022/5593147

Research Article

Opinion Mining-Based Term Extraction Sentiment Classification Modeling

Tae Yeun Kim [1] **and Hyoung Ju Kim** [2]

[1] *National Program of Excellence in Software Center, Chosun University, Gwangju, Republic of Korea*
[2] *Department of Computer Engineering, Chosun University, Gwangju, Republic of Korea*

Correspondence should be addressed to Hyoung Ju Kim; hyoungjukim@chosun.ac.kr

Received 24 January 2022; Accepted 2 April 2022; Published 27 April 2022

Academic Editor: Chun-Wei Tsai

The spread of social media has accelerated the formation and dissemination of user review data, which contain subjective opinions of users on products, in an e-commerce environment. Because these reviews significantly influence other users, opinion mining has garnered substantial attention in analyzing the positive and negative opinions of users and deriving solutions based on these analytical results. Terms that include sentimental information and used in user reviews serve as the most crucial element in sentimental classification. In this regard, it is crucial to distinguish the most influential terms in user reviews. This study proposed a document-level sentiment classification model based on the collection and application of user reviews generated in an e-commerce environment. Here, a term information extraction method was applied to the proposed model to select core terms, classify the selected terms according to parts of speech (POS), determine terms that can increase information power and influence, and adopt these terms in opinion mining research, based on SVM, SVM+, and SVM+MTL techniques. The results obtained from evaluating the proposed model indicate that it exhibited excellent sentiment analysis performance. The proposed model is expected to be effectively utilized in providing enhanced services for users and increasing competitiveness in the e-commerce environment.

1. Introduction

The development of information and communications technology has accelerated the advancement of social media, which facilitates communications among users, accumulates an enormous amount of knowledge, and enables users to easily exchange knowledge in online environments. Social media have provided online users with the ability to obtain and analyze data beyond the limitations of time and space. Users who purchase products in e-commerce environments often write reviews on these products, which serve as crucial sources of information for other online users who might be interested in purchasing products related to these reviews. In other words, these reviews are utilized as essential information for providing enhanced services to e-commerce users. However, owing to the rapid increase in the number of online reviews due to big data, it has become difficult to obtain useful information by analyzing these reviews. Hence, opinion mining based on text mining has garnered significant attention, as this technique can present the results obtained from analyzing sentiments included in a considerable amount of user reviews [1–3].

Previous studies on opinion mining applied to online reviews primarily focus on classifying the reviews or opinions of users on products generated in an e-commerce environment, according to sentimental polarity (positive or negative sentiments) [4–6]. Certain online reviews might provide users with incorrect information, such as advertisements or intentionally distorted information. Therefore, research on opinion mining should utilize core terms that represent the subjective opinions of users in the reviews. In other words, it is necessary to extract core terms, which function as the most significant elements in sentiment classification, from the entire terms included in reviews and determine core terms that can represent documents. Hence, numerous studies have been conducted on opinion mining based on term information extraction [7–9]. Because the roles of terms including sentimental information perform differently according to POS in

these studies, such roles should be analyzed according to POS [10]. Nouns tend to reflect properties and sentiments related to products, while adjectives and verbs tend to include information on the subjective opinions and evaluation of users. In addition, adverbs are applied as various types of expression methods and modification functions in the process of document classification [11, 12].

Existing studies have also adopted data mining techniques that can predict the sentiments of users by using text documents to analyze their sentiments based on opinion mining from a different perspective [13–15]. Although several researchers have utilized data mining techniques to derive enhanced prediction performance, they have failed in addressing the inherent limitations of these techniques. To address this problem, a number of studies improved prediction performance by integrating data mining techniques. Increasing attention has also been paid to a research, which proposed an integrated model considering the characteristics of each data mining technique to achieve outstanding prediction performance instead of performing the duplicated application of a simple model.

The objectives of this study are outlined as follows: First, in sentence-level analysis, a number of studies have been carried out that utilized various parts of speech to determine the main part of speech (POS) and conduct sentiment analysis. In particular, based on the approach of semantic analysis, adjectives bear the most sentiment information among different POS, and they are used as the core terms for classification of polarity. However, in a document-level analysis, optimal selection of terms is instrumental since selection of input variables determines the predictive performance of the model and has a significant impact on the classification of the sentiment of the entire document. Also, it has been reported that considering the entire document using various POS is more effective than considering only adjectives in a fragmented approach. Therefore, in this study, terms are extracted focusing on the four POS (adjectives, adverbs, verbs, and nouns) and the extracted terms are used as input variables for a model for document-level sentiment classification. Also, considering the superiority of adjectives in terms of sentiment bearing, we aim to examine whether there are differences in the relationship between the four POS and the prediction performance of different model types.

Second, in the document-level analysis, using an optimal number of terms as input variables for the model is important for effective classification of the document. If too many terms are used as input variables, it is highly likely that there will be missing values and high redundancy; thus, it is difficult to determine if sentiment information is included, and also, the efficiency of the analysis is reduced. Therefore, in this study, a variety of term information extraction techniques are used in order to select effective and efficient terms as input variables to be used in the model and compare the results and performance between different techniques.

Third, in opinion mining research, when data mining techniques are used, the selection of the variables used in the model and the method applied according to the characteristics of the data plays the central role. In addition, since different results are obtained depending on the use of various data mining techniques, the model performance may vary. That is, it is important to explore and select the most efficient model by adopting various data mining techniques. Therefore, in this study, SVM+ and SVM+MTL techniques, which have not been implemented in opinion mining research to date, are applied, and models for document-level sentiment classification are developed to compare the prediction performance of the developed models.

Lastly, in opinion mining research, different single models are combined to develop a new integrated model, which serves as a method for further improvement of the predictive performance of a single model. In general, an integrated model is applied with various methods to improve the performance of a single model, and a number of studies have been carried out on the development and application of the integrated models. Thus, in this study, through the results of the models for document-level sentiment classification using SVM, SVM +, and SVM+MTL techniques, we propose an integrated data mining model, a type of an integrated model, and comparatively analyze the difference in predictive performance between the single model and the integrated model.

To this end, core terms were selected by adopting a term information extraction technique, and then, the selected terms were classified according to POS to determine the final terms that can increase information power and influence. Subsequently, approach presented an integrated sentiment classification model by applying SVM+ and SVM+MTL techniques to the research on opinion mining. To evaluate the performance of the proposed model, this study collected 80 000 user reviews on movies, games, music, and books on Amazon (http://Amazon.com), eliminated unnecessary terms, and extracted terms based on POS tagging. It also calculated the values of document frequency, term frequency-inverse document frequency (TF-IDF), information gain, and chi-square statistic to rank the extracted terms based on the values derived. Subsequently, the top 20 optimal terms were determined according to the categories and classified sentiments included in the optimal terms selected by utilizing SVM, SVM+, and SVM+MTL techniques. The obtained evaluation results indicate that the sentiment classification model based on the chi-square statistic exhibited the most excellent prediction performance among the other models. In addition, the proposed sentimental classification model effectively analyzed sentiments included in terms. It is expected that the proposed term-based sentiment classification model can be effectively used to improve services, ensure competitiveness, and provide enhanced services for users in the e-commerce environment. The remainder of this paper is presented as follows: Section 2 reviews previous studies, while Section 3 introduces the proposed sentiment classification model based on terms at the document level. Furthermore, Section 4 presents the results obtained from conducting experiments on the proposed model. Finally, Section 5 provides conclusions and future research directions.

2. Related Research

2.1. Opinion Mining for Sentiment Classification. Opinion mining, which was developed from a data mining technique for document classification, refers to the process of

extracting, classifying, and analyzing the opinions of users expressed in various media [16]. This technique, also called sentiment analysis, analyzes sentiments expressed in text and a state of positive or negative expressions on a certain subject included in the text. Owing to such characteristics, the field of opinion mining for social network analysis has garnered increasing interests [17, 18]. Sentiment analysis generally determines the positive or negative tendency of an entire document by comprehensively analyzing opinions included in the document. To achieve this objective, the following analyses are performed at two different levels [19]: First, a document-level analysis is conducted to classify documents that express positive or negative opinions or sentiments. In this analysis, opinions on a certain subject are regarded as the basic unit of information. This analysis facilitates fast and comprehensive determination, as certain terms included in a document are selected according to frequency. Second, a sentence-level analysis is conducted to classify sentiments expressed in sentences that are included in a document. This analysis enables people to determine whether a positive or negative opinion is expressed, based on sentences in the document. Specifically, this analysis technique determines an opinion based on sentences, extracts sentences including sentiments, and examines paragraphs or phrases, considering different terms [20].

Sentiment classification is the process of extracting expressional terms in a document or a sentence from reviews or comments, identifying sentiments in various forms of documents, and classifying these sentiments according to polarity (positive or negative sentiments) [21]. Studies on sentiment classification adopt an approach based on vocabulary, machine learning (ML), or a method integrating vocabulary and ML [19]. In the vocabulary-based approach, a model analyzes the intentions of users based on positive and negative terms and the implications of these terms. In this process, it primarily adopts previously edited or well-known sentiment terms [22, 23]. In the ML-based approach, a model trains behavior or pattern data based on data collected and then predicts the future by classifying or analyzing these data. This process applies algorithms and linguistic variables used in the ML-based approach and depends on algorithms to solve document classification problems related to sentiment analysis. In the integrated approach, a model adopts both the vocabulary- and ML-based approaches. In general, sentiment terms serve as crucial elements in this approach. Figure 1 illustrates various approach techniques for sentiment classification.

2.2. Term Information Extraction Technique. Term information extraction is the most significant methodology for determining terms that contain the greatest amount of information on document classification and establishing a classification model to solve sentiment classification problems at the document level in opinion mining research [24, 25]. Terms used in reviews or comments contain sufficient information that can be adopted to classify linguistic properties and documents, thereby performing a crucial role in sentiment classification, and serving as an important basis for determining the main sentiment in a review. It is crucial to determine core terms, as these terms exert significant influence on determining the polarity of the entire document.

2.2.1. Document Frequency. Document frequency refers to the number of documents that include certain terms among the entire documents. In other words, it is a ratio of documents that contain terms that were used over a certain frequency in documents. This technique is simple and requires less calculation than other techniques. It also improves the accuracy of document categorization by eliminating low-frequency terms and discarding words under the assumption that terms that occur at low frequency do not contribute to document classification [26, 27]. Equation (1) defines document frequency, where N and i denote the entire number of documents and terms, respectively.

$$DF_i = \frac{n_i}{N}. \tag{1}$$

2.2.2. Information Gain. Information gain is one of the most frequently used techniques in opinion mining research, which is known to be more influential than document frequency. This technique calculates the amount of information included in terms according to categories by considering both the occurrence and absence frequencies of terms, respectively. It is applied as an evaluation standard for measuring the usefulness of terms in the ML field. Information gain is calculated via the following process. First, a set of the entire documents is divided into subsets. Second, entropy, which changes before and after the use of a certain term, is calculated. Third, the probability of a target is divided and calculated according to categories, to derive an information value on the certain terms. Fourth, the gain is calculated to derive normalized gain [28, 29]. Equation (2) defines information gain, where t denotes a term.

$$\mathrm{IG}(\mathrm{t}) = \sum_{i=1}^{c} P(c_i) \log \frac{1}{p(c_i)} - \sum_{\mathrm{t} \in \{\mathrm{t}_p, \overline{\mathrm{t}_p}\}} P(\mathrm{t}) \sum_{i=1}^{c} P(\mathrm{t} \mid c_i) \log \frac{1}{P(\mathrm{t} \mid c_i)}, \tag{2}$$

where t_p is the number of times the term t occurs in the document and $\overline{\mathrm{t}_p}$ is the number of times the term t does not occur in the document.

2.2.3. TF-IDF. TF-IDF is a method that can represent the significance of a term in individual documents. This method is frequently adopted because of its advantages, which includes its similarity with document frequency, simple calculation processes, and excellent performance [30]. The significance of a term is proportional to the number of times the term appears in a document and inversely proportional to the number of entire documents that contain the term [31, 32]. The TF-IDF technique was developed as a document representation method to evaluate the relative significance of terms in a document. This technique is frequently used in analyses based on the unit of a phrase or a paragraph [33]. It calculates the significance of a term by applying a

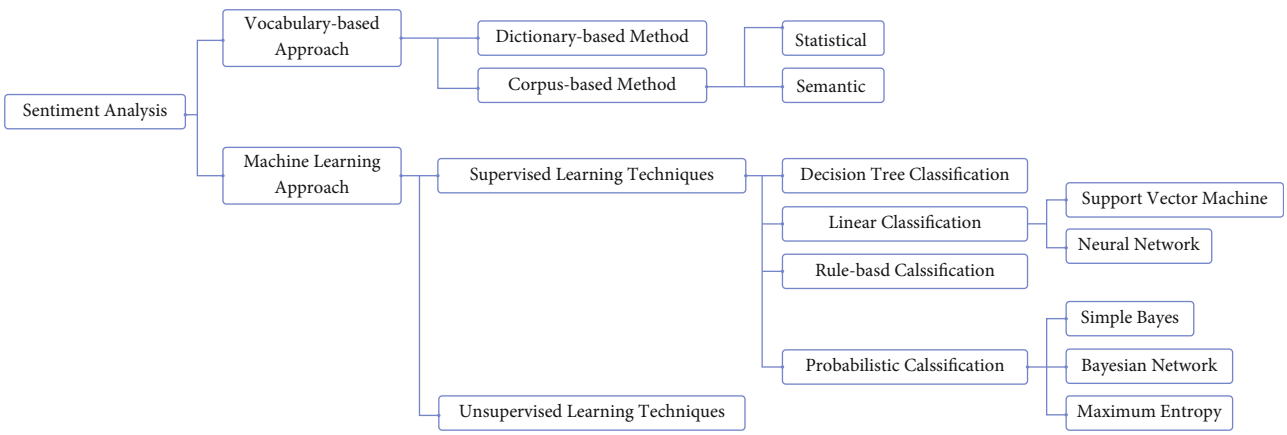


Figure 1: Sentiment classification technique.

weight to the term and employing TF, which refers to the number of times a word appears in a document, and DF, which is the number of the entire documents containing the term. Equations (3) and (4) express the TF-IDF calculation processes based on a term (t) and document (d).

$$\mathrm{TF-IDF}_{t_i d} = \mathrm{TF}_{t_i d} * \mathrm{IDF}_t, \tag{3}$$

where

$$\mathrm{IDF}_t = \log \frac{N}{\mathrm{DF}_t}, \tag{4}$$

where N denotes the total number of documents, $\mathrm{TF}_{t_i d}$ the number of documents (d) and terms (t) that appear, and DF_t the total number of documents containing the term (t).

TF alone cannot be used to calculate the significance of a term because a high-frequency term is likely to be meaningless in a document. Consequently, both TF and IDF are reflected in the calculation process. IDF divides the number of the entire documents by the number of documents that contain the given term. Therefore, a term that appears in numerous documents has a low IDF value. In contrast, a term that is biased to certain documents has a high IDF value.

2.2.4. Chi-Square Statistic. The chi-square statistic, which analyzes a correlation between categorical variables, is the most frequently used technique among cross-tabulation analysis techniques. Specifically, the value of a correlation is "1" if a term appears more than once in the entire documents; otherwise, it is "0." In other words, this technique is used to identify the significance between a term (t_i) and a category (C_j).

$$x^2 = \frac{N \times (\mathrm{AD} - \mathrm{CB})^2}{(\mathrm{A}+\mathrm{C}) \times (\mathrm{B}+\mathrm{D}) \times (\mathrm{A}+\mathrm{B}) \times (\mathrm{C}+\mathrm{D})}, \tag{5}$$

where
$N = \mathrm{A} + \mathrm{B} + \mathrm{C} + \mathrm{D}$ (N the total number of documents),
$i = 1, 2, \cdots, n$ (n the total number of terms used),
$j = 1, 2, \cdots, m$ (m the total number of categories).

The chi-square statistic is similar to mutual information; however, the former exhibits better performance for term extraction than the latter, owing to the standardized value of the former. Equation (5) defines the chi-square statistic (x^2) between C_j and t_i in a range of the consistent properties applied. Table 1 presents the document frequency between C_j and t_i [34, 35].

2.3. Data Mining Techniques

2.3.1. SVM. The SVM technique, which was developed by Vapnik based on the statistical learning theory, exhibits outstanding performance in solving classification problems [36]. While most learning algorithms focus on empirical risk minimization, this technique aims at structural risk minimization. Because this technique can easily generalize classification problems, it has been applied in various fields. The SVM technique can also match nonlinear problems on input space with linear problems on specific high-dimensional space, thereby facilitating convenient mathematical analysis [37, 38]. Kernel functions generally used for the SVM technique include the polynomial kernel and Gaussian radial basis function (RBF). These functions form input datasets that can be linearly separated in specific space by mapping data in high-dimensional space [39]. Figure 2 illustrates the decision boundary and margin of the SVM technique.

2.3.2. SVM+. Vapnik also proposed SVM+, which is an optimal SVM-based technique for managing learning with structured data (LWSD) or learning using privileged information (LUPI) to obtain hidden group data from training data [36]. In addition, Vapnik defined a slack variable in each group based on a modification function by considering group data and then presented a method of mapping input vectors in two different Hilbert spaces simultaneously. Figure 3 presents the mapping method.

The slack variable is restricted by a correction variable under the applied SVM+ environment. A mapping sample in a modification space should be placed on one side of a correspondence function. The SVM+ technique simultaneously maps data in both decision and modification spaces. When a

Table 1: Document frequency between term (t_i) and category (C_j).

	Documents in which category (C_j) occurred	Documents where category (C_j) does not occur
Documents with the term (t_i)	A	B
Documents that do not have the term (t_i)	C	D

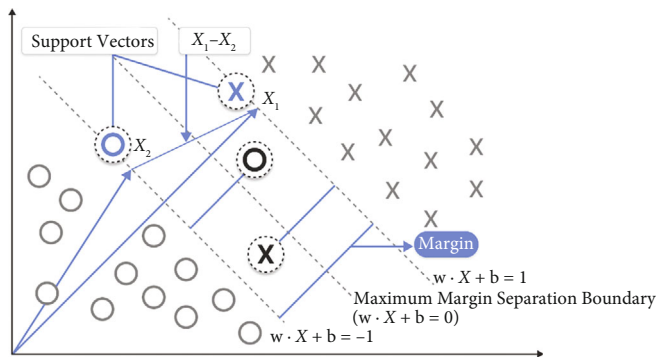


Figure 2: Decision boundary and margin of the SVM.

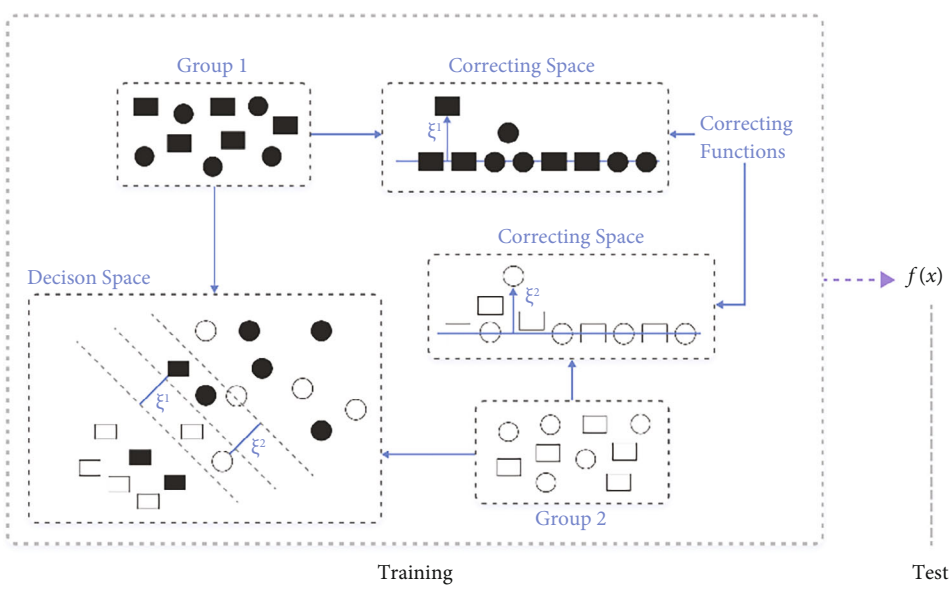


Figure 3: Processing using the SVM+.

slack modification function is defined in the same decision space, a decision function is defined in the decision space. Data in other groups are mapped in the same decision space, whereas a modification function can be defined in the same modification space or a different modification space, owing to the application of different correction variables in different groups [40]. Equations (6) and (7) define the SVM+ technique:

$$\min_{W,b\in R,\xi^r} \phi\left(W, b, \xi^r\right) = \frac{\|W\|^2}{2} + \frac{r}{2}\|W_r\|^2 + C\sum_{r=1}^{t}\sum_{i\in T_r}^{l} \xi_i^r \quad (6)$$

Subject to:

$$y_i\left(W^T\phi(X_i + b)\right) \geq 1 - \xi^r, i \in T_r,\ r = 1, \cdots, t \quad (7)$$

$$\xi^r \geq 0, i \in T_r, r = 1, \cdots, t \quad (8)$$

$$\xi^r = \left(W^T\phi_r(X_i) + d_r\right), i \in T_r, r = 1, \cdots, t. \quad (9)$$

R denotes the number of groups, while y adjusts the capacity of modification space and importance of capacities of the modification space. ξ^r represents the slack variables of each group and differs according to groups. A penalty coefficient (C) maintains a balance between the complexity of a model

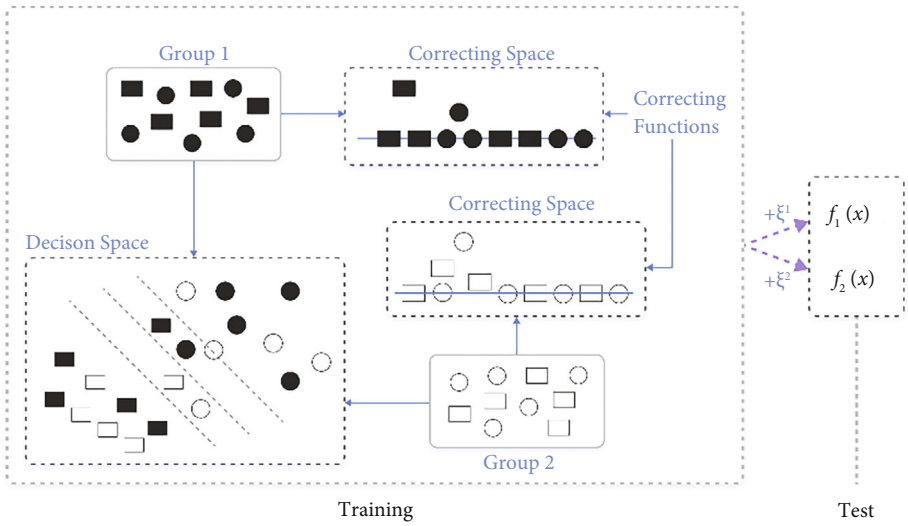


Figure 4: Processing using the SVM+MTL.

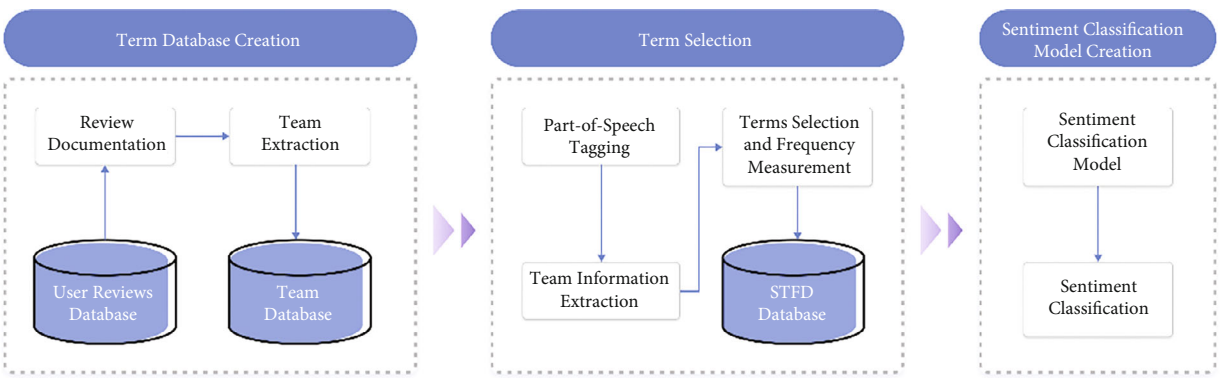


Figure 5: Term-based sentiment classification model at the document level.

(i.e., training capability) and the error rate of a sample. That is, the SVM+ technique can be divided into two parts, where W denotes the capability of the decision space and W_r denotes the capability of the crystal space. However, since W_r does not determine the size of the margin, the SVM+ technique adds $(r/2)\|W_r\|^2$ compared to the SVM technique.

The specific term information extraction processes of the SVM+ technique are consistent with those of the SVM technique. In the term information extraction process, special attention should be paid to the following issues:

(1) The SVM+ technique should divide training data by selecting appropriate group data according to professional or general knowledge, instead of performing statistical analysis

(2) The amount of samples can significantly influence the response effects of the SVM+ technique. Hence, the SVM technique might exhibit the best performance when the amount of samples is small, owing to a small number of parameters to be adjusted

(3) The SVM+ technique can significantly influence the selection of both kernel functions and parameters

(4) The SVM+ technique is consistent with the SVM technique when each training vector among the entire data belongs to independent groups that are not related to each other

(5) The SVM+ technique requires a longer time for experimental calculation than the SVM technique

Ribeiro et al. adopted the SVM+ method as a model for predicting corporate bankruptcy based on financial data obtained in France and then examined the prediction performance of the model by considering authority information [41]. Specifically, they analyzed the prediction performance of the corporate bankruptcy prediction model by using

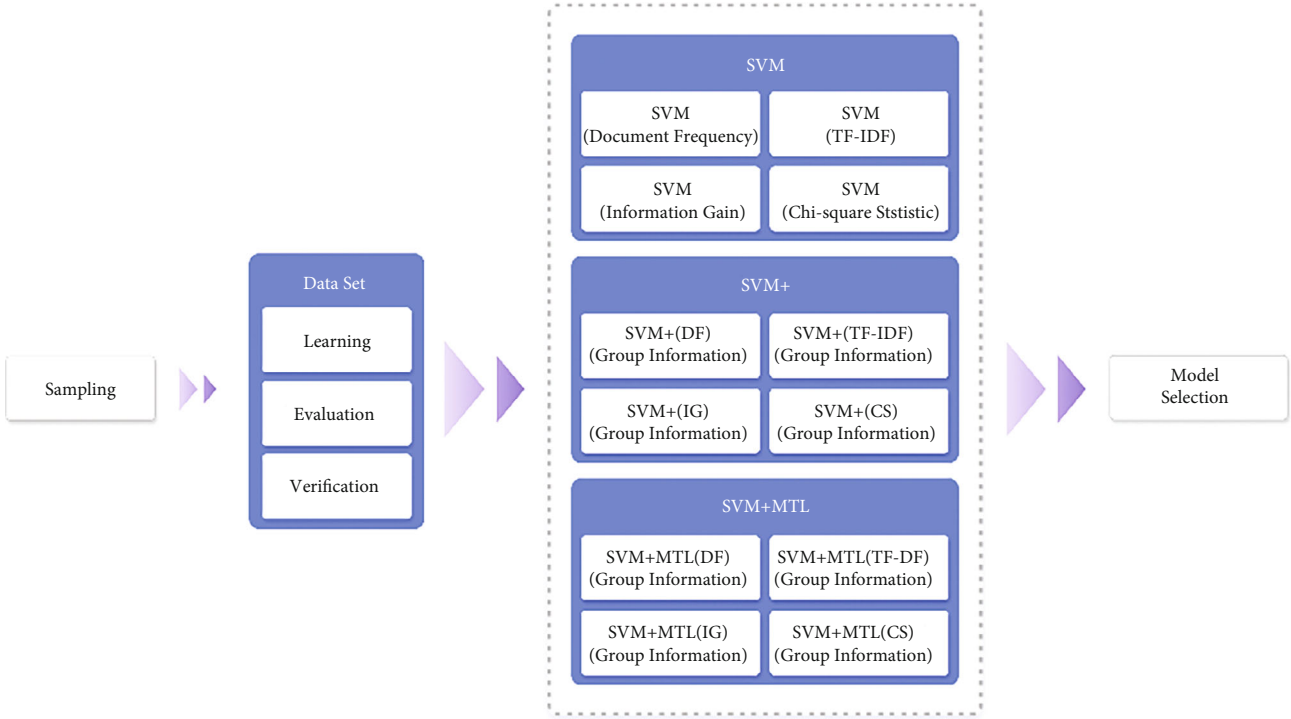


Figure 6: Sentiment classification model.

heterogeneous information grouped according to the size categories of 30 companies relative to financial ratios, number of employees, and annual profits. Based on the results obtained from their analyses, they argued that the SVM+ technique always exhibits a more optimal performance than the SVM and SVM+MTL techniques, regardless of kernel functions. Moreover, the corporate bankruptcy prediction model that applies the Gaussian RBF as a kernel function exhibited better performance than the model that applies a linear kernel function based on the application of structured information in training data.

Serra-Toro et al. elucidated the relationship between LUPI and the SVM+ technique, thereby indicating that the performance of this technique depends on a subtle relationship between normal data and authority information [42]. They also stated that a randomly generated variable can also perform a crucial role when it contains authority information on a specific problem.

In addition, they argued that the performance of the SVM+ technique can exceed that of the SVM technique according to data preprocessing, data set segmentation, validation protocols, experiments, and a range of parameters and search procedures.

2.3.3. SVM+MTL. Liang et al. developed an SVM+MTL technique by combining the SVM+ algorithm and the multitask learning (MLT) technique. Specifically, the SVM+ technique was adopted to define different decision functions for each group and identify a relationship between groups [43]. To solve problems related to the MTL technique, the SVM+MTL technique also maps data in two Hilbert spaces simultaneously, similar to the SVM+ technique. Equation (10) defines the decision function.

$$f_r(x) = W^T \phi_Z(X_i) + b + W_r^T \phi_{Z_r}(X_i) + d_r \; r = 1, \cdots, \mathrm{t}. \quad (10)$$

The SVM+MTL technique has the following advantages over the SVM+ technique:

(1) Group data also exist in prediction data

(2) A slack variable indicates an error in the entire model after a modification function is added, instead of being defined by the modification function

(3) A decision function has a correction term unlike those used for the SVM and SVM+ techniques

Liang et al. proposed an MTL algorithm for the SVM +MTL technique. This study utilized group data to solve MTL-related problems by applying the SVM+ technique [43]. Via this process, the prediction accuracy of the proposed model increased. Moreover, this study elucidated the relationships between SVM+ and SVM+MTL techniques and compared the performance of sentiment classification models that apply these techniques by using comprehensive datasets to empirically compare the performance of these models. The results obtained from the comparison indicated that the sentiment classification model that applies the XVM +MTL technique exhibited excellent prediction performance when the amount of data was sufficient. However, when the

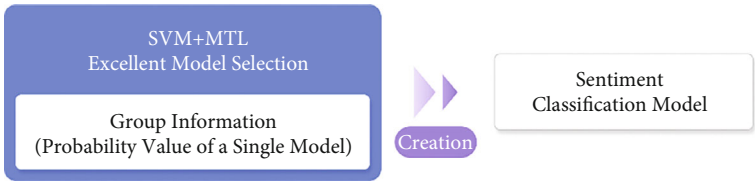


Figure 7: Sentiment classification model construction.

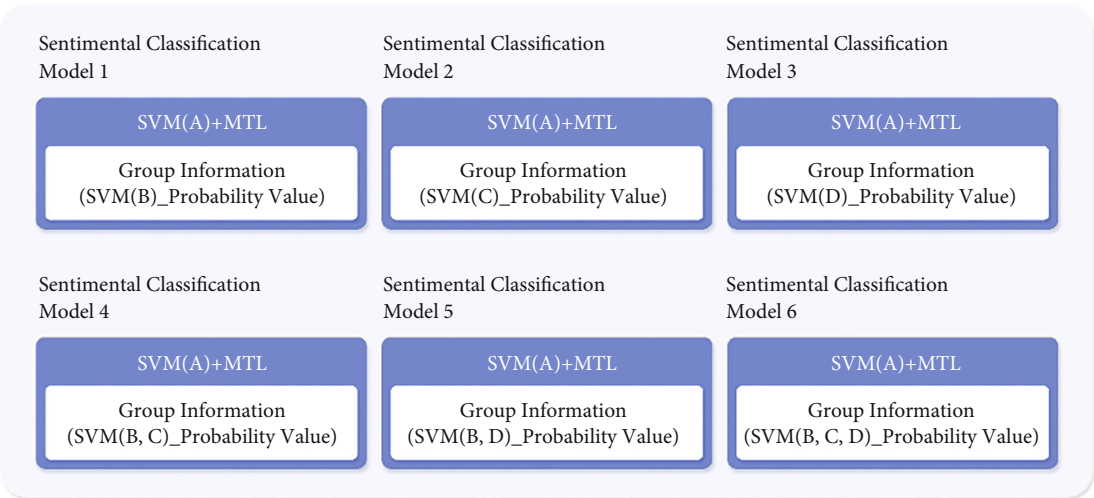


Figure 8: Application result of sentiment classification model.

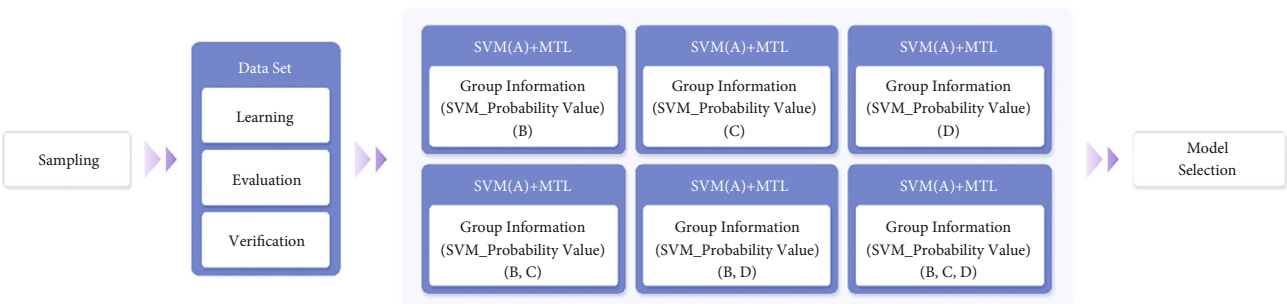


Figure 9: Document-level sentiment classification model.

amount of data was insufficient, sentiment classification models that apply the SVM or SVM+ technique exhibited outstanding performance. Figure 4 presents processing using the SVM+MTL..

3. Sentiment Classification Modeling

3.1. Sentiment Classification Model. The sentiment classification model proposed in this study determines and classifies core terms, which contain a significant amount of information and influence the entire document, for prediction. Figure 5 illustrates the three stages required to establish a term-based sentiment classification model at the document level, which include the formation of a term database (DB), selection of terms, and establishment of a sentiment classification model.

3.1.1. Processes of a Term DB Formation. The processes of a term DB formation are as follows: First, customer review data generated on the web are crawled, collected, and stored in a customer review DB.

Second, these data are preprocessed, owing to the unstructured data. In general, unstructured data obtained from the web should be converted to structured data to be used in an experiment. In other words, a customer review can accurately reflect a subjective opinion of an individual only when this review is considered as a document.

Third, structured data obtained via preprocessing are adopted to classify reviews as positive and negative reviews.

Table 2: Number of terms extracted by item.

	Movie	Game	Music	Book
Number of documents	30,000	30,000	20,000	20,000
Number of terms collected	77,222	78,523	66,190	85,456
Adjective	132	146	117	140
Noun	400	402	319	309
Adverb	57	55	61	63
Verb	161	147	169	154

This classification process is required to identify positive or negative tendencies of customers in their opinions. Hence, this study developed a simple application program based on the Microsoft foundation class (MFC) to document individual customer reviews.

Fourth, the entire terms included in the documents are extracted. The extracted terms contain stop words, which appear frequently, exhibit weak information power, and do not provide special information. Hence, it is necessary to eliminate such stop words. Stop words that appear frequently in a document consist of articles (e.g., a, an, and the), conjunctions (e.g., that, and, and when), pronouns (e.g., I and you), and the verb to be (e.g., is and are). The process of eliminating stop words is performed to select highly influential terms, develop a term DB based on the selected terms, and store the DB.

3.1.2. Term Selection. Subsequently, four POS (adjectives, adverbs, verbs, and nouns) are utilized to tag the terms stored in the term DB. Information on POS serves as a crucial index for deriving sentimental information from terms and can also be effectively used to determine terms with significant information power and influence by accurately analyzing the characteristics of terms. However, when POS tagging is not performed, the term-selection process requires numerous manual tasks, including term selection based on the opinions of experts and a considerable amount of cost and time.

In this study, an experiment was conducted in which a Stanford POS Tagger program, a grammatical structure analysis tool, was used to apply 4 POS tagging to the entire terms and measure the frequency of occurrence of these terms in the entire documents. Subsequently, term information extraction techniques were adopted to calculate term information values. Document frequency, TF-IDF, information gain, and chi-square statistic were applied as term information extraction techniques in this study. After the information values of each term were calculated, these terms were ranked according to term information extraction techniques to determine the terms to be used as input variables and store the selected terms in a sentiment terms for documents database (STFD).

3.1.3. Sentiment Classification Model Formation. The selected terms presented in Figure 6 are adopted as input variables, and the dataset including these terms is divided into three different datasets for training, evaluation, and verification. Next, SVM, SVM+, and SVM+MTL techniques for term information extraction are applied to establish document-level sentiment classification models.

3.2. Term-Based Sentiment Classification Model. The term-based sentiment classification model proposed in this study is an integrated model that exhibits better performance than individual sentiment classification models, beyond the limitations of these models.

Figure 6 presents the integrated sentiment classification model based on the SVM+MTL model, which reflects the experimental results of individual sentiment classification models based on the aforementioned term information extraction techniques, to exhibit more excellent sentiment analysis performance. The integrated model based on the SVM+MTL technique utilizes and trains group data as probability values (prediction performance rates) obtained in the experimental results on individual sentiment classification models.

Figure 7 presents the processes of establishing the integrated sentiment classification model proposed in this study. These processes are as follows:

(1) The experimental results based on sentiment classification models applying SVM, SVM+, and SVM +MTL techniques are derived, as illustrated in Figure 6

(2) The model, which exhibited the best result among sentiment classification models that apply the SVM +MTL technique, is selected

(3) Group data used for the SVM+MTL technique are used as probability values for the remaining three models, except for the best model

(4) Models are divided into specific models with group data applied in various ways for training, as illustrated in Figure 8

(5) Prediction results derived by the sentiment classification model are applied to the decision-making processes

When A, B, C, and D refer to document frequency, TF-IDF, information gain, and the chi-square statistic, respectively, and a model based on A exhibits the best performance in the experimental results based on sentiment classification models that apply the SVM+MTL technique, the sentiment classification model applying the SVM+MTL technique based on A is selected to establish the integrated sentiment classification model. Accordingly, probability values of sentiment classification models that apply the SVM technique based on B, C, and D are calculated and adopted as group data. Consequently, sentiment classification models are subdivided and defined. Figure 8 illustrates the defined models.

Figure 9 presents document-level sentiment classification models. Accordingly, sentiment classification models 1, 2, 3, 4, 5, and 6 are generated.

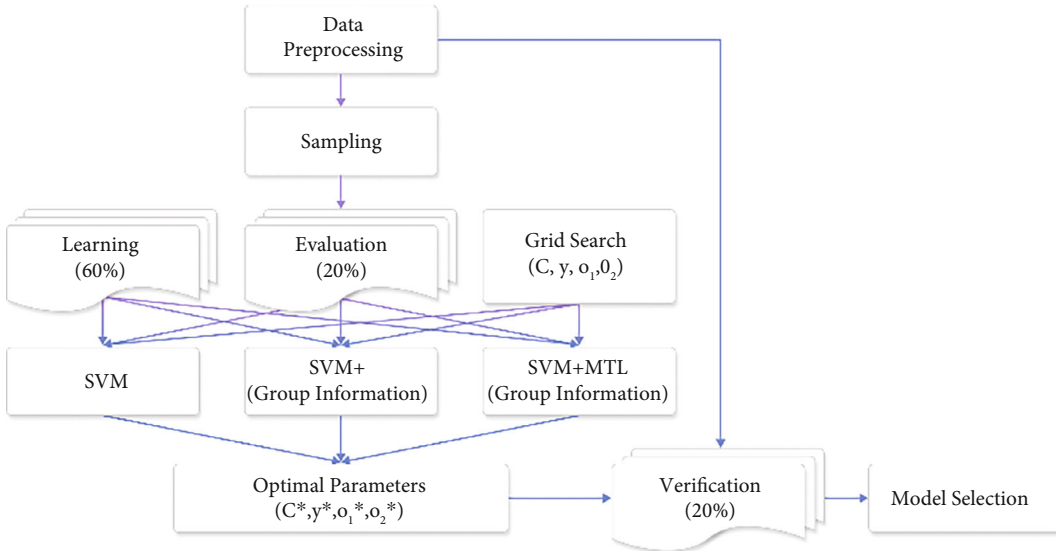


Figure 10: Single model experiment process for sentiment classification.

Table 3: Confusion matrix.

	Predicted positive	Predicted negative
Actual positive	TP (true positive)	FN (false negative)
Actual negative	FP (false positive)	TN (true negative)

4. Sentiment Classification Model Experiment and Results

4.1. Experiment Data. This study employed user review data generated on Amazon (http://amazon.com). Particularly, it collected user review data associated with the categories of movies, games, music, and books, in which users performed enormous purchase activities for relevant products. The data of these user reviews were directly crawled from July 2020 to August 2021, and MFC was utilized to document each user review. Accordingly, an application program for document classification was developed and utilized. First, the list of each item was formed. Then, documents corresponding to Score 4 or higher and documents corresponding to Score 2 or lower were classified and collected as positive and negative documents, respectively. Because documents corresponding to Score 3 were classified as neutral documents, they were excluded. Specifically, 150 movie reviews, 150 games reviews, 126 music reviews, and 134 book reviews were collected. For movie and game reviews, 30 000 review documents were generated for each category. Regarding music and book reviews, 20 000 review documents were generated for each category. A program R was utilized to calculate the occurrence frequency of the remaining terms, except stop words in the entire documents.

Table 2 presents the number of documents for each category, number of terms collected via POS tagging performed by the Stanford POS Tagger program, and results obtained from extracting 1000 terms of 4 POS that appeared frequently. Term information extraction techniques were applied to calculate term information values, and the terms were ranked according to each technique. Via this process, this study selected 20 terms of 4 POS that appeared frequently, according to each technique.

Numerous input variables can trigger problems such as an increase in missing values and difficulties in model management. Therefore, a model with a low number of parameters for input variables is more affected by input variables. A significant number of input variables can also generate variables that are correlated to each other. The elimination of these variables reduces the variance of prediction values, thereby enhancing a model with increased prediction power. Hence, this study selected 20 input variables, which had significant effects on classifying the properties of the entire documents and contained the greatest amount of information on document classification, to be used in the document-level analysis.

The experimental environment of this study is described as follows. Experiments were carried out in "Matlab" to calculate and evaluate experimental results. The specifications of the computer used in this experiment are as follows: Windows 10, Intel® Core™ i5-10500 CPU@3.10GHz 3.10GHz processor, 16.00-GB RAM, and a 64-bit operating system.

4.2. Experiment Design

4.2.1. Term-Based Sentiment Classification Model Experiment Design. To apply balanced data in this experiment, the entire datasets were classified as positive and negative documents. Document-level sentiment classification models that apply SVM, SVM+, and SVM+MTL techniques adopt the Gaussian RBF kernel function and grid search to identify optimal parameters. C denotes the upper limit value and γ represents a parameter that adjusts the weight between decision and modification function, while σ_1 and σ_2 are parameters the Gaussian RBF kernel function in the decision and modification functions, respectively. The

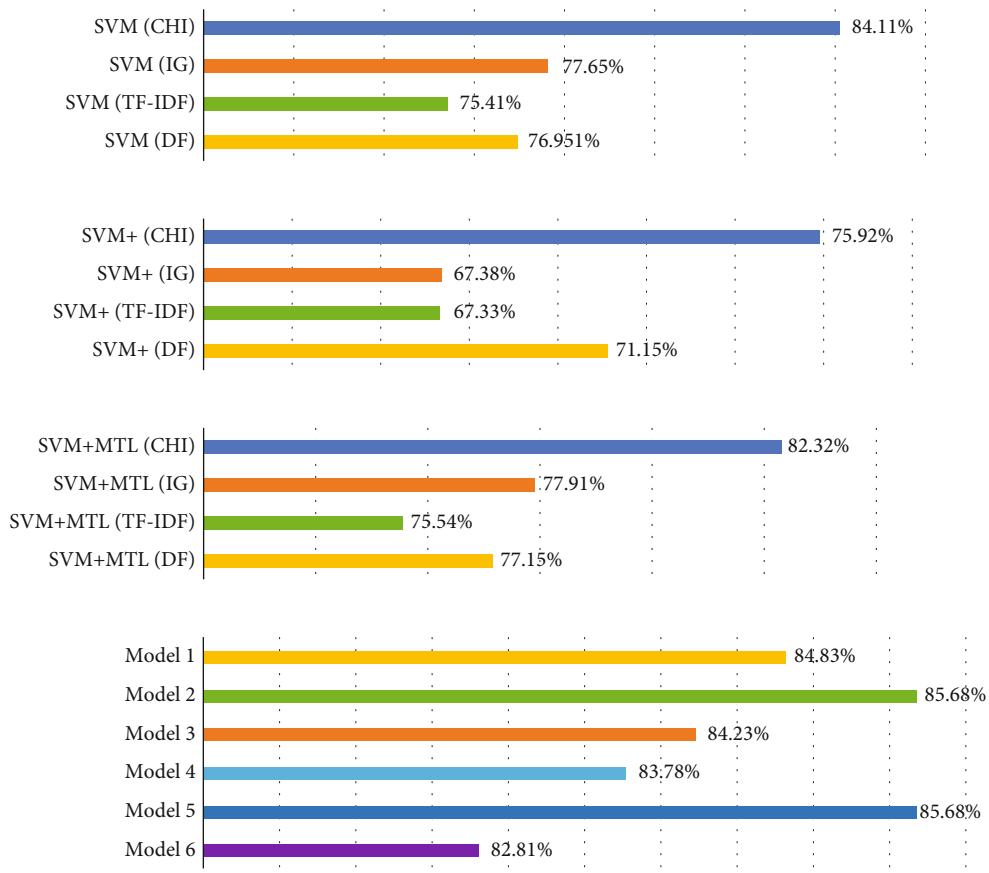


Figure 11: Comparison of overall model performance in movie reviews.

parameters used in this experiment were established as follows: $C = \{1, 25, 50, 75, 100\}$, $\gamma = \{0.001, 0.01, 0.1, 1, 10\}$, and $\sigma_1, \sigma_2 = \{1, 5, 7, 8, 10\}$.

Regarding the group data adopted by the sentiment classification model that applies the SVM+ technique and the sentiment classification model that applies the SVM+MTL technique, information on the year of issue was employed according to items. The confusion matrix was applied to calculate prediction results derived by document-level sentiment classification models that apply SVM, SVM+, and SVM+MTL techniques. Figure 10 presents single model experiment process for sentiment classification.

As presented in Table 3, the confusion matrix shows a relationship between the actual categories and those predicted and classified by models. Accuracy refers to the accuracy of sentiment classification results. With an increase in accuracy, the model improves in its performance. Reproducibility is a ratio of the actual values to the existing sentiment classification results, while precision refers to the accuracy of the prediction. Equations (11), (12), and (13) define accuracy, reproducibility, and precision, respectively.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}, \tag{11}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \tag{12}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}. \tag{13}$$

F-measure is a method that evaluates the validity of actual and predicted data by utilizing the confusion matrix to evaluate the performance of a model in the field of data mining. This method can measure classification performance by considering both precision and reproducibility, including reflecting effects triggered when actual data are inclined to a certain target. It is frequently utilized to verify sentiment classification performance. Equation (14) expresses the process of calculating F-measure:

$$\text{F} - \text{measure} = \frac{2 \times (\text{Precision} \times \text{Recall})}{(\text{Precision} + \text{Recall})}. \tag{14}$$

4.3. Experiment Results. The results obtained from analyzing each sentiment classification model based on movie reviews in experiments are presented as follows: the sentiment classification models that apply the SVM, SVM+, and SVM +MTL techniques based on the chi-square statistic exhibited prediction accuracies of 84.11%, 75.92%, and 82.32%, respectively. Based on these analytical results, it was verified that the sentiment classification model that applies the SVM +MTL technique exhibited the best prediction accuracy among the other single models. Sentiment classification models 2 and 5, which are integrated models proposed in

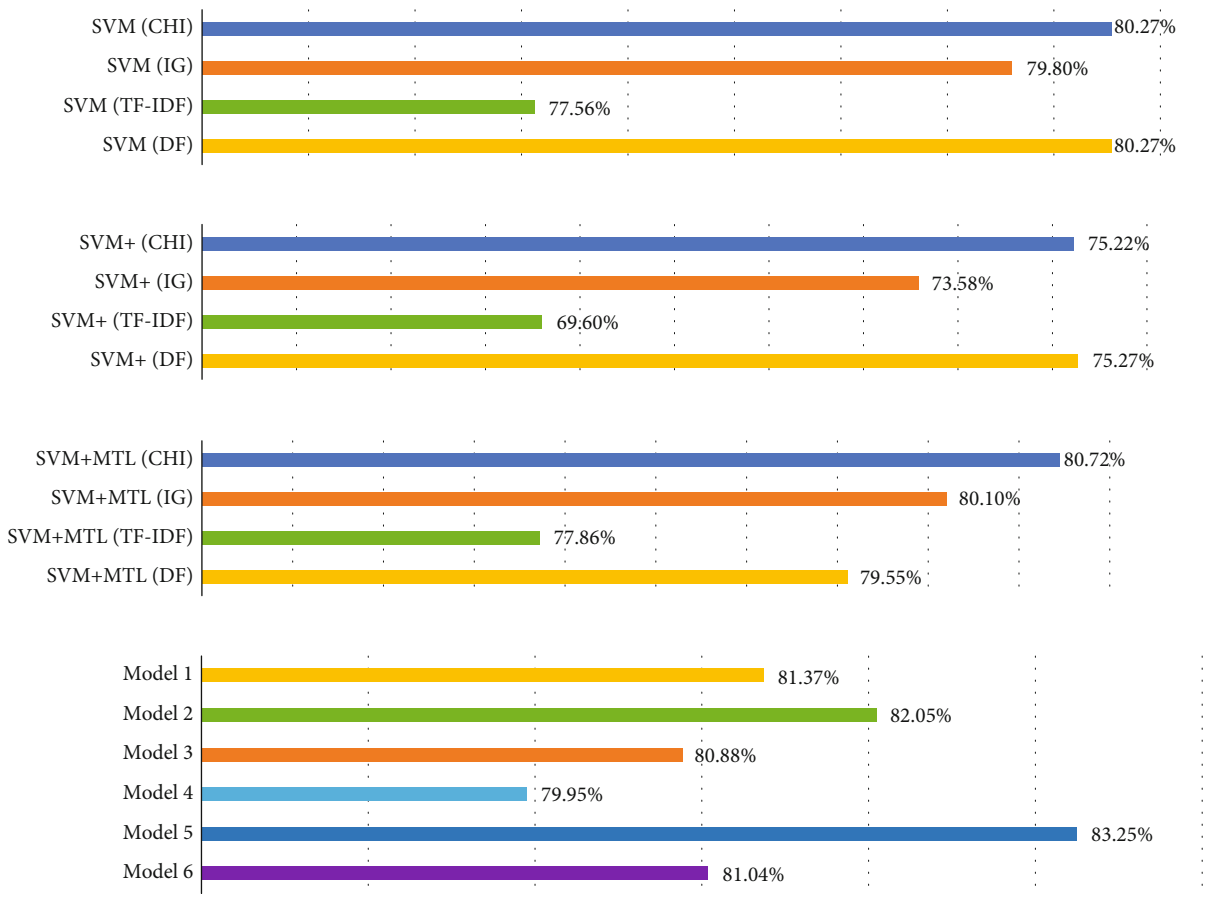


Figure 12: Comparison of overall model performance in game reviews.

this study, both exhibited a prediction accuracy of 85.68%, thereby exhibiting better prediction performances than single models. Figure 11 presents the results obtained from comparing the performance of the entire models based on movie reviews.

The results obtained from analyzing each sentiment classification model based on game reviews in the experiments are presented as follows: the sentiment classification models that apply the SVM technique based on the chi-square statistic and document frequency both exhibited a prediction accuracy of 80.27%, while the sentiment classification models that apply the SVM+ and SVM+MTL techniques based on the document frequency and chi-square statistic exhibited prediction accuracies of 75.27% and 80.72%, respectively. Based on these analytical results, it was verified that the sentiment classification model that apply the SVM+MTL technique based on the chi-square statistic exhibited the best prediction accuracy among the other single models. The sentiment classification model 5, which is an integrated model proposed in this study, exhibited a prediction accuracy of 82.33%, thereby indicating that it exhibited a better prediction performance than single models. Figure 12 presents the results obtained from comparing the performances of the entire models based on game reviews.

The results obtained from analyzing each sentiment classification model based on music reviews in the experiments are presented as follows. The sentiment classification models that apply the SVM, SVM+, and SVM+MTL techniques based on the chi-square statistic, TF-IDF, and chi-square statistic exhibited prediction accuracies of 80.45%, 75.23%, and 79.85%, respectively. Based on these analytical results, it was verified that the sentiment classification model that applies the SVM+MTL technique based on the chi-square statistic exhibited the best prediction accuracy among the other single models. The sentiment classification model 4, which is an integrated model proposed in this study, exhibited a prediction accuracy of 80.96%, thereby exhibiting a better prediction performance than single models. Figure 13 presents the results obtained from comparing the performances of the entire models based on music reviews.

The results obtained from analyzing each sentiment classification model based on book reviews in the experiments are presented as follows: the sentiment classification model that applies the SVM, SVM+, and SVM+MTL techniques based on the chi-square statistic, document frequency, and chi-square statistic exhibited prediction accuracies of 80.93%, 72.29%, and 81.00%, respectively. Based on these analytical results, it was verified that the sentiment classification model that applies the SVM+MTL technique based on the chi-square statistic exhibited the best prediction accuracy among the other single models. The sentiment classification model 6, which is an integrated model proposed in this

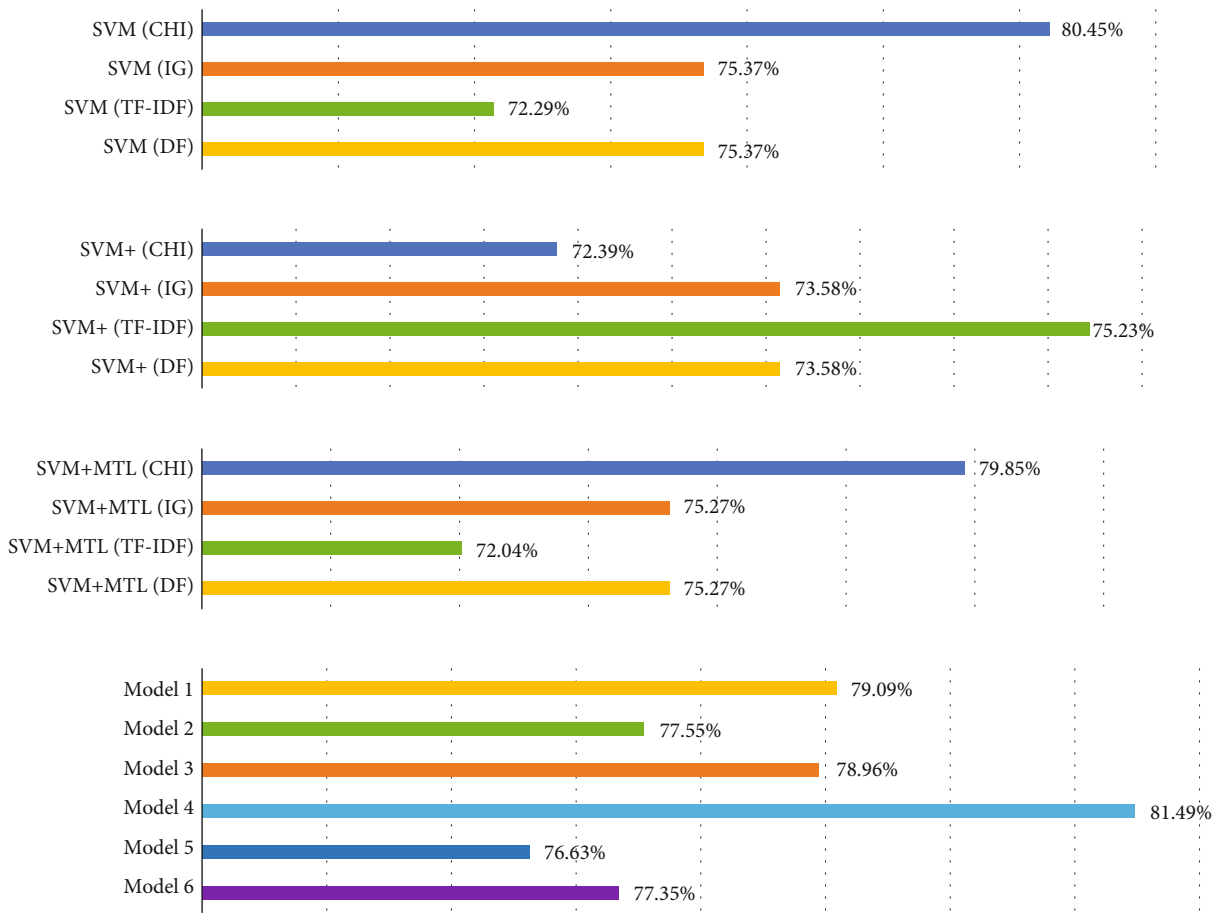


Figure 13: Comparison of overall model performance in music reviews.

study, exhibited a prediction accuracy of 81.11%, thereby exhibiting a more improved prediction performance than single models. Figure 14 presents the results obtained from comparing the performance of the entire models based on book reviews.

Tables 4 and 5 present the results obtained from analyzing sentiment classification models based on four cases (i.e., movies, games, music, and books). The analytical results indicate that the sentiment classification model that applies the chi-square statistic exhibited the most effective performance in terms of extracting core terms. This result was obtained because the chi-square statistic exhibited a close relationship with documents and terms. In other words, because this technique is closely related to terms with significant information power, the significance of these terms was reflected in documents.

This study predicted the performance of sentiment classification models that apply SVM, SVM+, and SVM+MTL techniques according to the amount of data and application of techniques. It was inferred that the sentiment classification model proposed in this study exhibited the best performance among the others. In addition, because the sentiment classification model that applies the SVM+MTL technique adopted group data for training, verification, and evaluation, it exhibited better performance than the sentiment classification models that apply the SVM and SVM+ techniques, respectively.

It was expected that the sentiment classification model that applies the SVM+ technique might exhibit better performance than the sentiment classification model that applies the SVM technique. However, the following problems were observed in this experiment. The first problem is related to the selection of group data. In this experiment, the year of product issue was selected as group data in the same experimental environment. Because the groups did not have a close relationship with each other in the sentiment classification model that applies the SVM+ technique, this model exhibited the lowest prediction performance among other models. The second problem is related to the amount of data.

This study conducted an experiment by extracting sample data and employing a small amount of data. The results obtained from the experiments indicate that the model that applies the SVM+MTL technique exhibited the best performance, while the model that applies the SVM+ technique exhibited a better prediction performance than the model that applies the SVM technique. However, as the amount of data increased, the model that applies the SVM+

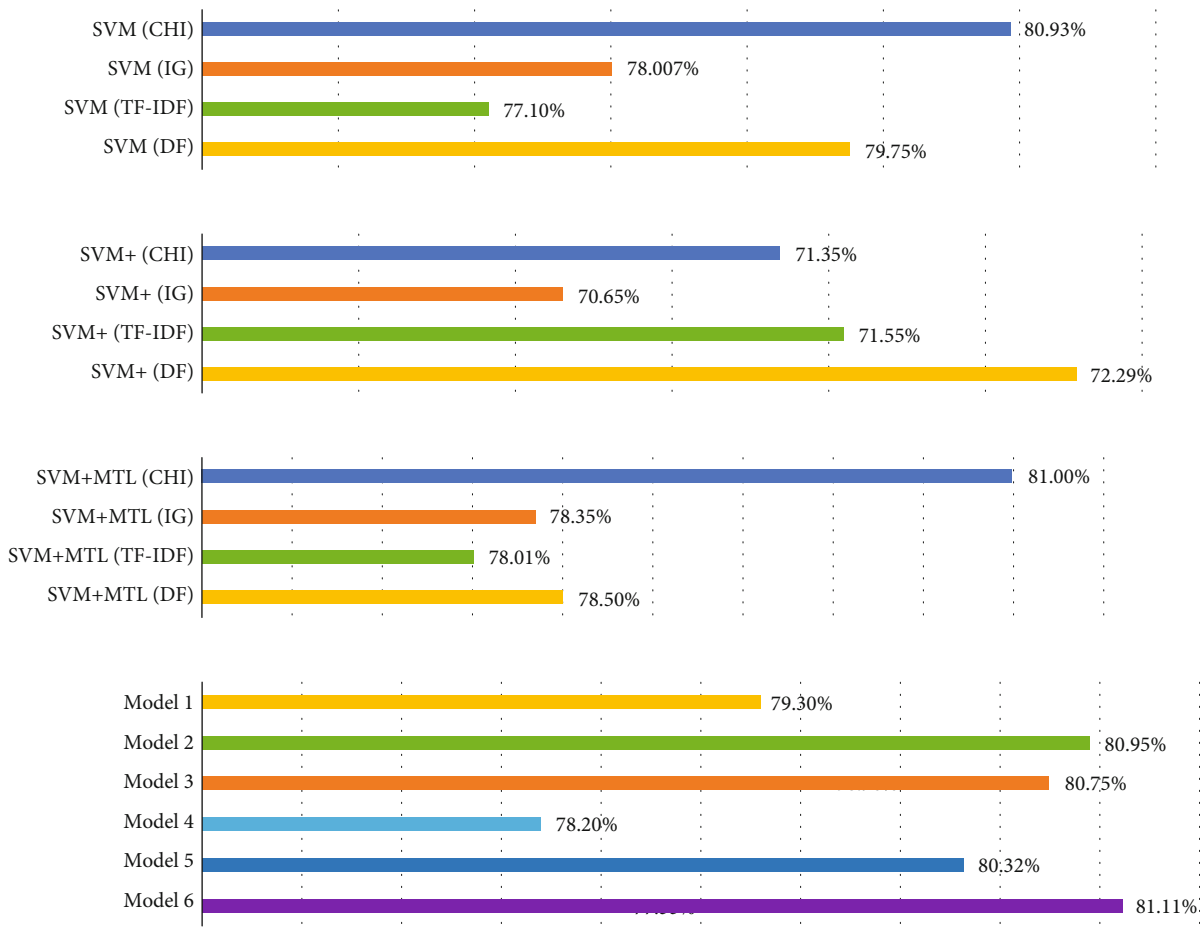


Figure 14: Comparison of overall model performance in book reviews.

Table 4: Experiment results of sentiment classification models in four cases.

	Document frequency			TF-IDF			Information gain			Chi-square statistic		
	SVM	SVM+	SVM+MTL	SVM	SVM+	SVM+MTL	SVM	SVM+	SVM+MTL	SVM	SVM+	SVM+MTL
Movie												
Accuracy	76.95	71.15	77.15	75.41	67.33	75.54	77.65	67.38	77.91	*84.11*	*75.92*	*82.32*
F-measure	77.18	75.52	72.37	72.99	66.43	73.49	76.80	66.78	77.18	75.52	75.24	81.78
Game												
Accuracy	*80.27*	*75.27*	79.55	77.56	69.60	77.86	79.80	73.58	80.10	*80.27*	75.22	*80.72*
F-measure	81.06	75.03	78.45	78.62	68.61	78.79	80.97	73.29	81.37	82.06	74.93	81.38
Music												
Accuracy	75.37	73.58	75.27	72.29	*75.23*	72.04	75.37	73.58	75.27	*80.45*	72.39	*79.85*
F-measure	75.06	75.37	75.03	72.14	74.93	71.37	55.06	73.73	75.03	81.88	73.30	79.63
Book												
Accuracy	79.75	*72.29*	78.50	77.10	71.55	78.01	78.00	70.65	78.35	*80.93*	71.35	81.00
F-measure	77.48	70.99	78.28	75.40	67.30	75.50	77.44	70.15	79.93	78.42	70.90	79.55

technique exhibited the poorest performance, contrary to expectations. The final problem is related to the selection of the kernel function and parameter adjustment. Accordingly, the results obtained from the experiments conducted in this study indicated that an appropriate kernel function and parameter should be selected.

In this study, the performance of the proposed sentiment classification model was evaluated by comparing it with the prediction performance of single models. The results obtained from analyzing the performance of sentiment classification models based on movie, game, music, and book reviews indicate that the sentiment classification model that

Table 5: Experiment results of document-level sentiment classification models in four cases.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Movie						
Accuracy	84.83	*85.68*	84.23	83.78	*85.68*	82.81
F-measure	84.09	85.61	82.09	82.05	85.59	82.10
Game						
Accuracy	81.37	82.05	80.88	79.95	*83.25*	81.04
F-measure	80.23	81.29	79.48	79.27	82.33	80.38
Music						
Accuracy	79.09	77.55	78.96	81.49	76.63	77.35
F-measure	78.46	76.23	75.79	*80.96*	75.73	76.65
Book						
Accuracy	79.30	80.95	80.75	78.20	80.32	*81.11*
F-measure	77.42	78.90	78.55	77.44	70.15	78.93

applies the chi-square statistic exhibited the best performance among the term information extraction techniques compared.

5. Conclusion

This study established a document-level term-based sentiment classification model by collecting user reviews generated in an e-commerce environment and using reviews that contained subjective opinions of users. In addition, this study analyzed the proposed model according to four cases to increase efficiency and information power, perform effective term extraction, and present the comparative performance of sentiment classification models. Accordingly, an experiment was conducted to verify the prediction performances of single models and the proposed model, respectively.

Specifically, this study developed sentiment classification models that apply SVM+ and SVM+MTL techniques and compared the prediction performance of these models with the sentiment classification model that applies the SVM technique, which is frequently used in existing studies. The aforementioned techniques are frequently applied in research to solve binary classification problems. This study is distinguished from existing studies in that it applied these techniques to opinion mining research. Furthermore, it proposed an integrated sentiment classification model that addressed the limitations of single models and exhibited enhanced prediction performance, which was verified in experimental results.

Furthermore, this study applied the proposed method to an opinion mining research, which requires the analysis of a considerable amount of data, and verified that the proposed model effectively analyzed reviews. The proposed model identified sentimental expressions of existing users toward products in an e-commerce environment based on various POS. Accordingly, it is deduced that research on various POS can be adopted as an effective methodology for analyzing decisions to be made by potential users in the process of purchasing products. The proposed model has advantages because it can collect, use, and monitor information on certain products, surveys, or industries from users on social media in actual time, to reflect analytical results to customers. It is expected that the proposed model can contribute to increasing the spread of opinion mining for sentiment classification based on user reviews. In addition, this model can also improve e-commerce services and the quality of these services by enabling users to obtain desired information more easily and purchase or use desired products based on various types of recommendation information provided to them.

Further research should be conducted to combine studies at the document level with those at the sentence level. As this experiment did not reflect the latest user reviews updated, the flow of positive or negative tendencies of opinions over time was not accurately predicted. This study also has a limitation in that highly influential terms adopted as input variables cannot represent documents. Considering this limitation, follow-up research should be conducted to form a sentiment dictionary at the sentence level, apply various input variables to the document-level sentiment classification model, and ultimately increase prediction performance.

Data Availability

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

Conflicts of Interest

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

Acknowledgments

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) Funded by the Ministry of Education (Nos. NRF-2021R1I1A3049061 and NRF-2017R1A6A1A03015496).

References

[1] M. Kang, J. Ahn, and K. Lee, "Opinion mining using ensemble text hidden Markov models for text classification," *Expert Systems with Applications*, vol. 94, pp. 218–227, 2018.

[2] S. Sun, C. Luo, and J. Chen, "A review of natural language processing techniques for opinion mining systems," *Information fusion*, vol. 36, pp. 10–25, 2017.

[3] S. B. Marzijarani and H. Sajedi, "Opinion mining with reviews summarization based on clustering," *International Journal of Information Technology*, vol. 12, no. 4, pp. 1299–1310, 2020.

[4] M. Husnain, M. M. S. Missen, N. Akhtar, M. Coustaty, S. Mumtaz, and V. S. Prasath, "A systematic study on the role of SentiWordNet in opinion mining," *Frontiers of Computer Science*, vol. 15, no. 4, pp. 1–19, 2021.

[5] P. Sánchez-Núñez, M. J. Cobo, C. De Las Heras-Pedrosa, J. I. Peláez, and E. Herrera-Viedma, "Opinion mining, sentiment analysis and emotion understanding in advertising: a bibliometric analysis," *IEEE Access*, vol. 8, pp. 134563–134576, 2020.

[6] F. Alattar and K. Shaalan, "A survey on opinion reason mining and interpreting sentiment variations," *IEEE Access*, vol. 9, pp. 39636–39655, 2021.

[7] A. Kumar, A. S. Veerubhotla, V. T. Narapareddy, V. Aruru, L. B. Neti, and A. Malapati, "Aspect term extraction for opinion mining using a Hierarchical Self-Attention Network," *Neurocomputing*, vol. 465, pp. 195–204, 2021.

[8] J. Yu, J. Jiang, and R. Xia, "Global inference for aspect and opinion terms co-extraction based on multi-task neural networks," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 27, no. 1, pp. 168–177, 2019.

[9] F. Ali, K. S. Kwak, and Y. G. Kim, "Opinion mining based on fuzzy domain ontology and Support Vector Machine: A proposal to automate online review classification," *Applied Soft Computing*, vol. 47, pp. 235–250, 2016.

[10] Y. Cai, K. Yang, D. Huang et al., "A hybrid model for opinion mining based on domain sentiment dictionary," *International Journal of Machine Learning and Cybernetics*, vol. 10, no. 8, pp. 2131–2142, 2019.

[11] C. Hung and S. J. Chen, "Word sense disambiguation based sentiment lexicons for sentiment classification," *Knowledge-Based Systems*, vol. 110, pp. 224–232, 2016.

[12] H. Ha, H. Han, S. Mun, S. Bae, J. Lee, and K. Lee, "Visualization of movie recommendation system using the sentimental vocabulary distribution map," *Journal of The Korea Society of Computer and Information*, vol. 21, no. 5, pp. 19–29, 2016.

[13] M. H. Arif, J. Li, M. Iqbal, and K. Liu, "Sentiment analysis and spam detection in short informal text using learning classifier systems," *Soft Computing*, vol. 22, no. 21, pp. 7281–7291, 2018.

[14] D. H. Pham and A. C. Le, "Learning multiple layers of knowledge representation for aspect based sentiment analysis," *Data & Knowledge Engineering*, vol. 114, pp. 26–39, 2018.

[15] N. F. Da Silva, E. R. Hruschka, and E. R. Hruschka Jr., "Tweet sentiment analysis with classifier ensembles," *Decision Support Systems*, vol. 66, pp. 170–179, 2014.

[16] B. Liu, "Sentiment analysis and opinion mining," *Synthesis lectures on human language technologies*, vol. 5, no. 1, pp. 1–167, 2012.

[17] D. Bi, J. Kong, X. Zhang, and J. Yang, "Analysis on health information acquisition of social network users by opinion mining: case analysis based on the discussion on COVID-19 vaccinations," *Journal of Healthcare Engineering*, vol. 2021, 2021.

[18] M. Keyvanpour, Z. K. Zandian, and M. Heidarypanah, "OMLML: a helpful opinion mining method based on lexicon and machine learning in social networks," *Social Network Analysis and Mining*, vol. 10, no. 1, pp. 1–17, 2020.

[19] W. Medhat, A. Hassan, and H. Korashy, "Sentiment analysis algorithms and applications: a survey," *Ain Shams Engineering Journal*, vol. 5, no. 4, pp. 1093–1113, 2014.

[20] A. Tripathy, A. Anand, and S. K. Rath, "Document-level sentiment classification using hybrid machine learning approach," *Knowledge and Information Systems*, vol. 53, no. 3, pp. 805–831, 2017.

[21] E. Cambria, B. Schuller, Y. Xia, and C. Havasi, "New avenues in opinion mining and sentiment analysis," *IEEE Intelligent Systems*, vol. 28, no. 2, pp. 15–21, 2013.

[22] P. D. Turney and M. L. Littman, "Measuring praise and criticism," *ACM Transactions on Information Systems (tois)*, vol. 21, no. 4, pp. 315–346, 2003.

[23] T. Y. Kim, S. B. Pan, and S. H. Kim, "Sentiment digitization modeling for recommendation system," *Sustainability*, vol. 12, no. 12, p. 5191, 2020.

[24] N. F. Ibrahim and X. Wang, "A text analytics approach for online retailing service improvement: evidence from Twitter," *Decision Support Systems*, vol. 121, pp. 37–50, 2019.

[25] F. Hu, L. Li, X. Xu, J. Wang, and J. Zhang, "Opinion extraction by distinguishing term dependencies and digging deep text features," *Neural Computing and Applications*, vol. 31, no. 9, pp. 5419–5429, 2019.

[26] A. Thakkar and K. Chaudhari, "Predicting stock trend using an integrated term frequency-inverse document frequency-based feature weight matrix with neural networks," *Applied Soft Computing*, vol. 96, article 106684, 2020.

[27] N. S. M. Nafis and S. Awang, "An enhanced hybrid feature selection technique using term frequency-inverse document frequency and support vector machine-recursive feature elimination for sentiment classification," *IEEE Access*, vol. 9, pp. 52177–52192, 2021.

[28] E. O. Omuya, G. O. Okeyo, and M. W. Kimwele, "Feature selection for classification using principal component analysis and information gain," *Expert Systems with Applications*, vol. 174, article 114765, 2021.

[29] Y. Liu, X. Yi, R. Chen, Z. Zhai, and J. Gu, "Feature extraction based on information gain and sequential pattern for English question classification," *IET Software*, vol. 12, no. 6, pp. 520–526, 2018.

[30] C. Z. Liu, Y. X. Sheng, Z. Q. Wei, and Y. Q. Yang, "Research of text classification based on improved TF-IDF algorithm," in *2018 IEEE International Conference of Intelligent Robotic and Control Engineering (IRCE)*, pp. 218–222, Lanzhou, China, 2018.

[31] Z. Zhu, J. Liang, D. Li, H. Yu, and G. Liu, "Hot topic detection based on a refined TF-IDF algorithm," *IEEE access*, vol. 7, pp. 26996–27007, 2019.

[32] C. Iwendi, S. Ponnan, R. Munirathinam, K. Srinivasan, and C. Y. Chang, "An efficient and unique TF/IDF algorithmic model-based data analysis for handling applications with big data streaming," *Electronics*, vol. 8, no. 11, p. 1331, 2019.

[33] A. Jalilifard, V. F. Caridá, A. F. Mansano, R. S. Cristo, and F. P. C. da Fonseca, "Semantic sensitive TF-IDF to determine word relevance in documents," in *Advances in Computing and Network Communications*, pp. 327–337, Springer, Singapore, 2021.

[34] M. Labani, P. Moradi, F. Ahmadizar, and M. Jalili, "A novel multivariate filter method for feature selection in text classification problems," *Engineering Applications of Artificial Intelligence*, vol. 70, pp. 25–37, 2018.

[35] P. Sur, Y. Chen, and E. J. Candès, "The likelihood ratio test in high-dimensional logistic regression is asymptotically a rescaled chi-square," *Probability Theory and Related Fields*, vol. 175, no. 1-2, pp. 487–558, 2019.

[36] V. Vapnik, *Estimation of dependencies based on empirical data*, Springer-Verlag, 2nd edition, 2006.

[37] A. Altan and S. Karasu, "The effect of kernel values in support vector machine to forecasting performance of financial time series," *The Journal of Cognitive Systems*, vol. 4, no. 1, pp. 17–21, 2018.

[38] J. Xu, C. Xu, B. Zou, Y. Y. Tang, J. Peng, and X. You, "New incremental learning algorithm with support vector machines," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 49, no. 11, pp. 2230–2241, 2019.

[39] T. Y. Kim, S. H. Bae, and Y. E. An, "Design of smart home implementation within IoT natural language interface," *IEEE Access*, vol. 8, pp. 84929–84949, 2020.

[40] M. Lapin, M. Hein, and B. Schiele, "Learning using privileged information: SVM+ and weighted SVM," *Neural Networks*, vol. 53, pp. 95–108, 2014.

[41] B. Ribeiro, C. Silva, A. Vieira, A. Gaspar-Cunha, and J. C. das Neves, "Financial distress model prediction using SVM+," in *The 2010 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–7, Barcelona, Spain, 2010.

[42] C. Serra-Toro, V. J. Traver, and F. Pla, "Exploring some practical issues of SVM+: is really _privileged_ information that helps?," *Pattern Recognition Letters*, vol. 42, pp. 40–46, 2014.

[43] L. Liang, F. Cai, and V. Cherkassky, "Predictive learning with structured (grouped) data," *Neural Networks*, vol. 22, no. 5-6, pp. 766–773, 2009.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 6154614, 10 pages
https://doi.org/10.1155/2022/6154614

Research Article

Research on Vehicle Detection and Recognition Based on Infrared Image and Feature Extraction

Ke Li,[1,2] Rongchun Deng,[3] Yongkang Cheng,[3] Rongqun Hu,[4] and Keyong Shen[4]

[1]*School of Mechanical and Electrical Engineering, Nanchang Institute of Technology, Nanchang, 330044 Jiangxi, China*
[2]*Nanchang Key Laboratory of Intelligent Manufacturing System Simulation Planning and Construction, Nanchang Institute of Technology, Nanchang, 330044 Jiangxi, China*
[3]*School of Electronics and Information, Nanchang Institute of Technology, Nanchang, 330044 Jiangxi, China*
[4]*College of Computer Information and Engineering, Nanchang Institute of Technology, Nanchang, 330044 Jiangxi, China*

Correspondence should be addressed to Keyong Shen; jsj@nut.edu.cn

Received 29 January 2022; Accepted 1 April 2022; Published 21 April 2022

Academic Editor: Hoon Ko

Vehicle detection and identification and safe distance keeping technology have become the main content of current intelligent transportation system research. Among them, vehicle detection and recognition is one of the most important research contents, and it is also crucial to the safe driving of vehicles. Real-time detection and recognition of current vehicles can effectively prevent the occurrence of malignant traffic accidents such as rear-end collision. Because the infrared image has some shortcomings such as poor contrast, loud noise, and blurred edge, this paper mainly studies the color space preprocessing of the image and uses threshold segmentation method and infrared image enhancement to segment the front vehicle and background. That is to say, by analyzing the infrared image captured by infrared CCD, we use median filter to remove noise from the collected infrared image and then use the improved histogram equalization to enhance the contrast of the image. Vertical Sobel operator is selected to enhance the vertical edge of the image, and the image is segmented by binary segmentation method. Finally, vehicle detection and recognition are realized by vertical edge symmetry, aspect ratio, and gray-scale symmetry. The experimental image and experimental data analysis results show that the image processing technology studied in this paper has achieved the intended research purpose.

1. Introduction

With the development of economy, the number of automobile and traffic accidents is increasing, so it is urgent to take measures to reduce traffic accidents and protect the safety of life and property [1]. Vehicle detection and recognition and safe distance keeping technology have become the main content of current intelligent transportation system research. Real-time detection and identification of the current vehicle can effectively prevent the occurrence of vicious traffic accidents such as rear-end collisions. The predecessor of the intelligent transportation system is the intelligent vehicle road system. The intelligent transportation system effectively and comprehensively applies advanced information technology, data communication technology, sensor technology, electronic control technology, and computer technology to the entire transportation management system, so as to establish a large-scale, all-round function, real-time, accurate, and efficient integrated transportation and management system [2]. The intelligent transportation system (ITS) project was formally proposed in 1990. It mainly focuses on improving vehicle safety, intelligence, and providing friendly human-vehicle interface. Transportation is facing many problems, such as high accident rate, traffic congestion, traffics, and carbon emission air pollution. Researchers combine virtual technology with transportation called intelligent transportation system, and it is effectively and comprehensively applied to transportation, service control, and vehicle manufacturing and strengthens the connection between vehicles, roads, and users, thereby forming an integrated transportation system that guarantees safety, improves efficiency, improves the environment, and saves energy [3]. It

has become a research hotspot to improve the utilization rate of existing roads, the safety degree of road traffic, and the comfort degree of road use by increasing the technology content. The growing demand for mobility has brought about significant changes in transport infrastructure [4, 5]. Intelligent transportation system (ITS) is becoming an important part of the society, and reliable and efficient vehicle communication is the key driving factor for good operation of ITS. Vehicle detection has been an important research area for many years. There are many valuable applications from traffic planners'support to real-time traffic management. Because of the high traffic volume and limited space, especially in dense urban areas, it is meaningful to detect cars. In order to meet the needs of various ITS applications, it is necessary to jointly consider the configuration and optimization of vehicle-to-vehicle and infrastructure communications [6, 7]. The application of wireless access technology in vehicle environment leads to the improvement of road safety and the reduction of the number of deaths caused by road traffic accidents through the development of road safety applications and the promotion of information sharing among mobile vehicles on roads. It is the important application of intelligent transportation system in highway management, so the research on vehicle detection and recognition based on infrared image and feature extraction is particularly important.

In recent years, video surveillance and monitoring system has been widely used in traffic management, mainly for traffic density estimation and vehicle classification. In recent years, many algorithms have been proposed to detect, recognize, and track vehicles in front of them based on the corresponding relationship between regions and vehicles established by moving vehicles through image sequences. In the past decade, vision-based vehicle detection technology for road safety improvement has attracted more and more attention [8]. Guido et al. [9] introduce a method of tracking mobile vehicles, which combines unmanned aerial vehicles with video processing technology. For vehicle distance measurement in front, Wu et al. [10] use the shadow of the vehicle in front to identify the location of the vehicle in front and use functional fuzzy neural network to estimate the actual distance. The experimental results show that the system runs successfully in real-time environment. The actual distance is estimated by using functional fuzzy neural network. It is a challenging task. Zhang [11] uses HMAX to recognize vehicle types from images. The database includes more than 2,000 vehicle images from 26 categories involving various complex photographic conditions recorded by surveillance cameras.

As an important way of information expression and storage, image has wide application value. Certain parking spaces are used in vehicle management through monitoring and image recognition. Cameras are used for the identification of vehicles by their plates based on smart and visual character recognition technology. The symmetric disk drive proposed by Hsieh et al. [12] is used to solve the problems of multiplicity and fuzziness and determines the areas of concern of each vehicle on the road without mobile function. Xiaoling [13] presents a method of vehicle location for accident-causing operation based on computer dynamic image processing technology, which can accurately locate the moving vehicle. Experiments show that the computer dynamic visualization method can greatly improve the accuracy of key frame location results of separated vehicle images. Neto et al. [14] introduced a new system that can sense and recognize Brazilian license plates. The digital image processing technologies such as Hough transform, morphology, threshold and Canny edge detector are used to extract characters, and the multilayer concept theory of least square, least average square, limit learning machine, and neural network is used to recognize numbers.

At present, the detection and recognition of vehicles in front of us are mainly based on monocular vision, infrared sensors, lidar, and so on. Wang et al. [15] studied the digital signal processing technology of visual perception, ultrasonic sensor, and radar technology. Ahmed et al. [16] use real-time velocity data collected from AVI to check the identification of highway locations with high collision potential. Sivaraman and Trivedi [17] introduced the progress of vehicle detection in detail and discussed the application of monocular. The research involved the use of space-time measurement, trajectory, and various features to characterize road behavior. Cheng et al. [18] designed an automatic vehicle detection system for aerial surveillance based on pixel classification, which retains the relationship between adjacent pixels in the region during feature extraction. Vehicle color and nonvehicle color are effectively separated by color transformation. The threshold of canny edge detector is adjusted automatically for edge detection. The results show that the method has flexibility and good generalization ability. Tian et al. [19] propose a vehicle recognition method using multiple sensor nodes.

Teoh and Bräunl [20] describe the use of a single front-view camera and vehicle edge and symmetry features to detect vehicle and road information. This paper introduces a method of extracting symmetrical regions in images by using multisize windows and clustering techniques. According to the detected symmetrical regions, the position of vehicles in images is assumed, and then, these regions are further processed to enhance their symmetrical edges. Vehicle boundaries are detected from enhanced projections of vertical and horizontal edges. Liu and Li [21] study the parameters of the optimal solution obtained by gray-scale image acquisition by CCD to identify lane lines, use parabolic model as objective function to fit lane lines, use genetic algorithm to optimize parabolic parameters, and use binary coding, multipoint crossover, and mutation genetic process. Tang et al. [22] use laser, ultrasonic or radar sensor space sensor systems, and communication technology to exchange information between HISS servers and vehicles, thus providing road information around vehicles. Biqing et al. [23] describe the function of vehicle recognition and location mainly by two core components, binocular vision system and camera calibration model. The results show that it can accurately locate the target position in recognition function. Xi et al. [24] study the depth estimation algorithm for monocular infrared images based on a nonlinear learning model. Experimental results show that most of the depth estimated

by this model is consistent with the original information on the depth of infrared images.

CCD camera is in the security system, the generation of image is mainly from CCD camera, CCD is the abbreviation of charge coupled device, and it can turn light into electric charge and store and transfer the electric charge and can also take out the stored electric charge to make the voltage. Therefore, it is an ideal CCD camera element. The CCD camera formed by it has the characteristics of small size and light weight, not affected by magnetic fields, has the characteristics of antivibration and impact, and is widely used. In this paper, infrared CCD is used to collect images. Under this background, this paper is aimed at the malignant traffic accidents caused by vehicle rear-end collision on expressway at night. By using the theory and technology of infrared CCD and image processing, the vehicle identification method in front of the vehicle at night is systematically studied in order to provide technical support for reducing the occurrence of similar traffic accidents.

2. Proposed Method

The process of vehicle recognition mostly follows the flow chart of Figure 1. As can be seen from Figure 1, it can be seen that each module affects the final results. This paper mainly studies vehicle detection and recognition based on infrared image analysis and feature extraction process from vehicle edge features. Figure 1 is an analysis of the extraction of infrared images, from the extraction of raw color features to the construction of feature sets to the tenfold classification using SVM, and the entire process is strictly carried out.

2.1. Infrared Image Enhancement. In the process of vehicle recognition, infrared image preprocessing is an indispensable link. Infrared image preprocessing can remove or reduce noise and clutter in infrared image and improve image quality and signal-to-noise ratio. The general infrared picture preprocessing methods can be summarized as follows: some operations on the original infrared picture, such as a transformation or calculation from the space or frequency range, can improve the target in the infrared picture, suppress noise in the infrared image, and provide a good input for subsequent processing, improving the overall performance of the system [25]. This entry uses the spatial domain method for image enhancement.

A known image $S(n, m)$ is decomposed into reflecting object image $R(n, m)$ and incident light image $L(n, m)$, and any point (n, m) in image S can be shown $S(n, m) = R(n, m) * L(n, m)$. Image enhancement based on Retinex theory can enhance the visual effect of some low-quality infrared images, improve their brightness uniformity, further enhance the brightness of local dark areas, and highlight the outline information of the objective. Not only the gray level of the whole area of the equipment has been improved, but also the detailed information of the enhanced equipment has been greatly strengthened and highlighted, especially the brightness of some darker areas caused by illumination and the connecting areas between the parts has been further enhanced, which further restores the formal structure information of the equipment. Among them, $(m, n)(a, b)(x, y)$ are image pixel coordinates. The basic steps of image enhancement are as follows:

(a) The arithmetic functions are performed in the original image to separate the elements of the incoming and outgoing light and the reflected light:$S(a, b) = (Y(a, b)) + \log(L(a, b))$

(b) Convolution of the original image using a Gauss template is equivalent to low-pass filtering of the original image, and the image $D(a, b), F(x, y)$ denotes the Gauss filtering function: $D(a, b) = S(a, b) * F(a, b)$

(c) In the log domain, the image obtained in the previous step is subtracted from the original image. $G(x, y) = R(x, y) - \log(D(x, y))$

(d) The final image $R(x, y)$: $R(x, y) = \exp(G(x, y))$ is obtained by taking the antilogarithmic operation of $G(x, y)$.

2.2. Smooth Enhancement of Infrared Image. The improved adaptive bilateral filtering divides the image into the base layer and the detail layer. In the base layer, the histogram specification based on Gaussian mixture model is used to achieve brightness preservation, and in the detail layer, the enhancement function is adaptively selected by using human visual characteristics to enhance weaker details. And protect the clear edges in the original image from distortion, and then, restore to the original gray-scale space. For infrared images, noise can be generated either by random interference of external environment or by random variation of internal physical quantities. The purpose of image smoothing and enhancement is to eliminate noise, improve image quality, and extract object features. In this paper, two kinds of smoothing algorithms are studied.

2.2.1. Mean Filtering. Average filtering is also called linear filtering. Its basic principle is to replace each pixel value in the original image with the average value. Then, assign the mean value to the current pixel point (n, m) as the gray level $g(n, m)$ of the processed image at that point, i.e., $g(n, m) = ''8721f(n, m)/\text{mm}$ is the total number of pixels in the template, including the current pixel, namely, $h(n, m) = 1/M(\sum_{n,m\in s} f(n, m))$.

Filter could be regarded as a low-pass spatial filter, which can effectively remove noise. It can be seen from the above formula that the cost of noise reduction by mean filtering is to make the image unclear, and the larger the neighborhood, the better the ability of noise elimination, but the more serious the image is, the impulse noise cannot be suppressed. The common mean filter algorithm is convolution operation with image pixels. The mean filter can remove noise to some extent. Therefore, when noise is removed, image details become blurred, especially the edges of vehicle images become blurred, which does not favour the detection of extremities and the identification of features of the later images of the vehicle, which is shown as Figure 2.

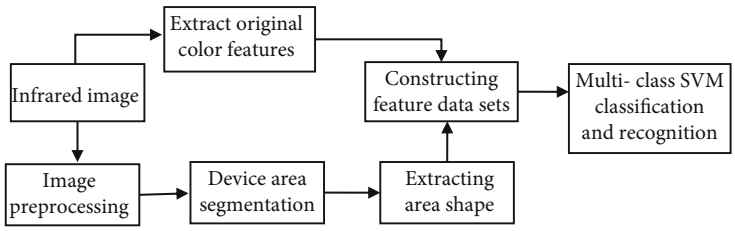


Figure 1: Identification plan process.

Figure 2: Vehicle picture mean filtering.

2.2.2. Median Filtering. The average filtering is a nonlinear image isolation method based on statistical theory. According to the different dimensions of sliding window, median filtering can be divided into one-dimensional and two-dimensional. In order to suppress the noise better, this paper uses median filtering method to filter the noise in the image. Because of the low contrast of the infrared image, the contour of the vehicle in the infrared image is blurred, and the median filtering is better than the mean filtering in noise processing. At the same time, the problem of blurring of vehicle contour image caused by mean filter is overcome to a certain extent, which enhances the details of vehicle image, provides high-quality image for vehicle image edge detection and feature recognition, and enhances vehicle contour features, and it is shown as Figure 3.

2.3. Improved Histogram Equalization Image Enhancement. Histogram equalization can greatly improve the dynamic range of image gray distribution, which is a typical spatial enhancement method. By increasing the overall contrast of the image, better visual effect can be obtained for infrared image. Most of the infrared images have the defects of low contrast, low resolution, and low signal-to-noise ratio. Image enhancement is to enhance its use value, not only to improve the overall quality, mainly reflected in the prominent degree of the goal. At present, a highly practical contrast enhancement technology is based on gray histogram correction technology, namely, histogram equalization.

Contrast is an important index of image quality. Contrast refers to the gray difference between the signal and the surrounding background. The calculation formula is as follows: $C = \sum_{\delta} \delta(i,j)^2 P_{\delta}(i,j)$, and among them, $\delta(i,j) = |i - j|$. $P_{\delta}(i,j)$ is the probability of the distribution of pixels whose gray difference value is delta between adjacent pixels.

Histogram equalization is to make the gray value range of the output image as wide as possible and as uniform as possible through a mapping relationship between gray value intervals. In this way, the processed image can be maximized to meet the human visual requirements. However, this method does not have the ability to identify the target and background. Blindly homogenizing the gray level of all regions will increase the contrast of the background region and reduce the contrast of the target region. The gray range of the processed image will be narrowed, and some local regions will be weakened. There are some peaks in the gray histogram of some images. After processing, the contrast of the image region has been excessively improved.

In order to eliminate the blindness of traditional histogram equalization in data processing, reduce the background gray interval and increase the target gray interval at the same time, so that histogram equalization can be applied to the target gray interval to the greatest extent. Gray scale stretching is the most direct means to change the range of gray scale in a region. Selective gray scale stretching can be achieved by using piecewise linear transformation function.

2.4. Edge Detection of Target Image. Vehicle recognition based on image features needs to recognize the image first. The basis of image recognition is picture separation. The essence of image segmentation is to divide a digital image into different parts and then express these different parts, that is, to extract the characteristics of specific parts [26]. This will help to reduce the amount of data operated in subsequent feature recognition, save storage space, and largely save the operation steps in the subsequent process. It not only weakens the impact of noise but also retains the basic shape features of the target feature structure. Edge detection plays an important role in subsequent processing. Given this nature, the picture edges are usually obtained quickly by

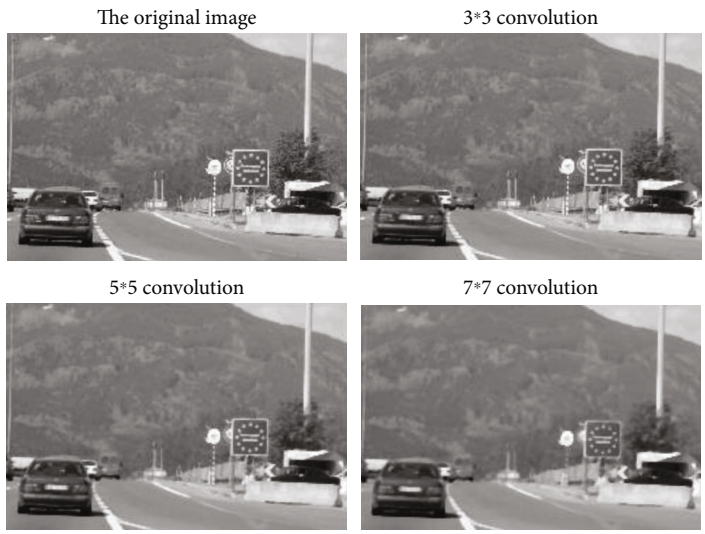


Figure 3: Median filter effect diagram.

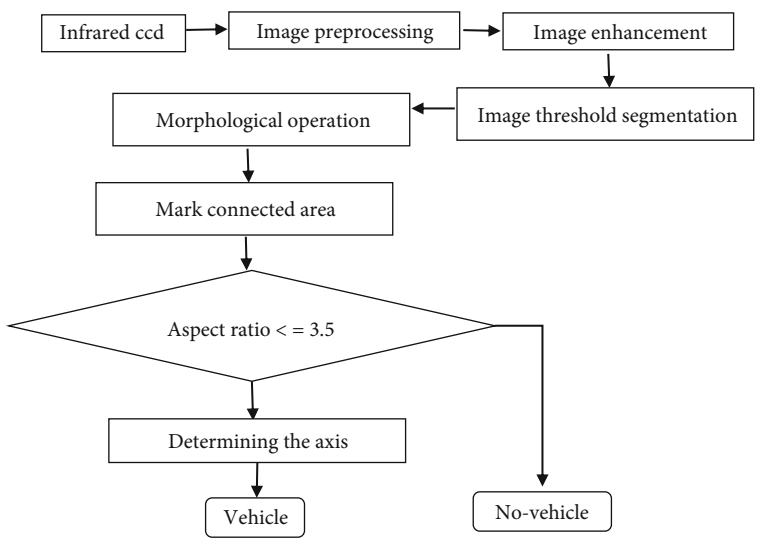


Figure 4: Vehicle identification method flow chart.

Table 1: Convolution kernel of Sobel operator.

	-1	-2	-1		1	0	-1
Horizontal convolution kernel	0	0	0	Horizontal convolution kernel	2	0	-2
	1	2	1		1	0	-1

derivation, and the first-order derivation and the second-order derivation are often used.

(1) *Sobel Operator*. That is the process of solving the first derivative, that is, the approximation of the two-dimensional gradient direction of the image. It is known that $g(m, n)$ is a continuous function. The inclination at a point (m, n) may be expressed by a vector. Its magnitude and angle are as follows:$\nabla f(x, y) = [G_x G_y]^T = [\partial f/\partial x \partial f/\partial y]^T$ The magnitude and orientation angles of the vectors are as follows:

$\nabla f = \mathrm{mag}(\nabla f) = [G_x^2 + G_y^2]^{1/2}$.The above formula gives the extreme value of the rate of change of the value $f(x, y)$ for each unit of distance added in the direction of ∇f. $\varnothing(x, y) = \arctan(G_y/G_x)$. When measuring the angle, the x-axis is chosen as the datum.

The above operators contain a matrix of 3×3 in horizontal and vertical directions. If convolution operation is

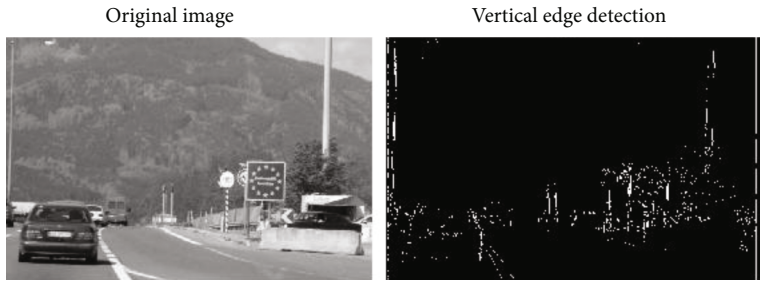


Figure 5: Infrared image vertical edge enhancement effect diagram.

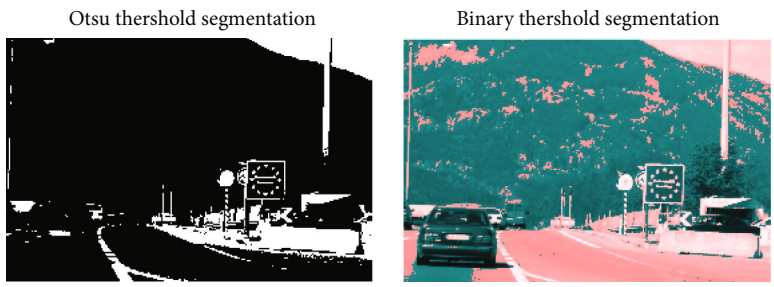


Figure 6: Comparison of the segmentation effects of two threshold segmentation methods.

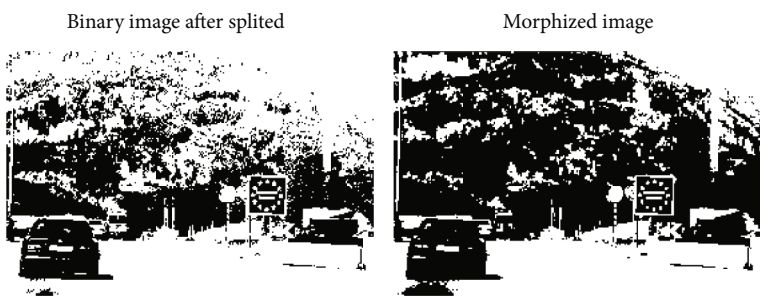


Figure 7: Morphized image.

Table 2: Rectangularity and ratio of width to height.

	1	2	3
Area	291	72	72
Perimeter	78	120	100
Width to height ratio	2.6	1.3	1.1
Rectangularity	0.768	0.627	0.76

made between them and the image, the brightness difference in horizontal and vertical directions can be calculated.

$$G_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} * A. \quad (1)$$

(2) Canny operator

The Canny edge detection operator is a multilevel edge detection algorithm developed by John F. Canny in 1986. More importantly, Canny created "Theory of Computing for Edge Detection" to explain how this technology works. Canny operator uses different thresholds and the first-order differential of Gauss function to extract the strong and weak edges of the image. Therefore, the algorithm can extract weak edges, has good edge feature recognition and location ability, but is very suspectible to noise. The algorithm of Canny operator will be listed below:

(1) To filter the image is to convolute the objective picture and the Gauss template

(2) Find out the extreme value and edge direction of the amplitude, that is, convolute the target image with the template

(3) Compare the magnitude of each pixel with its gradient direction, and judge whether the tilt of each pixel is the local maximum, and the nonmaximum suppression of the magnitude is also called "nonmaximum suppression"

(4) Take the upper and lower thresholds, compare the pixels in the image with the upper and lower

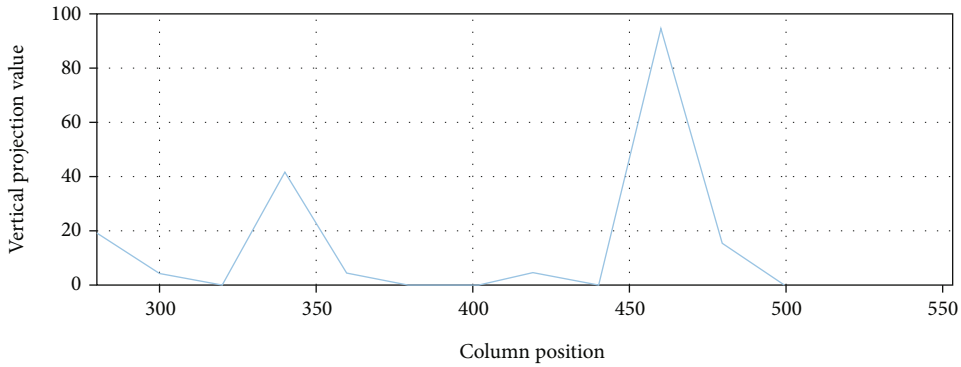


Figure 8: Vertical projection of image.

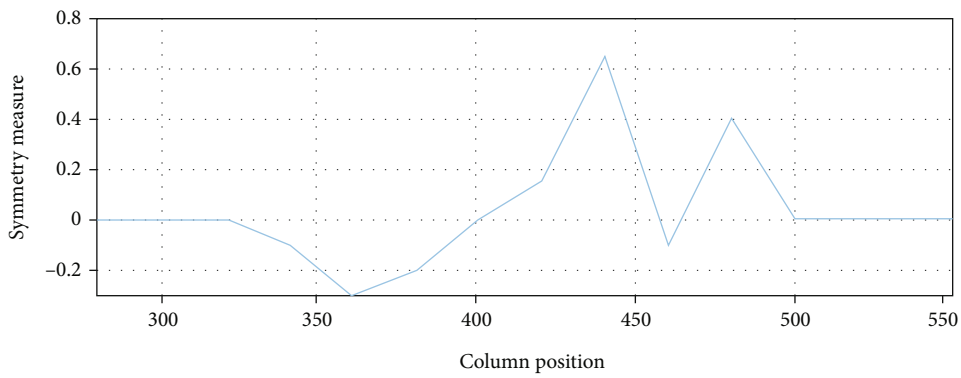


Figure 9: Vertical edge symmetry measure.

thresholds, and classify the upper bounds larger than the threshold as strong bounds, and the lower bounds smaller than the threshold as weak boundaries. Weak boundaries need to be processed later

(5) Take a larger threshold to identify the actual contour features, and then track all edges as a starting point. In the process of tracking, in order to track the fuzzy part of the edge of the target in the image, a small threshold is selected

2.5. Improved Mathematical Morphology. Traditional mathematical morphology can essentially be understood as a set of points in an image, and then, the appropriate structural element operation.

(1) *Extensions*. Extension operators is $\oplus$, f expands with b, and $f \oplus b$ is written, which is defined as follows: $(f \oplus b)(s,t) = \max\{f(s-x, t-y) + b(x,y) | (s-x), (t-y) \in D_f$ and $(x,y) \in D_b$, D_f and D_b both represent the domain of definition of f and b, and $f(x,y)$ in defining foreign territory is assumed to be $-\infty$.

(2) *Reading*. The corrosion operator is f corrodes with b, writes f, which is defined as follows: $(f\Theta b)(s,t) = \min\{f(s+x, t+y) - b(x,y) | (s+x), (t+y) \in D_f$ and $(x,y) \in D_b$.

D_f and D_b both represent the domain of definition of f and b, and $f(x,y)$ in defining foreign territory is assumed to be $+\infty$.

When using morphological algorithm to extract image edges, a single structural element cannot obtain better image edge detection results. According to the characteristics of the target image, the structural elements with appropriate shape and size are combined when selecting the structural elements. The simulation algorithm after twice processing the vehicle image is as follows, which is used to extract edge features. The stages of the improved algorithm are as follows.

(1) For rectangular elements, open and close operations are applied to rectangular images to remove noise

(2) For disc elements, open and close operations are performed to reduce the darker details

(3) Two images have been obtained by using expansion and corrosion operations, respectively

(4) Find out the difference between the two images in step 3, and get a better image edge

The improved mathematical morphology algorithm is used to analyze and further process the measured image, and the contour features of the vehicle image can be extracted more completely. Because the improved algorithm

has obvious advantages in contour extraction, it can eliminate the influence of noise and achieve better road width extraction effect.

3. Experiments

The near infrared imaging system used in this paper is an active infrared imaging system, which includes infrared CCD camera, video capture card, and scanning laser. Through data acquisition card and industrial computer data communication, vehicle detection is realized. For the use of infrared CCD cameras, video capture cards, scanning lasers, etc. as shown in Figure 4. Firstly, according to the characteristics of the original infrared image, image processing and threshold segmentation are used to process the image, and vehicle license plate recognition is based on shape features. In order to eliminate the influence of the factors on both sides of the road and shorten the time of vehicle detection, it is necessary to determine the region of interest for vehicle detection in front of the lane in order to further extract and detect the vertical edges of vehicles. The detection process is shown in Figure 4. Use openCV to extract image features, train svm classifier, and classify vehicles and nonvehicles. Use the trained model to identify the vehicle in the video recorded by the car's front camera.

4. Infrared Image Analysis

This paper mainly focuses on the analysis of the infrared image of the vehicle and carries out vehicle recognition for the vehicle image with obvious license plate characteristics. Firstly, the infrared image is segmented; secondly, the binary image is segmented by an improved morphological operation, and the connected area is marked; the region of interest of the front vehicle recognition is determined. In the region of interest, the area, rectangularity, aspect ratio, roundness, and symmetry of each connected area are analyzed to recognize and locate the vehicle in front of it.

(1) Preprocessing of the collected infrared image

To pick up the boundary between the target and the background in the image, it is necessary to remove the noise of the infrared image and increase the contrast of the image. In this paper, Sobel operator is selected as the edge enhancement operator for vehicle edge detection. According to the direction of derivation, the template is divided into horizontal and vertical directions, and the horizontal and vertical edges of the image are enhanced by finding the maximum gradient. The horizontal and vertical convolution kernels of Sobel operator are shown in Table 1. In this paper, only the vertical convolution kernels are calculated to highlight the vertical edges of vehicles when extracting the vertical edges of vehicles. Figure 5 is the result of vertical edge enhancement.

(2) Threshold segmentation of image edges

After the image is enhanced by Sobel operator, we can see from the result of edge enhancement that a lot of useless information is still included. In this paper, Otsu segmentation algorithm and binary threshold segmentation algorithm are compared. For the image enhanced by vertical Sobel operator edge, Otsu method and binary threshold segmentation algorithm are used to segment the result as shown in Figure 6.

The processing effect shows that the threshold method can retain the edge information, but the edge information of the image segmented by Otsu threshold method is not as rich as that of the binary threshold method.

(3) *Binary Image Processing*. The binary image after threshold segmentation often results in redundant segmentation. The vertical edges of the segmented image are mainly preserved. Figure 7 shows the effect of morphological processing of the original infrared image after image segmentation.

(4) Feature-based vehicle detection

For vehicles, the license plate is usually a regular rectangle with a certain aspect ratio, and the lamp on both sides usually has a certain circularity characteristic. The center of the license plate is the symmetrical axis, showing a certain degree of symmetry. Table 2 calculates the circumference, area, aspect ratio, and rectangularity of the target area of the extracted image.

The license plate is a regular rectangle with a width-to-height ratio of 2 or 3.2. From the data in Table 2, it can be seen that the rectangularity of 1 and 3 is greater than 0.75. Considering the influence of reflection and illumination, the range of width-height ratio in this paper is [1.5, 3.9], and the threshold of rectangularity is 0.7.

(5) Symmetry feature extraction based on vertical edge

In this paper, we first select the edge points of the region of interest in the vertical height direction and then add the edge points together to get the vertical projection value of the edge. Figure 8 is to calculate and analyze the symmetry on the basis of the vertical edge of the vehicle. The value range of symmetry measure is [-1,1]. When the value of measure is 1, it means complete symmetry, and when the value of measure is -1, it means complete asymmetry.

The threshold of symmetry measure chosen in this paper is 0.5. From Figure 9, it can be seen that the maximum symmetry measure in this measurement image is 0.65, and the corresponding x-axis is 434 columns. The threshold of symmetry measure is 0.5. Because the symmetry measure is more than 0, the vehicle area meets the symmetry requirement.

5. Conclusions

In order to extract the boundary between the target and the background in the image, it is necessary to remove the noise in the infrared image and improve the contrast of the image. For the collected infrared image, the median filter is used to

remove noise, and the gray linear enhancement is used to enhance the contrast of the image, and then, the image is segmented. The method of vehicle license plate recognition is also discussed. The research shows that the license plate lamp features of the vehicle image collected by infrared CCD are obvious in most cases. On the basis of denoising and segmentation of vehicle image, this paper identifies and determines the region of interest in vehicle recognition. Vehicle recognition and location are realized according to the rectangularity and aspect ratio of license plate and its central symmetry and based on rectangularity characteristics. For the recognition of vehicle image with obvious edge features, not only in the tail of the car body, but also in the infrared image and vehicle recognition and location, is carried out symmetrically from the vertical edge.

This paper is based on hypothetical conditions, so this study is only applicable to the identification of single vehicle on the road. There is no in-depth study on multilane vehicle identification in the urban area and the identification of front vehicles connected by vehicle back shadow and road guardrail reflections, opposite headlights, and so on. Therefore, I hope to solve this problem in the future research.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that there is no conflict of interest with any financial organizations regarding the material reported in this manuscript.

Acknowledgments

This work was supported by the Science and Technology Research Projects of Jiangxi Provincial Department of Education, GJJ219310 and GJJ212101.

References

[1] R. B. Chen, N. Li, H. F. Xiao, and W. Hou, "Research on safety vehicle distance recognition based on stereo vision and DSP// Applied Mechanics and Materials," *Trans Tech Publications*, vol. 333-335, pp. 805–810, 2013.

[2] Z. Lv, S. Zhang, and W. Xiu, "Deep Learning for security in Digital Twins of Cooperative intelligent transportation Systems," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, pp. 1–10, 2021.

[3] K. N. Qureshi and A. H. Abdullah, "A survey on intelligent transportation systems," *Middle-East Journal of Scientific Research*, vol. 15, no. 5, pp. 629–642, 2013.

[4] G. Dimitrakopoulos and P. Demestichas, "Intelligent transportation systems," *IEEE Vehicular Technology Magazine*, vol. 5, no. 1, pp. 77–84, 2010.

[5] Z. Lv, D. Chen, and Q. Wang, "Diversified technologies in internet of vehicles under intelligent edge computing," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 4, pp. 2048–2059, 2021.

[6] S. Tuermer, F. Kurz, P. Reinartz, and U. Stilla, "Airborne vehicle detection in dense urban areas using HoG features and disparity maps," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 6, no. 6, pp. 2327–2337, 2013.

[7] Z. Lv, R. Lou, and A. K. Singh, "AI empowered communication systems for intelligent transportation systems," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 7, pp. 4579–4587, 2021.

[8] X. Cheng, L. Yang, and X. Shen, "D2D for intelligent transportation systems: a feasibility study," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 4, pp. 1784–1793, 2015.

[9] G. Guido, V. Gallelli, D. Rogano, and A. Vitale, "Evaluating the accuracy of vehicle tracking data obtained from unmanned aerial vehicles," *International Journal of Transportation Science and Technology*, vol. 5, no. 3, pp. 136–151, 2016.

[10] C. F. Wu, C. J. Lin, and C. Y. Lee, "Applying a functional neurofuzzy network to real-time lane detection and front-vehicle distance measurement," *IEEE Transactions on Systems Man and Cybernetics-Part C-Applications Reviews*, vol. 42, no. 4, pp. 577–589, 2012.

[11] B. Zhang, "Classification and identification of vehicle type and make by cortex-like image descriptor HMAX," *International Journal of Computational Vision and Robotics*, vol. 4, no. 3, pp. 195–211, 2014.

[12] J. W. Hsieh, L. C. Chen, and D. Y. Chen, "Symmetrical surf and its applications to vehicle detection and vehicle make and model recognition," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 1, pp. 6–20, 2014.

[13] X. Xiaoling, "Based on computer vision escape vehicle identification system design and research," *Bulletin of Science and Technology*, vol. 11, p. 37, 2012.

[14] E. C. Neto, S. L. Gomes, P. P. Rebouças Filho, and V. H. C. de Albuquerque, "Brazilian vehicle identification using a new embedded plate recognition system," *Measurement*, vol. 70, pp. 36–46, 2015.

[15] W. Wang, Y. Song, J. Zhang, and H. Deng, "Automatic parking of vehicles: a review of literatures," *International Journal of Automotive Technology*, vol. 15, no. 6, pp. 967–978, 2014.

[16] M. M. Ahmed and M. A. Abdel-Aty, "The viability of using automatic vehicle identification data for real-time crash prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 2, pp. 459–468, 2012.

[17] S. Sivaraman and M. M. Trivedi, "Looking at vehicles on the road: a survey of vision-based vehicle detection, tracking, and behavior analysis," *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 4, pp. 1773–1795, 2013.

[18] H. Y. Cheng, C. C. Weng, and Y. Y. Chen, "Vehicle detection in aerial surveillance using dynamic Bayesian networks," *IEEE Transactions on Image Processing*, vol. 21, no. 4, pp. 2152–2159, 2012.

[19] Y. Tian, H. Dong, L. Jia, and S. Y. Li, "A vehicle re-identification algorithm based on multi-sensor correlation," *Journal of Zhejiang University SCIENCE C*, vol. 15, no. 5, pp. 372–382, 2014.

[20] S. S. Teoh and T. Bräunl, "Symmetry-based monocular vehicle detection system," *Machine Vision and Applications*, vol. 23, no. 5, pp. 831–842, 2012.

[21] G. Liu and X. Li, "Lane line identification based on genetic algorithm," *Journal of Hubei University of Technology*, vol. 1, p. 14, 2013.

[22] X. Tang, F. Gao, G. Xu et al., "Sensor systems for vehicle environment perception in a highway intelligent space system," *Sensors*, vol. 14, no. 5, pp. 8513–8527, 2014.

[23] L. Biqing, Z. Shiyong, and Q. Ming, "Research on picking identification and positioning system based on IOT," *International Journal of Online Engineering (iJOE)*, vol. 14, no. 7, pp. 149–160, 2018.

[24] L. Xi, S. Sun, L. Li, and F. Y. Zou, "Depth estimation from monocular infrared images based on SVM model," *Laser Infrared*, vol. 42, no. 11, pp. 1311–1315, 2012.

[25] B. Wang, B. F. Zhang, X. W. Liu, and F. C. Zou, "Novel infrared image enhancement optimization algorithm combined with dfocs," *Optik-International Journal for Light and Electron Optics*, vol. 224, p. 165476, 2020.

[26] Y. Zhao, H. Li, S. Wan et al., "Knowledge-aided convolutional neural network for small organ segmentation," *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 4, pp. 1363–1373, 2019.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 1938651, 9 pages
https://doi.org/10.1155/2022/1938651

Research Article

Sustainable Utilization Mode of International Communication of Cultural Tourism Resources Based on the Concept of Green Growth

Lu Wang

Tourism College, Hunan International Business Vocational College, Changsha, 410000 Hunan, China

Correspondence should be addressed to Lu Wang; 201701111000021@stu.hubu.edu.cn

Received 9 February 2022; Revised 26 February 2022; Accepted 7 March 2022; Published 19 April 2022

Academic Editor: Hoon Ko

Cultural tourism is gradually showing a rapid development momentum, but there are also some corresponding problems in the use and development of cultural resources and management mode, so how to maintain the sustainable development of cultural tourism resources and environment on the basis of economic stability is particularly important. Therefore, this paper takes the tea culture resources of Hunan Province as the research object, adopts the methods of questionnaire survey, and conducts a sample survey on the tourists in Hunan of 2020 Oct. It also analyzes the types of tourists' travel behavior, the reasons for tourists' travel, the most important tourist attraction elements, tourists' satisfaction, awareness of tea culture tourism, and awareness of Hunan Province. The value evaluation of cultural resource development, the evaluation of cultural tourism resources in Hunan Province, the evaluation of cultural tourism commodities and the shortcomings in tourism are analyzed. The results show that tourists pay more attention to the comprehensive experience, cultural resources, and environment; they are willing to learn more about cultural resource tourism; at the same time, they also reflect that there are insufficient development efforts, backward supporting facilities, and lack of relevant talents and laws in cultural resource tourism. Finally, this paper puts forward the sustainable utilization and management mode of cultural resources. By comparing before and after the trip, whether the respondents are "willing" to further understand the tea culture, 26.6% of them are "unwilling" before the trip and 73.4% are "willing," and 12.2% are "unwilling" after the trip and 87.8% of them choose "willing."

1. Introduction

With the development of transportation and national economy, people's living standard has been greatly improved, and consumption concept has also changed greatly. China has gradually entered the era of national tourism. In Hunan, in 2020, the number of tourists reached 693.1903 million, the total domestic tourism revenue was 825.842 billion yuan, the number of inbound tourists was 17400, and foreign exchange income from tourism was US $51.1662 million [1]. With the development of tourism, it is more comprehensive and relevant. It is driving the development of catering, accommodation, and transportation industries, expanding domestic demand and enhancing economic vitality. Tourism has made great contributions to China's economic growth and has become the spiritual pillar of the tertiary industry in China [2]. Accordingly, we need to maintain the sustainable development of tourism resources and protect the ecological environment, which requires us to master the causes, development, and evolution of all kinds of tourism resources and make rational development and utilization of tourism resources.

In Ted research, it is found that local residents in cultural tourism areas are more inclined to maintain the original local style and relatively closed living space [3]. In robes Li's survey, it is pointed out that residents can also participate in tourism planning and development and reduce the possibility of commercialization of local culture, and it is an effective means to achieve sustainable development [4]. Yu et al. classified the water culture and sports tourism and analyzed the current situation of the development of water culture and sports tourism resources in the research

area and the advantages of the development of cultural tourism [5].

In order to guarantee the green economic growth and realize the sustainable development of cultural tourism resources and environment, this paper first discusses the related concepts of green growth, cultural tourism resources, and sustainable development and explores the relationship between economic green growth and cultural tourism resources. Then, this paper conducts a questionnaire survey, makes an in-depth study on the utilization and management of cultural resources, and analyzes the main problems existing in the development, utilization, and management of cultural tourism resources at this stage. Based on this, the paper puts forward the management mode of promoting green economic growth and sustainable development of cultural resources. The research in this paper is not only beneficial to the balanced development of economy and resources but also lays a theoretical foundation for the future research.

2. Related Overview

2.1. Connotation and Characteristics of Green Growth. Green development is based on the development of green economy. As the mode of realizing the coordination of human, material, and environmental production in the process of sustainable development, economic development is the target practice and embodiment of the current sustainable concept [6]. Green growth emphasizes the unity of economic development and environmental protection. It not only changes the way of resource utilization but also protects resources and protects and recovers the system of resource regeneration. It also emphasizes the unity of economic development, social progress, and ecological protection and attaches importance to the connection between human and nature as well as social development.

The main characteristics of green growth are as follows [7]: first, overall characteristics. Green growth does not belong to a certain class or group. It belongs to all mankind and determines the development of all mankind. Second characteristic is comprehensive characteristics. Green growth not only represents economic development but also the common development of society, environment, economy, and resources. Third characteristic is endogenous characteristics. Green growth is the result of interaction of various variables in the whole resource. Fourth is continuity. Green development emphasizes intergenerational equity, which is the organic unity of human social development. Different countries are in different stages of development and have different green growth models and systems. Green growth should follow two development priorities: on the one hand, the balance between human and nature; on the other hand, maintaining harmony between people.

Internationally, "green growth" is regarded as an important way to solve the resource crisis and achieve sustainable development. South Korea and the European Union are the first to put forward and implement the concept of green growth. Only by building a harmonious relationship between economic growth and natural resources can we bring healthy and happy life to the world. Therefore, "green growth" is finally defined to protect the environment and ecological balance to achieve economic growth. The background of green economy is a kind of balanced economy aiming at maintaining the living environment of human beings and making rational use of energy resources [8]. It is the performance of green growth and green development in various economic fields, and it is an important economic form to coordinate environment and development issues. Its inevitability is that the development of green economy is the inevitable requirement of economic development and transformation. Second, the development of green economy is the only way to promote the construction of resource-saving and environment-friendly "two oriented society."

2.2. Concept and Characteristics of Cultural Tourism. Cultural tourism refers to the tourism with the purpose of appreciating foreign traditional culture, cultural relics, ancient buildings, cultural celebrity sites, and celebrity deeds [9]. It not only includes traditional tourism activities but also involves art, religion, customs, and other activities.

As a product of the combination of culture and tourism, cultural tourism is a new trend of tourism development, and it is inevitable for the development of tourism at this stage [10]. Actively explore and integrate the historical culture, cultural characteristics, and resource characteristics of the city, adapt to local conditions, conform to the trend of the times, and integrate with economic development [11, 12]. It has become an important direction of industrial development in the era of "green GDP," adding internal power to the sustainable and harmonious development of the city [12, 13]. With the increasing demand of the tourism market, cultural tourism gradually presents a rapid development momentum. People pay more attention to the pursuit of spiritual level and the perception of mind.

There is no unified understanding of cultural tourism in domestic and foreign studies. There are more cultural tourism in China and 37 world cultural heritages. It has national characteristics, regional characteristics, and unique cultural attraction, mainly including landscape culture and social human resources. Cultural resources bear the traditional culture for five thousand years. Appropriate development can make people understand history, understand hometown, love motherland, enhance national cohesion, enhance national cultural identity, and enhance national self-confidence and pride. It is of great significance to carry forward the traditional culture of the Chinese nation and strengthen globalization [14, 15]. It is far more meaningful than the sermon in the textbook. However, while appreciating culture, we are faced with various problems, such as the problems of the dilapidation of ancient buildings, the graffiti of places of interest, and the littering of garbage.

2.3. Core and Content of Sustainable Development Theory. With the rapid development of global economy and the deterioration of environment, sustainable development is derived. Sustainable development is to meet the needs of human development from generation to generation. The core of the theory is first strive to balance the relationship between man and nature and seek the rational nature for the harmonious development of man and nature. At the

same time, we should link human development with resource consumption, environmental deterioration, ecological environment, and so on. The essence of pressure reflects the harmony between human beings and coevolution between human beings and nature. Second, we should strive to achieve the harmony between human beings. Realize the harmony between human and nature. Its theoretical system is to require human beings to meet their own needs and ecological balance [16, 17]. Sustainable development includes three aspects: first, economic sustainable development; second, ecological sustainable development; third, social sustainable development.

Sustainable development follows three basic principles [18, 19]: first, the principle of fairness: to ensure the sustainability of resources for generations; second, the principle of sustainability: human development cannot cause damage to environmental resources; human development is constrained by population, environment, resources, technology, and social organizations; third, the principle of commonality: all countries in the world share resources and should follow the principle of sustainable development. Sustainable development involves many aspects, such as natural environment, ecological balance, and economic development [20]. While developing cultural tourism, we should keep the unity of social and ecological benefits [21, 22].

3. Research on the Sustainable Utilization Mode of International Communication of Cultural Tourism Resources Based on the Concept of Green Growth

3.1. Research Object. China is the birthplace of tea, which has thousands of years of history, forming a strong tea culture. Tea culture tourism refers to the project with tea and tea culture as the theme, through a series of colorful and complex tourism behaviors, to finally achieve physical and mental relaxation [23, 24]. Broad and profound tea culture is not only an important part of China's excellent traditional culture but also the basis for the development of tea culture tourism. Generally speaking, the influence of tea on cultural tourism is a tea area with rich and beautiful natural landscape and unique and profound historical and cultural landscape on the basis of increase and a series of scientific and feasible planning and design guided by various tourist experiences. The unique charm of tea cultural resources has been highly concerned by the tourism industry. However, there are few theoretical and practical researches on tea culture tourism in Hunan Province. Based on the theory of tourism science and Chinese tea culture, this paper studies the tea culture tourism resources, which not only can enrich and develop the existing tea culture but also has great practical significance [25].

3.2. Research Methods

(1) Literature collection method: combing the domestic and foreign scholars' research on tea cultural tourism resources in Hunan Province, to analyze and summarize the current situation of cultural rural tourism, as well as the utilization of cultural resources and existing problems

(2) Field investigation method: in order to master the local geographical and cultural resources of Hunan Province, it is necessary to carry out field investigation and interview with local residents

(3) Questionnaire survey method: in order to understand the tourism situation of tea culture resources in Hunan Province, the author conducted a questionnaire on local residents and foreign tourists in Hunan Province. During the national day of 2020, visitors to Hunan will be randomly distributed for sampling survey. The questionnaire mainly includes tourist motivation, service quality, and satisfaction. A total of 750 questionnaires were issued and 620 valid questionnaires were collected

3.3. Research Results

3.3.1. Analysis of Tourist Types. Among the questionnaires, the largest number of tourists was Chinese, including China, Hong Kong, and Macao compatriots, Taiwan compatriots accounted for 66.61%, 7.42%, and 9.52%, respectively, and the proportion of foreigners was 16.45%. The most foreign tourists are Americans, accounting for 47.06%. In fact, there are South Koreans and Japanese, accounting for 34.31% and 18.63%, respectively, shown in Tables 1 and 2.

3.3.2. Why Tourists Go Out. There are many channels for tourists to obtain tea culture; 50% of which are affected by their relatives and friends; 21.6% are through multimedia advertising and Internet; 26.4% are through hotels; 2% are other channels. The channels for tourists to learn about tea culture are shown in Figures 1 and 2.

3.3.3. Tourist Attraction Elements Most Valued by Tourists. The most important tourist attractions of tourists are shown in Figure 3. In the effective investigation, it is found that tourists pay the most attention to the comprehensive services in the tourism process, accounting for 43%, of which 2%, 4%, and 3% value catering features, accommodation, and transportation, respectively. It is worth thinking that culture and environment account for 32% and 16% of the total. It can be seen that people's assessment of tourism quality is mainly attracted by the local environment and culture.

3.3.4. Tourist Satisfaction Survey. It can be seen from Table 3 that tourists are most satisfied with tea cultural sites and natural scenery, with an average of 4.7521 and 4.7192. The average value of communication with local residents is the lowest, 3.2712, and the average value of tourism explanation is medium, but the tourist souvenirs are relatively satisfactory, which is also a surprise. On the one hand, tourists affirm the status of tea culture in tourism; on the other hand, tourists are not very satisfied with the local dialect or habits, which are the key points in the process of tea culture reconstruction.

Table 1: Composition of tourists from different countries.

Country	Number	Percentage
China (inland)	413	66.61%
Foreigners	102	16.45%
Hong Kong and Macao	46	7.42%
Taiwan	59	9.52%

Table 2: The top three foreign tourists in Hunan Province.

Country	Number	Percentage
America	48	47.06%
Korea	35	34.31%
Japan	19	18.63%

3.3.5. An Analysis of the Awareness of Tea Culture Tourism. By comparing before and after the trip, whether the respondents are "willing" to further understand the tea culture, 26.6% of them are "unwilling" before the trip and 73.4% are "willing," while 12.2% are "unwilling" after the trip and 87.8% of them choose "willing." Comparison of the importance of tea culture communication before and after tourism is shown in Figure 4.

3.3.6. Evaluation on the Development Value of Cultural Tourism Resources in Hunan Province. In the questionnaire, tourists rated the development value of cultural tourism resources in Hunan Province and scored the following 14 evaluation factors, with a total score of 10 points. Table 4 shows that the comprehensive average value of 14 evaluation factors is 7.3. The average is in the "good" class. Hunan Province has many types of cultural resources and high cultural value, which has certain development value. The highest score was given to social customs and practices, thus indicating that social customs and practices have a greater influence on the development of cultural tourism resources in Hunan.

3.3.7. Tourists' Evaluation of Cultural Tourism Resources in Hunan Province. In the questionnaire, 100 samples were randomly selected to evaluate the cultural tourism resources of Hunan Province. It can be seen from Table 5 that, in the survey, most tourists do not know much about the cultural tourism resources before they come, but after the tourists know about the cultural tourism resources of the scenic spot, they have a great interest in the cultural tourism resources of the scenic spot and think that the cultural tourism resources of the scenic spot are unique; although they are unique, the development is not comprehensive, and the whole development level is low.

3.3.8. Tourists' Evaluation of Cultural Tourism Commodities. In the questionnaire, 100 samples were randomly selected to evaluate the cultural tourism products. According to the statistics of the following data, tourists have a high recognition of cultural tourism commodities, especially Jingdezhen ceramics, with a good evaluation; in addition, tourists also have a good evaluation of root carving and oil painting; but the evaluation of tea is low, mainly because the tourism commodities are still in the early stage of development, many tourists do not know much, and their quality needs to be improved. Tourists' evaluation of cultural tourism commodities is shown in Table 6 and Figure 5.

3.3.9. Deficiencies Found in Tourism and Relevant Suggestions. In the process of investigation and interview of tea cultural resource tourism in Hunan Province, the author summarizes the factors that restrict the development of cultural tourism resources in scenic spots in the questionnaire and finds that there are mainly the following aspects:

(1) Insufficient and unreasonable development of domestic cultural resources: China's culture is broad and profound, with rich resources, and there are about ten thousand scenic spots available for tourism. Then, there are only a thousand scenic spots for tourism operation. Therefore, China's cultural tourism resources have not been fully explored, and the cultural tourism industry is still in the primary stage. Part of the scenic spots pursue commercialization, but lost the national flavor, so that tourists cannot understand the significance of tourism

(2) Backward supporting infrastructure: cultural tourism and its supporting facilities include transportation, accommodation, and catering. High-quality infrastructure can bring high-quality sense of body. Many counties and villages have rich culture, but due to the impact of economic development, the development of transportation is backward, which brings many inconveniences to tourists and loses its cultural significance. However, the backwardness and lack of catering, accommodation, and other supporting facilities also limited passenger flow

(3) Lack of talents related to cultural tourism operation, management, and development: China's cultural tourism started late and lagged behind in personnel training. The talents needed for tourism development cover its development, management, and operation fields, but at present, the threshold of China's tourism industry is low and the personnel are uneven. Its development and utilization mainly rely on government power; private tourism enterprises are few and lack special talents

4. Research on the Mode of Sustainable Utilization of International Communication of Cultural Tourism Resources

4.1. Sustainable Development and Utilization of Cultural Tourism Resources. Development refers to the mining of the existing undeveloped cultural resources or the deepening of the development and utilization of the resources with small development intensity [18]. From the perspective of

Figure 1: Age composition of tourists.

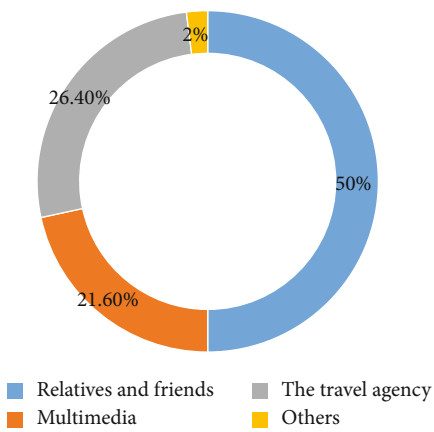


Figure 2: Channels for tourists to learn about tea culture.

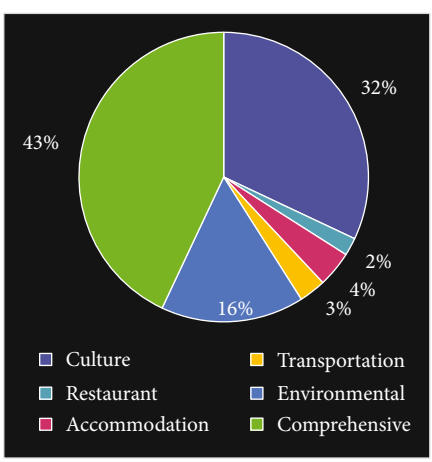


Figure 3: Tourist attraction elements that tourists value most.

protection objects, the protection and utilization of cultural resources mainly include two main bodies: the government and the city [19], which are divided into four modes: the government's full control type, the government's leading community participation, the community's independent type, and the laissez faire type. Based on the concept of green growth, the development and utilization modes of cultural tourism resources are as follows:

(1) "Government's complete control type" has low utilization rate and high degree of protection for cultural resources, which is characterized by government's monopoly management, mandatory, low risk of community holding the post, but low enthusiasm for participation. This model is suitable for the management of endangered cultural resources, the areas with serious resource damage, or the areas with improper development

(2) "Government-led community participation" has the highest utilization rate and protection degree of cultural resources, which is characterized by macro monitoring and guidance by the government. The development and utilization of resources by the community should meet the requirements of the government and be suitable for areas with good protection awareness and great development potential

(3) "Community autonomy" has a high utilization rate of cultural resources and a low degree of protection, which is characterized by the city's independent development and utilization in accordance with market demand, high enthusiasm, and prone to resource damage. It is suitable for the development of renewable resources with a strong sense of protection, regeneration, and a large number of renewable resources

(4) "Laissez faire" has the lowest utilization and protection of cultural resources. The characteristic of "laissez faire" is simply driven by economic interests, without any advantages, and will eventually be replaced by other models, which is not in line with the principle of sustainable development, combining the advantages and disadvantages of the four models

No matter which mode is adopted, the utilization and development of cultural resources need to be planned in the following three aspects:

Table 3: Average tourist satisfaction in Hunan Province.

	Mean	Tourist Se	Sd
Environmental health status	3.7253	0.8754	0.82156
Infrastructure	3.6721	0.8542	0.89542
Local transportation	3.5421	0.9952	0.95637
Accommodation and catering	3.4821	0.9174	1.0264
Introduction to tourism	3.4126	0.8657	0.89531
Souvenirs	3.7432	0.9852	0.8521
Natural scenery	4.7192	0.816	0.92617
Cultural relics	4.7521	0.6508	0.64261
Communication with local residents	3.2712	0.9521	0.92617

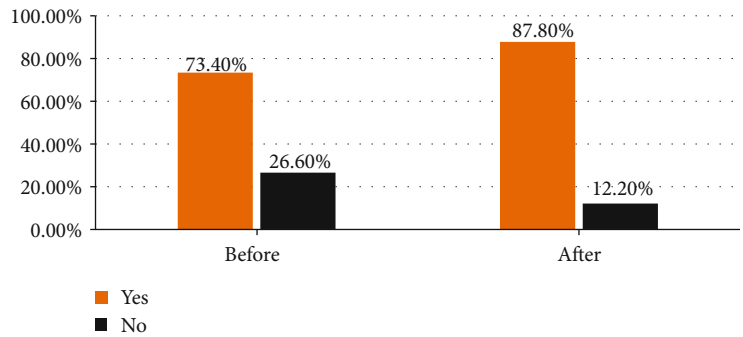


Figure 4: Comparison of the importance of tea culture communication before and after tourism.

Table 4: Development value of cultural tourism resources in Hunan Province.

	Evaluation factors	Mean
1	Infrastructure conditions	6.8
2	Value of sightseeing and recreation	6.2
3	Degree of rarity and strangeness	3.6
4	Integrity	5.7
5	Visibility and influence	7.8
6	Ability to experience	6.7
7	Technology and personnel	5.5
8	Geographic conditions	7.5
9	Natural environment	9.0
10	Local openness	9.5
11	Social customs and habits	10.5
12	Protection status	8.7
13	Local life and cultural tradition	4.8
14	To regional economy	9.4
	Comprehensive	7.3

(1) The development and utilization of resources need to be subdivided first. According to its category, structure, grade, and endangered degree, the comprehensive assessment and subdivision are carried out in order to select the appropriate development and utilization mode

(2) Divide the area: according to the distribution of cultural resources, the cultural resources should be protected in different areas, and corresponding systems and protection measures should be formulated to effectively protect the corresponding cultural resources. Guarantee the standardization and institutionalization of management process

(3) Complete supporting protection measures: determine supporting protection funds, responsible units, and responsible persons for different cultural resources, and formulate corresponding assessment management methods

4.2. Management Mode of Sustainable Utilization of Cultural Tourism Resources. Tourism system is composed of subject, object, intermediary, and carrier [19, 20]. The management of tourism resources involves systematic management. In essence, it is a harmonious coexistence between people and nature. From the point of view of purpose, it is based on the concept of green growth to maintain the sustainable development of economic development and ecological environment. Ensure the balance of ecological, economic, environmental, and tourism benefits. Based on the concept of green growth, my cultural tourism resource management mode is as follows.

Table 5: Statistics of tourist perception of Hunan cultural tourism resources.

Attribute	Features	Frequency ($N = 100$)	Percentage
How much do you know before traveling	Yes	39	39%
	So-so	23	23%
	No	38	38%
Interested or not	Yes	71	71%
	So-so	26	26%
	No	3	3%
Is it fully developed	Very full	0	0%
	Full	11	11%
	So-so	15	15%
	Not full	74	74%
Where to improve	Tourist participation	46	46%
	Souvenirs	38	38%
	Scenic spot explanation	57	57%
	Sign of scenic spot	62	62%
	Cultural performance	76	76%
	Others	56	56%

Table 6: Evaluation of cultural tourism commodities by tourists.

Tourist goods	Very good	Good	So-so	Bad	Very bad
JDZ ceramics	46	36	12	5	1
Tree-root carving	16	29	28	15	12
Bamboo weaving	15	14	36	16	19
Lacquer painting	16	25	41	8	10
Tea	25	20	30	10	15

4.2.1. On the Management of Tourists. First of all, build a perfect green interpretation system, including the propaganda slogans, billboards, and cultural explanations of cultural tourism areas. Increase green propaganda and voice to remind tourists to protect the environment and enhance their awareness of protection. Make full use of tour guide resources, improve tour guide's understanding of green tourism, and make tour guide become mobile propagandist, advocate, and supervisor.

Secondly, improve the rules and regulations. The administrative department shall give full play to its legal rights, resolutely crack down on the illegal acts of destruction of cultural resources, educate the tourists in laws, regulations, and rules before entering the cultural tourism area, and restrict or even fine the uncivilized acts of damaging the building appearance and the image of historical celebrities.

4.2.2. The Management of Tourism Resources. First, strengthen the protection of cultural tourism resources. The main reasons for the destruction of cultural resources are divided into natural factors and human factors. For the damage caused by natural factors, we can adopt advanced simulation technology and scientific methods to repair and establish a sound compensation mechanism for cultural resources. For the damage caused by human factors, it is necessary to establish and improve the management rules and regulations of the scenic spot, improve the laws and regulations, strengthen the inspection of law enforcement management personnel, and stop the behavior in time.

Secondly, strengthen the management of the scenic area community. In addition to the education of tourists, we should also strengthen the education of laws and regulations for the management departments and improve their awareness of the protection of cultural resources. When the foreign culture has a strong impact on the local culture, we should protect the local culture, prevent the loss of the local culture, and enhance the sense of identity and pride of the local culture. Only let the community benefit from the development of tourism can let them have stronger willingness to participate in the protection of cultural scenic spots. First, let the community and residents benefit from the development and management of cultural resources and improve the living standards of community residents; second, improve the quality of residents, carry out necessary cultural edification and education for residents, enhance the residents' understanding of the protection culture, and make residents participate in the protection of cultural resources.

4.2.3. The Management of Tourism Intermediary. Tourism intermediary mainly refers to the participation of tourists in various industries in the tourism process, such as catering, accommodation, and transportation. In order to realize the

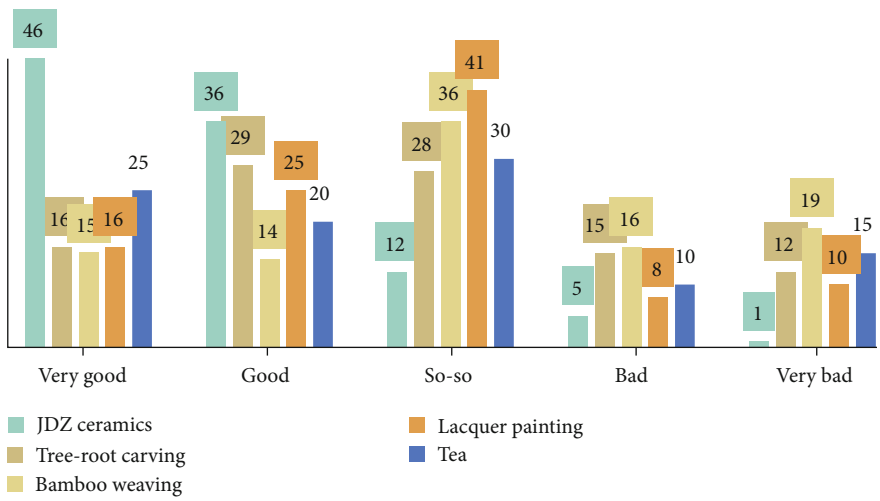


Figure 5: Tourists' evaluation of cultural tourism commodities.

sustainable development of ecotourism, we must adhere to the principles of regional management, government involvement, measurement dependence, and information dissemination: (1) establish a special management organization to carry out unified management of ecotourism, improve the management system, define the management objectives, formulate the management plan, improve the management system, strictly enforce the law, and strengthen supervision and feedback to avoid the bad situation of mutual prevarication; (2) the combination of tourism and environmental management makes environmental protection and business operation depend on each other and promotes the creation of green restaurants, green travel tools, and green hotels; green travel, reduce pollution; (3) strengthen green marketing management; develop green products and build a green promotion and distribution system.

4.2.4. The Management of Tourism Environment. Tourism environment is closely related to the subject, object, and intermediary of tourism. The management of each factor directly affects the health and sustainable development of tourism environment. In order to protect the tourism environment, we should coordinate the relationship among all elements of tourism, make full use of the means and methods of law, economy, planning, administration, science and technology, education, etc., control all behaviors and activities that may damage the tourism environment, maintain the high quality of the tourism environment, and coordinate the relationship between tourism activities and tourism protection so that tourism activities can meet the needs of tourists. In addition, it can keep its original culture from being damaged, so as to realize a series of control activities with unified environmental, social, and economic benefits. The main management measures are to establish and improve the green management system and green environmental policies and regulations system, to establish scientific and technological means of green environmental management, to develop the scientific research system of ecotourism environmental protection, and to improve the environmental awareness of the public.

4.3. Suggestions on Green Economic Growth and Sustainable Utilization of Cultural Tourism Resources. The sustainable utilization and development of cultural resources promote economic growth, and the promotion of green economic growth depends on the sustainable development of cultural resources. Therefore, facing the problems of green economic growth and sustainable utilization of cultural tourism resources, we should focus on two aspects.

(1) Build a long-term mechanism of green development: improve the ecological environment legislation and improve the relevant standards

(2) Strengthen the position of sustainability in green development in economic development. The assessment standard of resource and environment utilization should be formulated, the overload should be restrained, and the development of regional economy should be considered on the premise of the concept of sustainable development. We should not blindly pursue economic growth and ignore the harm to resources and environment

(3) The development of tourism resources must be dealt with in a way that balances the interests of all parties involved. The development of cultural tourism resources must take into account profitability, but must not disregard the protection of the ecological environment and natural resources, as well as the legitimate rights and interests of local residents

5. Discussion

Rational use of cultural resources and timely protection and publicity work can bring certain economic benefits and improve people's awareness of protecting the ecological environment. The utilization of cultural resources always needs to follow the principle of protection first and sustainable development. In the process of using cultural resources, research, planning, development, management, and supporting implementation should be done well. The development, utilization,

and management mode of cultural tourism resources based on green management proposed by the author is the development, utilization, and management mode proposed after a comprehensive analysis of the whole regional ecotourism system, which is an ideal mode for cultural tourism to achieve the goal of sustainable development. The development of cultural tourism activities from a cultural and ecological perspective is based on the fundamental development principle of sustainable development, focusing on the question of how to develop and protect the cultural and ecological environment in cultural tourism and proposing measures to address this from a planning perspective so that the sustainable development of cultural tourism is achieved from the outset of design.

Data Availability

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

Conflicts of Interest

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

[1] V. E. Dann, "The Development of the Tourism Industry Has Been Neglected," *Problems of Economic Transition*, vol. 59, no. 6, pp. 435–442, 2017.

[2] C. C. Lee and C. P. Chang, "Tourism development and economic growth: A closer look at panels," *Management*, vol. 29, no. 1, pp. 92-93, 2008.

[3] P. Teo and B. S. Yeoh, "Remaking local heritage for tourism," *Annals of Tourism Research*, vol. 24, no. 1, pp. 192–213, 1997.

[4] R. Y. Li, "Ethnic tourism: A Canadian experience," *Annals of Tourism Research*, vol. 27, no. 2, pp. 115–131, 2000.

[5] W. Yu and L. Kaicao, "Research on the development of water culture and sports tourism resources in Jiangsu Province," *Journal of Nanjing Institute of Physical Education*, vol. 1, pp. 46–50, 2014.

[6] J. Maroušek, R. Zeman, R. Vaníčková, and S. Hašková, "New concept of urban green management," *Clean Technologies and Environmental Policy*, vol. 16, no. 8, pp. 1835–1838, 2014.

[7] T. T. Nguyen, D. Camacho, and J. E. Jung, "Identifying and ranking cultural heritage resources on geotagged social media for smart cultural tourism services," *Ubiquitous Computing*, vol. 21, no. 2, pp. 267–279, 2017.

[8] E. Drimili, Z. Gareiou, and E. Zervas, "Public perceptions of the concept of green growth: application in Athens, Greece, during a period of economic crisis," *Environment Development and Sustainability*, vol. 22, no. 7, pp. 6053–6076, 2020.

[9] C. Lu and S. Liu, "Cultural tourism O2O business model innovation: a case study of CTrip," *Journal of Electronic Commerce in Organizations*, vol. 14, no. 2, pp. 16–31, 2020.

[10] K. Chatzina, "Social media activism on cultural tourism: a proposal for Paleochora in Aegina, Greece," in *Strategic Innovative Marketing and Tourism*, pp. 511–520, Springer, Cham, 2020.

[11] M. Adil, M. K. Khan, M. Jamjoom, and A. Farouk, "MHADBOR: AI-enabled administrative distance based opportunistic load balancing scheme for an agriculture Internet of things network," *IEEE Micro*, vol. 42, no. 1, pp. 41–50, 2022.

[12] L. Ogiela, M. R. Ogiela, and H. Ko, "Intelligent data management and security in cloud computing," *Sensors*, vol. 20, no. 12, article 3458, 2020.

[13] J. Y. Hong, H. Ko, and J. H. Kim, "Cultural intelligence and ARCS model for digital era," in *Proceedings of The 9Th International Conference on Web Intelligence, Mining and Semantics*, pp. 1–4, Seoul, Korea, June 2019.

[14] Y. P. Guilarte, R. C. L. González, F. X. A. Quintá, and X. C. M. Arce, "Heritage information system to promote cultural tourism and the use of digital mapping in primary and secondary schools," in *Cultural and Tourism Innovation in the Digital Era*, pp. 17–35, Springer, Cham, 2020.

[15] X. Zheng, "Countermeasures for Development of Fujian Cultural Tourism Based on SWOT Analysis," in *3rd International Conference on Culture, Education and Economic Development of Modern Society (ICCESE 2019)*, pp. 2125–2128, Atlantis Press, 2019.

[16] D. Rosmalia and L. E. Prasetya, "Development of cultural tourism area based on the spiritual space of Cirebon Keraton," in *IOP Conference Series: Earth and Environmental Science*, vol. 126, article 012076, IOP Publishing, 2018.

[17] A. Ibon, P. Beatriz, and E. Marisol, "Sustainable cultural tourism in urban destinations: Does space matter?," *Sustainability*, vol. 8, no. 8, article 699, 2016.

[18] W. Zyzak, "Sustainable tourism - sensible tourism," *Didactics Ecology Metrology*, vol. 20, no. 1-2, pp. 105–112, 2015.

[19] M. Adil, M. A. Jan, S. Mastorakis et al., "Hash-MAC-DSDV: mutual authentication for intelligent IoT-based cyber-physical systems," in *IEEE Internet of Things Journal*, p. 1, 2021.

[20] Z. Lv, Y. Li, H. Feng, and H. Lv, "Deep learning for security in digital twins of cooperative intelligent transportation systems," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–10, 2021.

[21] K. Dube and G. Nhamo, "Sustainable development goals localisation in the tourism sector: Lessons from Grootbos Private Nature Reserve, South Africa," *GeoJournal*, vol. 86, no. 5, pp. 2191–2208, 2020.

[22] M. F. Popescu, B. C. Chiripuci, A. Orîndaru, M. Constantin, and A. Scrieciu, "Fostering sustainable development through shifting toward rural areas and digitalization—the case of Romanian universities," *Sustainability*, vol. 12, no. 10, article 4020, 2020.

[23] L. Xu, H. Pan, Q. Lei, W. Xiao, Y. Peng, and P. Xiao, "Insect tea, a wonderful work in the Chinese tea culture," *Food Research International*, vol. 53, no. 2, pp. 629–635, 2013.

[24] K. Mongkol, "The critical review of new public management model and its criticisms," *Research Journal of Business Management*, vol. 5, no. 1, pp. 35–43, 2017.

[25] B. Eberhard, "Phase i cultural resources investigation for the proposed access pointe development in perrysburg township, wood county,ohio," *Science*, vol. 304, no. 5672, pp. 847–850, 2018.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 6743070, 13 pages
https://doi.org/10.1155/2022/6743070

Research Article

Implementation and Design of a Zero-Day Intrusion Detection and Response System for Responding to Network Security Blind Spots

Won-Seok Choi,[1] **Si-Young Lee**,[2] **and Seong-Gon Choi**[3]

[1]*Research Institute for Computer and Information Communication (RICIC), Chungbuk National University, Chungdae-ro 1, Seowon-Gu, Cheongju, Republic of Korea*
[2]*Xabyss Corp., Seongnam-si, Republic of Korea*
[3]*College of Electrical and Computer Engineering, Chungbuk National University, Chungdae-ro 1, Seowon-Gu, Cheongju, Republic of Korea*

Correspondence should be addressed to Seong-Gon Choi; choisg@chungbuk.ac.kr

Received 25 January 2022; Revised 10 March 2022; Accepted 17 March 2022; Published 8 April 2022

Academic Editor: Kuo-Hui Yeh

We propose a zero-day intrusion detection and response system (ZDRS) for responding to network security blind spots. An existing detection and response system for the analysis of zero-day attacks uses a full-packet storage method; thus, the longer the time required to recognize a zero-day attack, the higher is the packet storage capacity and inspection cost. To solve the storage capacity and inspection cost problems, we design an architecture for ZDRS for a retroactive security check (RSC) using a first-N packet storage method. For fast verification of the RSC result, we propose a drill-down session metadata searching algorithm using session and flow metadata. The ZDRS comprises a network processing unit and a security processing unit. The ZDRS network processing unit generates metadata for the RSC verification and efficiently stores packets using the first-N packet storage method. The ZDRS security processing unit performs the RSC and RSC verification using the drill-down session metadata searching algorithm. For ZDRS performance analysis, we implemented ZDRS and analyzed the storage efficiency, detection efficiency, and detection speed of ZDRS at the campus level. As a performance analysis result of implementation, the amount of data storage decreased from 3.4 terabyte to 62 gigabyte compared to the full-packet storage method by 1.82%, and storage efficiency increased by 54.84 times. Furthermore, the detection rate of 99.55% based on the first 5-kilobyte size compared to the full-packet storage method was confirmed.

1. Introduction

Information and communication technology (ICT) has been rapidly developing into new network types such as digital twin (DT), Internet of Things (IoT), cloud computing, and vehicle networks. The appearance of new networks and technologies creates new applications and increases network traffic. With developments in ICT, new security requirements are required, and new types of attacks have also increased. However, the response to new and unknown network attacks has no choice but to react reactively [1–8].

Keeping up with hackers' efforts for attacks that mimic normal behavior is difficult. There is no perceptible manner to prevent hackers from attacking blind spots before the security detection rule is implemented. Therefore, to respond to such blind spots, a general procedure for recognizing, analyzing, and responding to hacker attacks, creating rules, and applying rules is necessary [9]. However, the application of the new detection rule only responds to attacks after that point in time, and checking for the attacks

previously received is not possible. For example, for an attack that leaks user information, analyzing which user information has been leaked is necessary in order to prepare future countermeasures. Therefore, analyzing the data from the point of attack until the detection rule is applied is necessary.

As mentioned above, a zero-day attack refers to an attack that occurs during the time from the creation of a new/variant security threat until the detection technology for blind spots is provided. When a new security threat occurs, it takes physical time for a new security rule or technology to be recognized and dealt with. This time is called the zero-day attack period, and a method that can quickly and efficiently detect and analyze attacks during the zero-day attack period is required.

Several studies have been conducted to counter zero-day attacks. Most research on zero-day attacks has focused on detecting security threats in the past. It can be classified into provenance tracking methods [10], detection method for software defined network (SDN) [11], signature-based detection methods [12, 13], machine learning-based method [14–16], anomaly detection methods [14–20], and specification-based detection methods [20–23]. However, the detection of a zero-day attack requires stored data during the zero-day attack period.

The detection of the zero-day period is performed through forensic analysis of all packets stored during the zero-day period. It stores all incoming packets for a certain period and performs forensic analysis on all stored packets using the abovementioned detection method-related studies [24–27].

For blind spots, inspection through forensic data analysis requires storing all packets during the zero-day attack period and subsequently checking and analyzing all packets. However, as the zero-day attack period increases, the storage capacity and cost increase rapidly. In addition, the analysis time increased.

Therefore, in this study, we classify and define blind spots for zero-day attacks. We propose a zero-day intrusion detection and response system (ZDRS) for coping with blind spots that can be solved. The ZDRS solves the problem of existing forensic storage by drastically reducing the storage cost through the first-N packet storage method.

We propose a recheck solution for blind spots using a system architecture for a regression security check, a metadata generation algorithm, and a metadata search algorithm. We solve the problem of existing forensic storage by drastically reducing the storage cost through the first-N packet storage method in the ZDRS. In addition, we analyze the number of detections, detection rate, storage efficiency, and detection speed through a campus-level implementation on the ZDRS.

The main contributions of this paper are as follows:

(i) We classify and define the type of blind spot

Existing studies have considered the security blind spot as one area without classification. However, we classified them into three categories to respond to more precise security blind spots. In particular, security blind spots were classified into three types: operational security blind spots, temporal security blind spots, and unknown security blind spots according to the timing of security responses. Existing studies have tried to solve the unknown security blind spot. However, we introduce the threats of operational security blindness and time security blindness and present solutions for resolving operational security blindness and security blindness in time

(ii) We present solutions for resolving operational security blind spots and time security blind spots

One problem of the solution to solve the operational security blind spot and temporary blind spot is packet storage issue. In existing researches, packets are inspected through a forensic method, so all packets for a certain period are required. However, the forensic method increases cost for packet storage because the longer the period to be inspected and the greater the flow of data, the more packets need to be stored. To solve this problem, we design an architecture for the ZDRS using the first-N packet storage method. Moreover, we propose the RSC method using the RSC algorithm for intrusion detection in a zero-day period and propose a drill-down session metadata searching algorithm and metadata generation method to support the correct verification of the RSC.

(iii) We verify the proposed ZDRS through implementation test

In real network environments where various protocols coexist, implementation is performed in a campus network to analyze the accuracy and efficiency of the ZDRS. In order to test in the real network environment, it was approved by Chungbuk National University, and we collected data at the egress point of the Chungbuk National University network. We also analyze the ZDRS test results based on the collected data.

The remainder of this paper is organized as follows. We review the related work in Section 2. We classified the blind spots in Session 3. The ZDRS is introduced in Section 3.2. We provide the implementation results of the ZDRS in Section 4, , and the concluding statements are presented in Section 5.

2. Related Work

In this section, we introduce research related to zero-day intrusion detection as shown in Table 1, such as the provenance tracking method, detection method of a software-defined network, signature-based detection method, anomaly detection method, and specification-based detection method.

Sun et al. proposed a provenance tracking method that is a probabilistic approach and implemented a prototype system ZePro for zero-day attack path identification. They used Bayesian networks to identify the zero-day attack paths. To capture the zero-day attack, a dependency graph called an

Table 1: Classification of related work on zero-day attacks.

Category	Reference
Provenance tracking method	[10]
Detection architecture of a software-defined network	[11]
Signature-based detection method	[12, 13]
Machine learning-based detection method	[14–16]
Anomaly detection method	[14–20]
Specification-based detection method	[21–23]

object instance graph is first built as a supergraph by analyzing system calls. To further reveal the zero-day attack paths hidden in the supergraph, the system builds a Bayesian network based on the instance graph [10].

Al-Rushdan et al. proposed a detection method designed and implemented for SDN using a modified sandbox tool called Cuckoo. The provenance tracking method is a probabilistic approach that implements a prototype system ZePro for zero-day attack path identification [11].

Hindy et al. and Bherde and Pund proposed a signature-based detection method [12, 13].

Hindy et al. proposed an autoencoder implementation for detecting zero-day attacks. CICIDS2017 and NSL-KDD were used for the evaluation. Hindy et al.'s model benefits greatly from autoencoder encoding–decoding capabilities [12].

Bherde and Pund proposed a technique for detecting zero-day attacks by using signature-based and knowledge-based methods. This system is a combination of signature-based and knowledge-based systems. [13].

Patidar and Khanderwal, Mirsky et al., and Bovenzi et al. proposed machine learning-based detection method [14–16]. In general, machine learning-based detection methods are included in anomaly detection methods that use machine learning as a method to find anomaly packets.

Patidar and Khanderwal proposed a solution based on the detection of zero-day malware using machine learning and artificial intelligence. They made an idea-level proposal, and the performance of the proposed method has not verified [14].

Mirsky et al. introduce a kitsune and evaluate its performance. The kitsune is a neural network-based network intrusion detection system which has been designed to be efficient and plug-and-play. It accomplishes this task by efficiently tracking the behavior of all network channels,and by employing ensemble of autoencoders for anomaly detection. They discussed the framework's online machine learning process in detail and evaluated it in terms of detection and runtime performance [15].

Bovenzi et al. proposed a hierarchical hybrid intrusion detection (H2ID), a two-stage hierarchical network intrusion detection approach. The H2ID performs anomaly detection via a novel lightweight solution based on a multimodal deep autoencoder and attack classification, using soft-output classifiers. The attack classification can use any machine/deep learning-based supervised classifier [16].

Patidar and Khanderwal, Mirsky et al., Bovenzi et al., Innab et al., Duessel et al., Aygun and Yavuz, and Bostani and Sheikhan proposed an anomaly detection method [14–20].

Innab et al. proposed a hybrid system between an anomaly-based detection system and a honeypot to detect a zero-day attack, consequently to integrate both approaches in one hybrid model as enhanced solution of detecting the zero-day attack that may occur in the system [17].

Duessel et al. proposed a detection method for zero-day attacks using context-aware anomaly detection at the application layer. They presented a new data representation, called cn-grams, that allows to integrate syntactic and sequential features of payloads in a unified feature space and provides the basis for context-aware detection of network intrusions. Also, they conducted experiments on both text-based and binary application-layer protocols which demonstrate superior accuracy on the detection of various types of attacks over regular anomaly detection methods [18].

Aygun and Yavuz proposed a network anomaly detection method with stochastically improved autoencoder-based models, to detect zero-day attacks with high accuracy. The key factor of their models is the threshold value which was determined using a stochastic approach rather than the approaches available in the current literature. Their models were tested using the KDDTest+ dataset contained in NSL-KDD, and their research achieved an accuracy of 88.28% and 88.65%, respectively [19].

Bostani and Sheikhan proposed a hybrid of anomaly-based and specification-based IDS for Internet of Things using unsupervised OPF based on the MapReduce approach. Specifically, the specification-based intrusion detection agents that are located in the router nodes analyze the behavior of their host nodes and send their local results to the root node through normal data packets. In addition, an anomaly-based intrusion detection agent that is located in the root node employs the unsupervised optimum-path forest algorithm for projecting clustering models by using incoming data packets. In experimental results, their method showed achieve true positive rate of 76.19% and false positive rate of 5.92% when both sinkhole and selective-forwarding attacks were launched [20].

Bostani and Sheikhan, Siu and Panda, Althubaity et al., and Altaf and Majeed proposed a specification-based detection method [20–23].

Siu and Panda proposed a specification-based detection method for attacks in a multiarea system. They described the implementation of a three-area system model. Also, they assessed the risk and devise several intrusion scenarios. Specifically, they injected false data into the frequency measurement and Automatic Generation Control (AGC) signals. And then, they developed a rule-based method to detect anomalies at the system-level [21].

Althubaity et al. proposed a hybrid specification-based intrusion detection system for rank attacks in 6TiSCH networks. Specifically, they proposed a hybrid specification-based intrusion detection system (IDS) that consists of centralized and distributed modules installed on the sink and RPL nodes, respectively, to prevent nodes from selecting an intruder as their successors. This method also eliminates

intruders' chances of becoming a time source and disrupts the synchronization of 6TiSCH networks [22].

Altaf and Majeed proposed a specification-based intrusion detection model for OLSR. The specification-based approach analyzed the protocol specification of an ad hoc routing protocol to establish a finite-state-automata (FSA) model that captures the correct behavior of nodes supporting the protocol. Then, this method extracted constraints on the behavior of nodes from the FSA model. Thus, this approach reduced the intrusion detection problem to monitoring the individual nodes for violation of the constraints [23].

These studies have been conducted to counter zero-day attacks. Most research on zero-day attacks has been focused on detecting security threats in the past. However, the detection of a zero-day attack requires stored data during the zero-day attack period. There are no studies on how to store zero-day data.

These studies used forensic analysis on all packets stored during the zero-day period or perform for real-time detection. Consequently, as the zero-day attack period increases, the storage capacity and cost increase rapidly. In addition, the analysis time is increased. Also, these studies do not consider how to store and retain data during the zero-day period.

Therefore, we propose a ZDRS to cope with blind spots that can be solved. The ZDRS solves the problem of existing forensic storage by drastically reducing the storage cost through the first-N packet storage method.

3. Zero-Day Intrusion Detection and Response System (ZDRS)

In this section, we first classify and define the type of blind spot. Subsequently, the architecture of the ZDRS is introduced, which is based on a first-N packet storage method, RSC algorithm, and drill-down metadata searching algorithm. In addition, for metadata generation and packet storage, the method of session metadata generation, first-N packet storage method, RSC algorithm, and drill-down searching algorithm are described.

3.1. Blind Spot. Figure 1 shows the blind spots of the network security area. t_1, t_2, t_3, and t_x denote points of time of event related to intrusion. Particularly, t_1, t_2, and t_3 are points of time where detection information is provided. t_x is the point of time where a new intrusion occurred without it being recognized. This intrusion may not be a zero-day intrusion. However, we consider for case of t_x is a zero-day intrusion.

t_1 is the point of time that provided detection information for any intrusion recognized before t_1.

t_2 is the point of time when detection information is provided regarding the intrusion of t_x. It means that zero day ends for the threat at t_x.

t_3 means the situation that detection information is not provided yet for the intrusion of t_x. It means that zero day does not end for the threat at t_x.

We classify security blind spots into three types: operational blind spots, temporal blind spots, and unknown blind spots.

An operational blind spot is defined as a blind spot that can detect an intrusion due to the presence of detection information; however, it was not recognized by mistake.

In Figure 1, detection information was released at the point of time t_1 for any intrusion that occurred before t_1, but there is a delay for reflecting the new detection rule or an operator's mistake. In the operational blind spot, detection rules have already been provided, so confirming an intrusion for stored past data is necessary. At this point of time, the important thing is a time of stored past data. The longer time is saved, the more data needs to be saved.

A temporal blind spot is defined as a blind spot where the intrusion occurred in the past, but detectable information is not recognized because that information is released today.

In Figure 1, a new intrusion occurred in t_x, but it was not recognized. In t_2, the detection rule for the intrusion was provided. Therefore, performing inspection on past data using the detection rule provided in t_2 is necessary. As for the temporal blind spot, the number of packets that need to be inspected increases as the time increases when the detection rule is delayed. Therefore, it is important how quickly detection rules are provided and applied, or how long period packets are stored for a zero-day attack period.

Unknown blind spots are defined as blind spots that have not been recognized because intrusion has occurred in the past, and detectable information has not yet been identified. In Figure 1, it is a situation corresponding to t_3, indicating a situation wherein not only the release of detection rules but also the intrusion was not recognized. In other words, because they do not even know that there was an attack, they cannot respond to anything.

Extensive research is being conducted to solve blind spots. However, most research on blind spots has been conducted for unknown blind spots. The unknown blind spot is an area that cannot be solved in the current system, which has no choice but to reactively respond to unknown network attacks. Therefore, research on unknown blind spots is being used as an abnormal behavior-based method. To fundamentally solve the unknown blind spot, studying a different method from the current mechanism is necessary. Operational blind spots are not a technical research area related to network attacks but rather a management tool research area. A study on the temporal blind spot used forensic data analysis to check packets during the zero-day period.

Most studies involving the blind spot are studies on the detection method, as mentioned in Section 2. However, the common problem with solutions is the storage space.

After recognition of the intrusion, a detection check should be performed on the stored packet using a released detection rule. However, as the time for intrusion recognition and provision of detection rules increases, the number of stored packets increases, and it takes considerable time to perform a check on the increased amount of data.

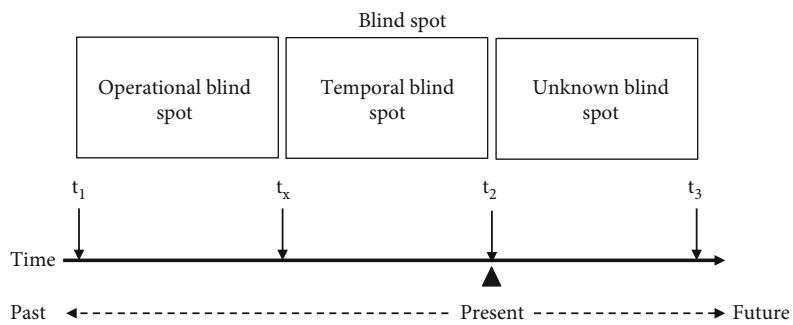


Figure 1: Classification of blind spot.

In general, the method of performing detection and checking for intrusion uses a forensic inspection method that stores all packets and checks all packets [24–27].

However, to solve the common storage problem of blind spots, we use the first-N packet storage method to reduce the packet storage capacity while maintaining a certain level of detection rate.

3.2. Architecture of ZDRS. The main purpose of the ZDRS is to reduce the amount of packet storage and to perform and verify fast RSC while maintaining a high detection rate in detecting and responding to zero-day attacks.

Figure 2 shows the architecture of the ZDRS to achieve the abovementioned purpose of the ZDRS. The ZDRS is generally located at the edge between the local area network (LAN) and the gateway.

The ZDRS is composed of a ZDRS network processing unit and a ZDRS security processing unit.

The ZDRS network processing unit performs packet capture, general packet processing, and flow and session metadata generation. This creates metadata to be used for RSC verification. In addition, it performs the operation of a general network processing unit.

The ZDRS network processing unit is composed of the following:

(i) *Packet Capture Interface Module*. It captures packets in the network.

(ii) *Packet-Processing Module*. It performs conventional packet processing.

(iii) *ZDRS Flow Processing Module*. It creates flow metadata of packets.

(iv) *ZDRS Session Processing Module*. It creates session metadata of packets.

The ZDRS network processing unit is a software module that operates in a host system in the case of a 1 Gbps bandwidth and operates in a separate hardware smart network interface card (NIC) if the bandwidth is above 10 Gbps.

For bandwidths of 10 Gbps, because a large amount of processor resources is required for packet capture, packet processing, and storage, additional hardware smart NIC is installed to satisfy the performance of the ZDRS network processing unit.

The ZDRS security processing unit performs intrusion detection and response to zero-day attacks that store packets, performs RSC, and verifies the results of the RSC.

The ZDRS security processing unit is composed of following:

(i) *Storage Management Module*. This module is used to set the configuration of the first-N packet storage method.

(ii) *Retroactive Security Check Module*. The RSC is performed, and the RSC result is verified using the drill-down session metadata searching algorithm.

(iii) *External Interface Module*. This module is used to set a retroactive security check module and to confirm results from the outside.

(iv) *Storage Device*. It is conventional storage I/O device.

The ZDRS captures incoming packets and performs general packet-processing procedures in the packet-processing module. In addition, according to the configuration of the first-N packet storage method set in the storage management module, the flow metadata and session metadata are created in the ZDRS flow processing module and the ZDRS session processing module. Then, the packets are stored in the storage device.

Thereafter, when a zero-day attack is recognized and a detection rule is generated, the corresponding rule is updated to the retroactive security check module by the external interface module. The retroactive security check module performs RSC for stored packets based on the updated detection rule. Then, the retroactive security check module verifies the RSC result using the drill-down session metadata search algorithm based on session metadata and flow metadata.

3.3. Metadata Generation and Packet Storage. In this section, we introduce the metadata generation method and the first-N packet storage method.

3.3.1. Metadata Generation. For verification of the RSC result, we used a drill-down searching algorithm. The drill-

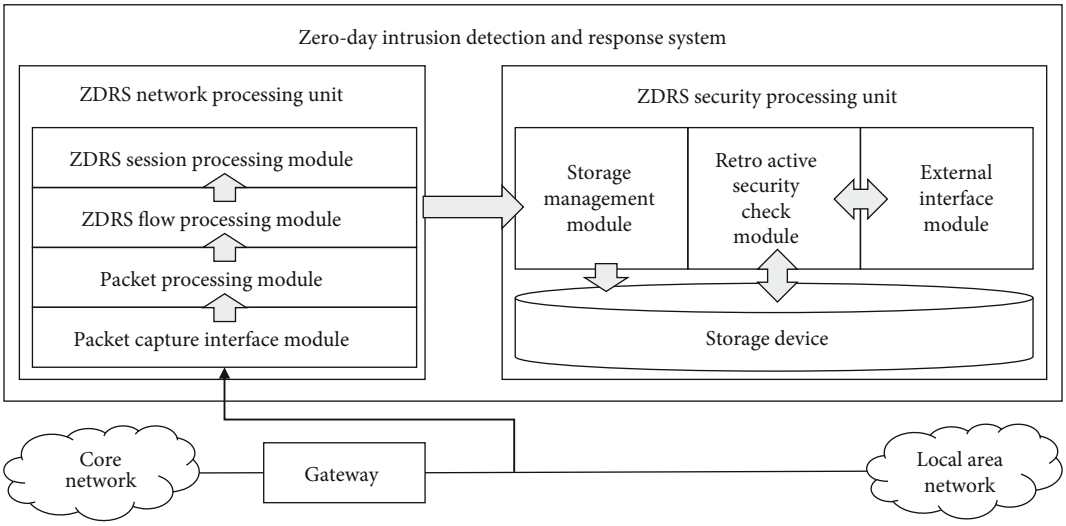


Figure 2: Architecture of the ZDRS.

```
struct flow _metadata
{
    uint64_t id; //Flow ID
    five_tuple tuple; //5-tuple information
    uint32_t packet_count; // The number of packet of the flow
    timestamp start_time; //Flow start time
    timestamp end_time; //Flow closed time
    tcpflag tcp_flag; //TCP flag information of the flow
    uint16_t flow_state; //Flag state information: active or closed
    uint64_t session_id; //Session id of the flow
    uint64_t bytes; //Total packet size(bytes) of the flow
}
```

Figure 3: Flow metadata structure.

```
struct session_metadata
{
    uint64_t id; //Session ID
    five_tuple tuple; //5-tuple information
    uint32_t flow_count; //The number of flow of the session
    uint32_t packet_count; //The number of packet of the session
    timestamp start_time; //Session start time
    timestamp end_time; //Session closed time
    tcpflag tcp_flag; //TCP flag information of the session
    uint16_t session_state; //Session state information: active or closed
    uint64_t bytes; //Total packet size (bytes) of the session
}
```

Figure 4: Session metadata structure.

down searching algorithm uses the flow metadata and session metadata for high-speed searching. We consider that the same flow is a set of packets having the same 5-tuple information within an active timeout and an inactive timeout. Therefore, to classify the flow, the active timeout and inactive timeout or TCP FIN flag of the packet is used. We consider that the same session is a set of packets with the same 5-tuple information until the application starts and ends communication. Therefore, to classify the session, we use the inactive timeout or TCP FIN flag of a packet. This criterion for classifying sessions and flows is used for packet storage in the first-N packet storage method and for generating the metadata. In general, the start and end of a session are recognized by analyzing the packet header information. Therefore, to reduce the burden of recognizing a session by analyzing all packets, we generate metadata and perform a metadata-based quick search.

First, we create the flow metadata from a packet. Figure 3 shows the structure of flow metadata. The flow metadata comprises an *id*, *tuple*, *packet_count*, *start_time*, *end_time*, *tcp_flag*, *flow_state*, *session_id*, and *bytes*.

The *id* is a flow identifier that is randomly allocated a value to not be duplicated for the classification of flows. The *tuple* is 5-tuple information comprising a source IP address, destination IP address, source port number, destination port number, and protocol. The *packet_count* denotes the number of packets in the flow. The *start_time* is the start time of the flow. The *end_time* is the closed time of the flow. The *tcp_flag* is the TCP flag of the flow. The *flow_state* is flag state information on whether the flag is in an active or a closed state. The *session_id* is the session id of the flow, and the *bytes* denote the total packet sizes of the flow.

Session metadata is created using flow metadata. Figure 4 shows the structure of session metadata. The session metadata comprises an *id*, *tuple*, *flow_count*, *packet_count*, *start_time*, *end_time*, *tcp_flag*, *session_state*, and *bytes*.

The *id* is a session identifier that is randomly allocated a value to not be duplicated for the classification of sessions. The *tuple* is 5-tuple information comprising a source IP address, destination IP address, source port number, destination port number, and protocol. The *flow_count* is the number of flows in the session. The *packet_count* is the number of packets in each session. The *start_time* is the start time of the session, and the *end_time* is the closed time of the session. The *tcp_flag* is the TCP flag information for the session. The *session_state* is a session state information on whether the session is in an active or closed state, and the *bytes* denote the total packet sizes of the session.

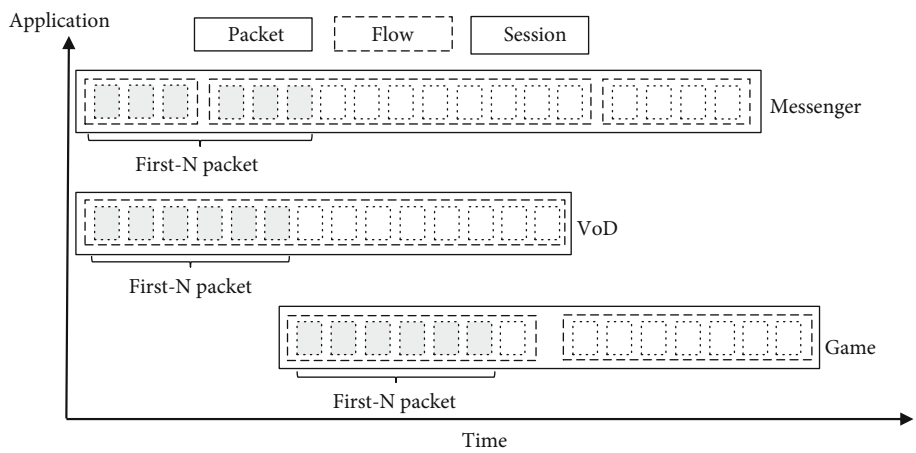


Figure 5: First-N packet storage method of ZDRS.

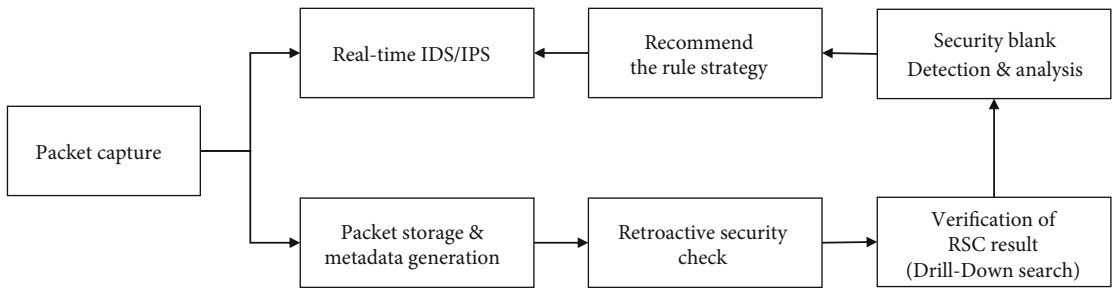


Figure 6: Operation of the RSC in the ZDRS.

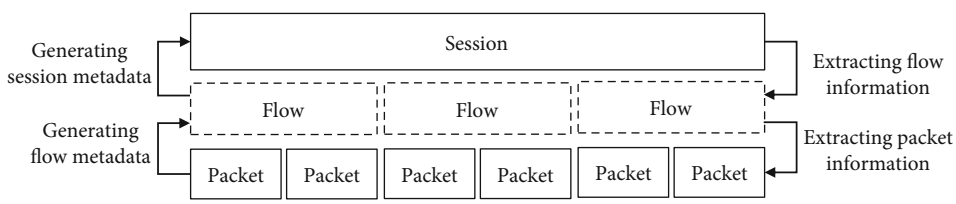


Figure 7: Drill-down search algorithm based on session and flow metadata.

The generated session metadata and flow metadata are used in the drill-down searching algorithm for fast packet search to verify the RSC result.

3.3.2. First-N Packet Storage Method. Figure 5 shows the first-N packet storage method of ZDRS, including examples of messengers, video on demand (VoD), and game applications [28]. The first-N packet storage method stores n-or n-sized packets for a session. Assuming that a session of a messenger application comprises 3 flows and a total of 18 packets, the initial 6 packets are stored to recognize the flow. In the case of VoD application, when a session is composed of 1 flow and a total of 14 packets, the initial 6 packets are stored. In the case of a game application, when a session comprises 2 flows and a total of 14 packets, the initial 6 packets are obtained. In other words, it stores the initial 6 packets for any session regardless of the application type and flow.

The length of the session differed depending on the application. In addition, the number of flows and packets included in each session is different. However, if the initial few packets can be checked for a session, it can confirm the network signature [29]. In Park et al.'s research, the optimal number of n packets was 5 [29]. However, the optimal number may vary depending on the network conditions and user requirements. The n unit can also be used as the number or length of packets to be stored. In particular, when used as the meaning of the length of a packet for a continuous syn attack, its characteristics can be better captured. Therefore, in this study, we implemented and tested its length.

When a detection rule for a zero-day attack is issued, the stored packets using the first-N packet storage method are checked by the RSC module in the ZDRS.

3.4. Retroactive Security Check. In this section, we introduce the RSC algorithm and the drill-down session metadata search algorithm in the RSC module. In the RSC module, the RSC operation and the drill-down session metadata

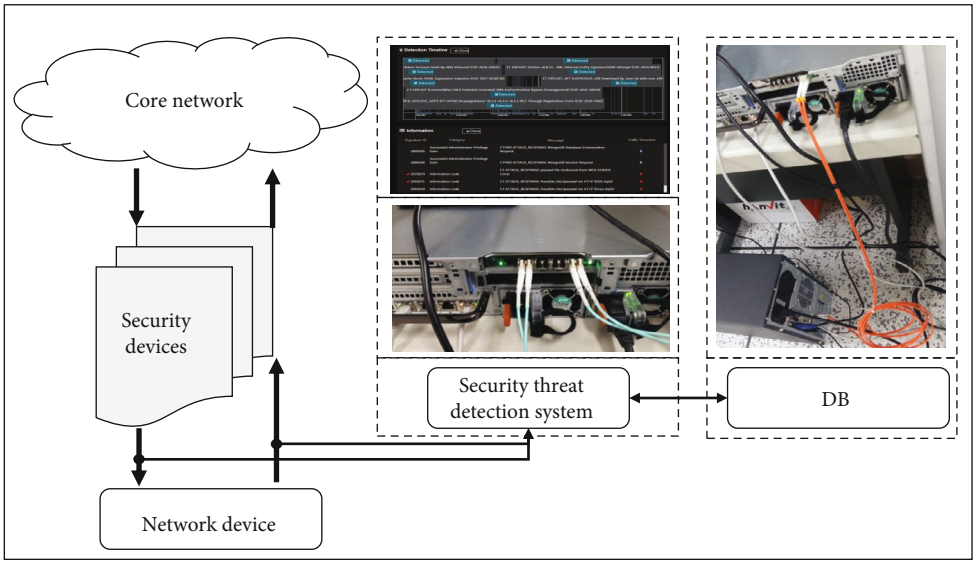


Figure 8: Implementation environment.

Table 2: Specifications of ZDRS security processing unit.

List	Specification
Processor	Intel Zeon Silver 4110 2.1 GHz (8 core; 16 thread)
Memory	16 GB (8 GB RDIMM ×2)
Operating system	Linux (Debian 8.9)
HDD	SAS 7.2 k RPM 48 TB (8 TB ×6)

Table 3: Specifications of ZDRS network processing unit.

List	Specification
Network processor	36-core 1.2 GHz Tile-Gx36 processor
Memory	DDR3 SODIMM type 8 GB
Interface	1/10 Gbps (SFP/SFP+) 4-port
PCIe type	PCIe Gen2 8-lane
Supported operating system	Linux: CentOS, Red Hat
Performance	Full 20 Gbps transfer to host

Table 4: Stored packets and storage rate of the ZDRS test.

First-N size	5 kB	10 kB	20 kB	30 kB	50 kB	100 kB	300 kB	1 MB	Full
Stored packets (GB)	62	85	128	159	209	291	445	662	3,400
Storage rate (%)	1.8	2.5	3.8	4.7	6.1	8.6	13.1	19.5	
Number of detections	13,882	13,885	13,890	13,895	13,895	13,900	13,900	13,902	13,944

searching algorithm were used to detect zero-day attacks in the past.

3.4.1. Operation of Retroactive Security Check. Figure 6 shows the operation of the RSC in ZDRS. The RSC conducts a check for false-negative packets for a zero-day period. When a packet is captured by the ZDRS in the network, the ZDRS conducts real-time IDS/IPS for all captured packets. Simultaneously, packets are stored according to the setting of the first-N storage method. Subsequently, the ZDRS flow and session metadata processing module create metadata before storing the packets.

In the ZDRS, the RSC module performs RSC to respond to operational blind spots and temporal blind spots. To respond to a temporal blind spot, the RSC module performs the RSC by updating a detection rule or confirming a zero-

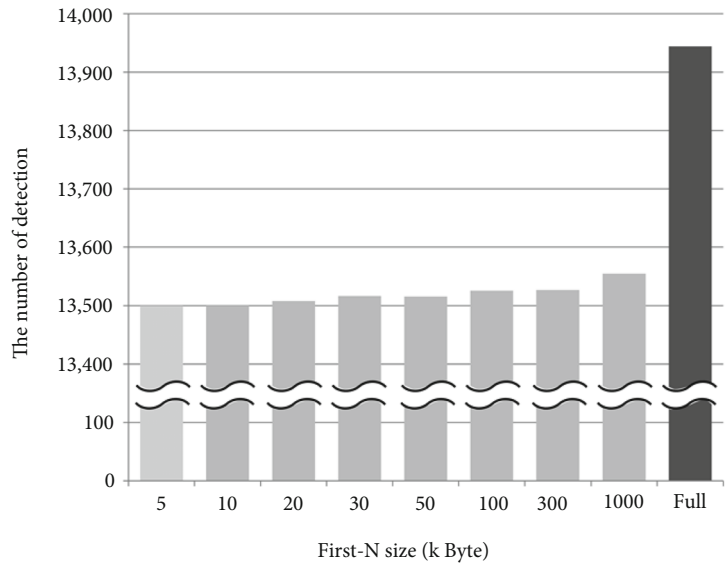


Figure 9: Number of detections of each first-N storage method versus the full-packet storage method.

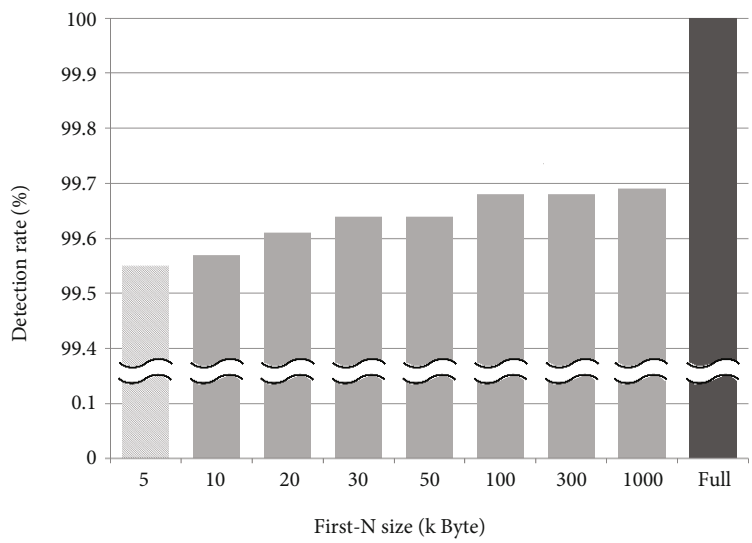


Figure 10: RSC detection rate for each first-N storage methods versus the full-packet storage method.

day attack. For an operational blind spot, the RSC module conducts periodic RSC according to the cycle set by the operator.

After the RSC is executed, the RSC execution result is verified through a drill-down search based on previously created metadata. Thereafter, a security check was performed through the continuous detection and analysis of the security blank. When a new rule strategy is created, it is reflected in the real-time IDS/IPS to remove the blind spots for a zero-day attack.

3.4.2. Drill-Down Session Metadata Searching Algorithm. Figure 7 shows the drill-down search algorithm based on metadata to quickly verify the RSC results. After the RSC module receives a search requirement owing to the predetermined detection rule or the updated detection rule, this module conducts a check for zero-day attacks in a blind spot. Subsequently, the drill-down search algorithm performs the following procedure to obtain false positives in the RSC module. The RSC module finds an objective session to compare false positive packets to session metadata by using the session metadata information. The RSC module then obtains an objective flow to compare the requested flow to the flow metadata. The RSC module subsequently obtains an objective packet to verify the RSC results to prevent false positive detection and then rechecks the objective packet.

The drill-down search algorithm does not search all packets or flows but instead searches for information in the sessions, flows, and packet order. Therefore, the effect of high-speed search was obtained.

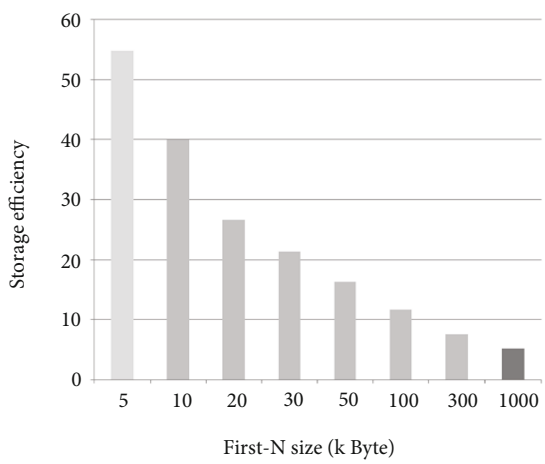


Figure 11: Storage efficiency for first-N storage methods versus the full-packet storage method.

4. Implementation

In this section, we introduce the implementation environment and the results of the proposed ZDRS. We analyzed the detection rate, storage efficiency, and RSC time in the ZDRS versus the full-packet storage method.

4.1. Implementation Environment. The implementation environment for the performance analysis of the ZDRS is described. Figure 8 shows the implementation environment at the campus level. The ZDRS was connected to the egress point of Chungbuk National University's local network to collect data. The collected data were approved by Chungbuk National University to be used only for this study.

In general, a campus-level network may consist of one or more LANs within the campus network and is connected to an external core network through a network device such as a router or gateway. Here, when connecting to the core network, it passes through a security device to block security threats. For test of the ZDRS, we mirror packets between the security device and the network device to the ZDRS. In addition, a database is configured to store rule sets, inspection history, and traffic information. We used the emerging threat Suricata open rule set for the detection rule. The flow inactive timeout, flow active timeout, and session inactive timeout were set as 15 s, 1800 s, and 1860 s, respectively.

Table 2 lists the specifications of the ZDRS security processing unit. The processor is an 8-core 16-thread Intel Zeon Silver 4110 with 16 gigabyte (GB) RDIMM memory. It operates on the Debian 8.9 Linux, and the HDD is a 48-terabyte (TB) SAS (7.4 k RPM).

The ZDRS security processing unit is a software module, but if it uses a 10 Gbps interface, it has to use an external hardware network processing unit; otherwise, if it uses a 1 Gbps interface, it can be implemented on the host system. In this implementation, we implement the ZDRS using an external hardware network processing unit to use a 10 Gbps interface.

Table 3 lists the hardware specifications of the ZDRS network processing unit. The network processor was a Tile-G36 processor.

Memory is a DDR3 type and is 8 GB. It has an SFP+4-port network interface. It uses 8-lane of the PCIe Gen2. It operates on CentOS or Red Hat. These hardware specifications make it transfer packets of 20 Gbps to the host. These hardware specifications allow packets of 20 Gbps to be transmitted to the host.

4.2. Implementation Result. We collected the raw 3.4 TB PCAP data during 9 hours and confirmed the storage and detection rates of the collected data in the ZDRS. And then, through replaying based on this full packet, we resaved the data as each first-N storage method. Using this sample data, the implementation results were compared with the results of the proposed first-N storage method and full-size storage method. In this implementation, the application classification method based on the port number was applied in all implementations to compare the results of the proposed first-N storage scheme with those of the full-size storage scheme.

In general, research on detection algorithms for zero-day attack suggests correct detection rate and false positive rate as experimental results. However, this paper does not propose a detection algorithm, but the ZDRS that solves the problems of the forensic storage method through the first-N packet storage method. Therefore, we used Suricata open rule set. That is, the purpose of this experiment is to confirm the detection rate of inspection for packets collected by the first-N method compared to the full-packet method. Therefore, true rate and false rate are not considered in this study.

Table 4 shows the stored packets, storage rate, and the number of detections of the ZDRS test. While the full-packet storage method stored and analyzed approximately 3.4 TB of data, the first 5-kilobyte (kB) size storage method stored and analyzed approximately 62 GB of data. The first 5 kB storage used approximately 1.8% storage space versus the full-packet storage method. Also, the number of detections increases as approaches the full packet.

Figure 9 shows the number of detections for each first-N storage method and the full-size storage method. The full-size storage scheme detected 13,944 cases, while the first 5 kB and 1 MB storage schemes detected 13,500 and 13,555 cases, respectively.

Figure 10 shows the detection rate of the first-N storage method and the full-packet storage method. We calculate the detection rate as follows using the number of detections in Table 4.

$$\text{Detection rate} = \frac{N_{df}}{N_{dn}} * 100. \quad (1)$$

N_{df} is the number of detections in full-packet storage method. N_{dn} is the number of detections in first-N storage methods. When the detection rate of the full-packet storage method was taken as 100%, the detection rates of the first 5 kB and 1 MB storage methods were 99.55% and 99.69%, respectively. The first-N storage methods are more than 99.55%, and there is little difference compared to the full-packet storage method. However, the difference between

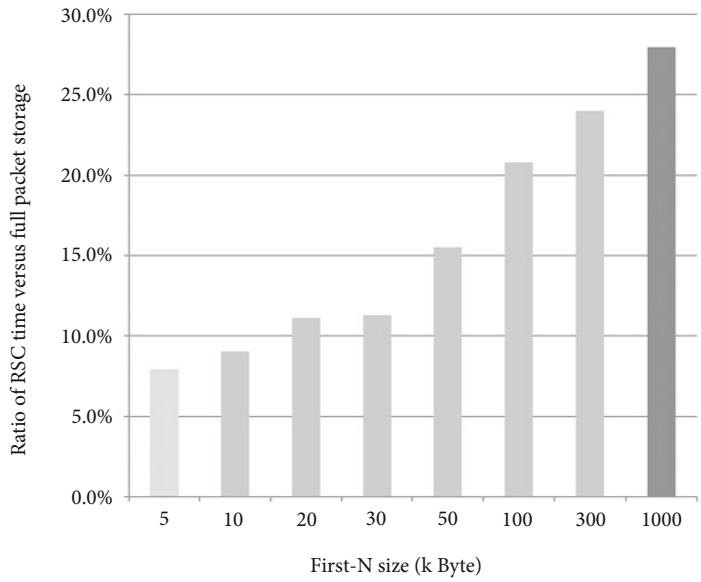


Figure 12: Ratio of RSC time for first-N storage methods versus full-packet storage method.

the first 5 kB size storage method and the first 1 MB size storage method was only 55 cases.

Figure 11 shows storage efficiency for the first-N size storage method versus the full-packet storage method. The result of the first 5 kB size showed that the data storage rate decreased approximately 55-fold compared with the full-packet method. Moreover, the result of the first 1 MB size showed that the data storage rate decreased approximately fivefold compared with the full-packet method.

Figure 12 shows the ratio of RSC time for first-N storage methods versus the full-packet storage method. The first 5 kB size storage method spent 7.9% RSC time versus the full-packet storage method. Moreover, the first 1 MB size storage method spent 28% RSC time versus the full-packet storage method. In other words, as the size of N decreases, the RSC time decreases.

To evaluate the performance of ZDRS, it is necessary to look at the all factors which are number of detections, detection rate, storage efficiency, and RSC time. For example, the first 5 kB storage method has 444 fewer detections than the full-packet method. However, the detection rate is only 0.45%, and the storage efficiency is 55 times higher. Moreover, the efficiency is only 7.59% of the time required to perform RSC in full-packet method. In other words, depending on the application target, it is better to take a small amount of first-N size in order to have the maximum storage efficiency, and it is better to take a large first-N size for a strict security check.

5. Conclusion

In this study, we designed and demonstrated a data storage and regression security check system for responding to network security blind spots. In detail, we propose a system structure for regression security inspection, a metadata generation algorithm, and a metadata search algorithm to provide a solution for operational blind spots and temporal blind spots. In addition, we used the first-N packet storage method to drastically reduce the storage cost. As a performance analysis result of implementation, the amount of data storage decreased from 3.4 TB to 62 GB compared to the full-packet storage method by 1.82%, and storage efficiency increased by 54.84 times. The detection rate of 99.55% based on the first 5 kB size compared to the full-packet storage method was confirmed. That is, depending on the application target, it is better to take a small amount of first-N size in order to have the maximum storage efficiency, and it is better to take a large first-N size for a strict security check. In conclusion, it was confirmed that even if the packet is saved with ZDRS's first-N technology, it can be sufficiently utilized for security check.

In this paper, we used the session-based first-N storage method. In a future study, we plan to apply the flow-based first-N storage method and compare the storage efficiency and accuracy with the session-based first-N storage method. This study used the session-based first-N method. However, the flow-based first-N may have different inspection volume and detection rate than the session-based first-N and may have different characteristics depending on the type of attack. Therefore, we plan to study the flow-based first-N storage method. In addition, we will collect and test more samples at various sites—research institutes and companies.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Authors' Contributions

Won-Seok Choi contributed to the conception of the main idea. Won-Seok Choi and Si-Young Lee contributed to the writing of manuscript and preparation of all figures. Won-Seok Choi and Si-Young Lee contributed to the implementation and test. Won-Seok Choi and Seong-Gon Choi contributed to the analysis of the results. Seong-Gon Choi contributed to the supervision and conceptualization.

Acknowledgments

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2020R1A6A1A12047945).

References

[1] V. Hassija, V. Chamola, V. Saxena, D. Jain, P. Goyal, and B. Sikdar, "A survey on IoT security: application areas, security threats, and solution architectures," *IEEE Access*, vol. 7, pp. 82721–82743, 2019.

[2] H. Tabrizchi and M. K. Rafsanjani, "A survey on security challenges in cloud computing: issues, threats, and solutions," *The Journal of Supercomputing*, vol. 76, no. 12, pp. 9493–9532, 2020.

[3] H. Tabrizchi and M. K. Rafsanjani, "Digital twin networks: a survey," *IEEE Internet Of Things Journal*, vol. 8, no. 18, pp. 13789–13804, 2021.

[4] K. B. Kelarestaghi, M. Foruhandeh, K. Heaslip, and R. Gerdes, "Intelligent transportation system security: impact-oriented risk assessment of in-vehicle networks," *IEEE Intelligent Transportation Systems Magazine*, vol. 13, no. 2, pp. 91–104, 2019.

[5] N. Sameera and M. Shashi, "Deep transductive transfer learning framework for zero-day attack detection," *ICT Express*, vol. 6, no. 4, pp. 361–367, 2020.

[6] A. Blaise, M. Bouet, V. Conan, and S. Secci, "Detection of zero-day attacks: an unsupervised port-based approach," *Computer Networks*, vol. 180, article 107391, 2020.

[7] T. Zoppi, A. Ceccarelli, and A. Bondavalli, "Unsupervised algorithms to detect zero-day attacks: strategy and application," *IEEE Access*, vol. 9, pp. 90603–90615, 2021.

[8] A. Ghosal and M. Conti, "Security issues and challenges in v2x: a survey," *Computer Networks*, vol. 169, article 107093, 2020.

[9] R. Singh, H. Kumar, R. K. Singla, and R. R. Ketti, "Internet attacks and intrusion detection system: a review of the literature," *Online Information Review*, vol. 41, no. 2, pp. 171–184, 2017.

[10] X. Sun, J. Dai, P. Liu, A. Singhal, and J. Yen, "Using Bayesian networks for probabilistic identification of zero-day attack paths," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 10, pp. 2506–2521, 2018.

[11] H. Al-Rushdan, M. Shurman, S. H. Alnabelsi, and Q. Althebyan, "Zero-day attack detection and prevention in software-defined networks," in *2019 International Arab Conference on Information Technology (ACIT)*, pp. 278–282, Al Ain, UAE, 2019.

[12] H. Hindy, R. Atkinson, C. Tachtatzis, J.-N. Colin, E. Bayne, and X. Bellekens, "Utilising deep learning techniques for effective zero-day attack detection," *Electronics*, vol. 9, no. 10, p. 1684, 2020.

[13] G. P. Bherde and M. Pund, "Technique for detecting zero day attack by using signature based and knowledge based method, International Journal of Scientific Research in Science," *Engineering and Technology*, vol. 4, pp. 649–652, 2018.

[14] C. Patidar and H. Khandelwal, "Zero day attack detection using machine learning techniques," *IJRAR*, vol. 6, no. 1, pp. 1364–1367, 2018.

[15] Y. Mirsky, T. Doitshman, Y. Elovici, A. Shabtai, and A. Kitsune, "An ensemble of autoencoders for online network intrusion detection," 2018, https://arxiv.org/abs/1802.09089.

[16] G. Bovenzi, G. Aceto, D. Ciuonzo, V. Persico, and A. Pescape, "A hierarchical hybrid intrusion detection approach in IoT scenarios, in proceedings of the GLOBECOM 2020, Taipei," in *GLOBECOM 2020 - 2020 IEEE Global Communications Conference*, pp. 1–7, Taipei, Taiwan, 2020.

[17] N. Innab, E. Alomairy, and L. Alsheddi, "Hybrid system between anomaly based detection system and honeypot to detect zero day attack," in *2018 21st Saudi Computer Society National Computer Conference (NCC)*, pp. 1–5, Riyadh, Saudi Arabia, 2018.

[18] P. Duessel, C. Gehl, U. Flegel, S. Dietrich, and M. Meier, "Detecting zero-day attacks using context-aware anomaly detection at the application-layer," *International Journal of Information Security*, vol. 16, no. 5, pp. 475–490, 2017.

[19] R. C. Aygun and A. G. Yavuz, "Network anomaly detection with stochastically improved autoencoder based models," in *2017 IEEE 4th International Conference on Cyber Security and Cloud Computing (CSCloud)*, pp. 193–198, New York, NY, USA, 2017.

[20] H. Bostani and M. Sheikhan, "Hybrid of anomaly-based and specification-based ids for Internet of Things using unsupervised OPF based on MapReduce approach," *Computer Communications*, vol. 98, pp. 52–71, 2017.

[21] J. Y. Siu and S. K. Panda, "A specification-based detection for attacks in the multi-area system," in *IECON 2020 The 46th Annual Conference of the IEEE Industrial Electronics Society*, pp. 1526–1526, Singapore, 2020.

[22] A. Althubaity, H. Ji, T. Gong, M. Nixon, R. Ammar, and S. Han, "Arm: A hybrid specification-based intrusion detection system for rank attacks in 6TiSCH networks," in *2017 22nd IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*, pp. 1–8, Limassol, Cyprus, 2017.

[23] I. Altaf and I. Majeed, "A specification-based intrusion detection model for OLSR, international journal of computer science," *Engineering Technology*, vol. 8, pp. 273–282, 2017.

[24] M. Ali, S. Shiaeles, N. Clarke, and D. Kontogeorgis, "A proactive malicious software identification approach for digital forensic examiners," *Journal of Information Security and Applications*, vol. 47, pp. 139–155, 2019.

[25] K. Demertzis, P. Kikiras, N. Tziritas, S. L. Sanchez, and L. Iliadis, "The next generation cognitive security operations center: network flow forensics using cybersecurity intelligence," *Big Data and Cognitive Computing*, vol. 2, no. 4, p. 35, 2018.

[26] P. Prasad, N. Sowmya, K. Reddy, and P. Bala, "Introduction to dynamic malware analysis for cyber intelligence and forensics," *International Journal of Mechanical Engineering and Technology*, vol. 9, no. 1, pp. 10–21, 2018.

[27] A. Nisioti, G. Loukas, A. Laszka, and E. Panaousis, "Data-driven decision support for optimizing cyber forensic investigations," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 2397–2412, 2021.

[28] S. M. Ryu, S. Y. Lee, C. W. Ryu, and S. G. Choi, "A method of the first-n packets storage by using the advanced-PCA, Contemporary," *Engineering Sciences*, vol. 9, no. 14, pp. 647–653, 2016.

[29] J.-S. Park, S.-H. Yoon, and M.-S. Kim, "Performance improvement of signature-based traffic classification system by optimizing the search space," *Journal of Internet Computing and Services*, vol. 12, no. 3, pp. 89–99, 2011.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 5374389, 11 pages
https://doi.org/10.1155/2022/5374389

Research Article

The Interaction of Internet Fun Design Resources Based on the Network Communication System

Xiaojie Min

Fashion and art Design Department, Gongqing College, Nanchang University, Gongqing, 332020 Jiangxi, China

Correspondence should be addressed to Xiaojie Min; i22021959@student.newinti.edu.my

Received 27 January 2022; Revised 8 March 2022; Accepted 15 March 2022; Published 7 April 2022

Academic Editor: Hoon Ko

The interactive function of the Internet is the basis of all data transmission. The functions of Internet products are becoming more and more complex, and the amount of information they carry also increases dramatically. People and products mainly interact with products through the product interface, and the operation of the interface becomes more and more complex and more important. With the help of the interactive function of the Internet, various "Internet +" businesses are carried out in an orderly manner. The realization of Internet functions also requires user-centric interactive design. This puts higher demands on the fun of Internet interaction. This article aims to study the interaction of interesting design resources on the Internet, so as to make the Internet interaction more interesting. This article proposes to add interesting elements to the interactive functions of the Internet on the basis of the network communication system and stimulate the user's feelings from many aspects such as actions, pictures, sounds, and colors, so as to make the users feel psychologically happy and trigger. The experimental results in this paper show that fun interaction design can enhance user experience by 20%.

1. Introduction

The network communication system is the basis for voice, data, and image transmission and is connected to external communication networks (such as public telephone networks, integrated service digital networks, computer Internet, data communication networks, and satellite communication networks) to ensure smooth information flow. With the development of the technological age, our production and life can no longer be completely separated from the Internet. More and more network products are overflowing our lives. Especially in the current era of focusing on user experience, facing a fast-paced and stressful living environment, modern users pay more attention to spiritual relaxation, and the design that reaches the depths of the soul will affect the user's own emotions and behaviors and enhance the interest and pleasure of the user experience. A very important evaluation criterion for interaction design is the product function and ease of use. For users, in the process of using the product, if the function of the product exceeds the user's expectations, this will become an attractive point to attract users to interact with the product in depth, reducing the dullness of system operation. As a perceptual animal, the center of people's attention is the emotional communication, and the most moving thing is also the emotional communication. In the process of product marketing, the survey found that consumers' emotional needs have increasingly become a marketing focus. Consumers enjoying basic product services when buying products are just the most basic marketing strategy. Emotional interaction is the key to moving consumers and stimulating consumers to buy. This is not only the strategy of product marketing today, but in the future, product marketing is also inseparable from the exploration of user needs. The interest is the prerequisite for the realization of user emotional interaction. Studying consumer interest can better find the true needs of consumers and create more product value. In the future, the fun of interactive design will surely become the direction of Internet product design.

For the research of interaction design, domestic and foreign experts and scholars have already achieved results. Woodbury studies this near-universal feature of changing parameterized interfaces to support the exploration of using multiple alternatives. It builds a prototype gallery system on a web browser that supports saving alternatives from three

graph-based parametric modeling tools. The users can retrieve alternatives from the gallery, share them with others, and combine them to generate more alternatives. This research proposes several directions for new user interface design [1]. Sequenzia developed a variable geometry saddle (VGS) prototype for amateur cyclists through an interactive redesign method, which can reduce saddle pathology and improve pedaling comfort. The development of VGS can be adjusted according to the rider's anatomical requirements and various riding conditions (uphill, flat, and downhill). Simple adjusters will affect the inclination of the nose and the width of the saddle backrest. In particular, the nose mechanism allows immediate adjustment. The VGS developed can also allow cyclists to identify the most suitable subjective geometric shapes to help choose among commercial alternatives [2]. Rizzuti proposes a unique noise reduction function that can be used in combination with all types of "loss functions." This article discusses an interactive program that can integrate Taguchi's method and axiomatic method, which can check the design matrix, the nature of which can hint at the effectiveness of the product being developed and find contradictions [3]. Abbood proposed a robot vision system that can distinguish the color of objects and their position coordinates and then classify the objects (products) on the right branch conveyor belt in real time based on the colors. The system is constructed based on the HVS mode algorithm and classifies products based on color. In addition, the system can also recognize the shape of the object, then find its position, pick up the shape of the object, and place it on the correct branch conveyor [4]. Kim proposed a new robust optimization design method, which realized the optimization procedure based on direct search. The Interactive Design Space Reduction Method (IDSRM) based on orthogonal arrays for experimental design was developed as a general optimization tool. Using this system, the designers can interactively adjust the design space according to experimental results during the search process to obtain the best solution. Compared with the initial design solution, the proposed method-assisted optimization design shows more effective and better results in terms of design robustness [5]. Preciado proposes to use the finite element method (FEM) to include the volume model in the interactive simulation of deformable objects, directly using the triangular mesh method. This method avoids the problems caused by the generation of tetrahedral meshes, such as unstable, poor-quality meshes or degraded meshes. In this research, he improved the grid intersection method by increasing the robustness to simulate complex grids and also showed how to realize user interaction in real time [6]. Amory's research involves the development and evaluation of a redesigned online and mobile application for the African Storybook Initiative service, which supports the creation and reading of publicly licensed storybooks to support the development of literacy in Africa. It provides an opportunity to combine the sociocultural theories used by African storybook researchers with the design of interactive artifacts and environments. The redesign utilizes many cultural-historical activity theory principles, including activity objects, tool intermediaries, and shared objects as part of the third-generation activity system. Three main activities (reading, author and research) have been identified [7]. Jim has developed an Art Game, which provides a very simple user experience. Most of the content of the game is to answer yes or no to each painting. But attribution is an extraordinary problem expert, and connoisseurs are related to it when verifying their findings. Like them, the art game players bring their knowledge and instincts (regardless of what they have recently acquired) to a problem. Jim reconstructed the science of appreciation; he focused on a kind of dual factor copyright to help users improve their visual thinking and recognition ability [8].

This article studies the integration of fun into the Internet interaction design and, on the basis of network communication technology, incorporates interesting elements, investigates the user experience, finds the best user experience, and optimizes the design of the plan to better increase the user's Internet experience. The user-centered design concept requires that in the process of designing the interactive interface of the product, the needs of the product users should be fully considered, and the interactive interface of the product should be designed according to the needs of the users. In order for the product to obtain a good user experience, it is necessary to conduct in-depth research on the product's interactive interface and evaluate the design results.

2. Interesting Introduction to Network Communication System and Interaction Design

2.1. Network Communication System. The communication system consists of two parts: hardware and software. The hardware consists of processor, clock and power circuit, reset configuration circuit, memory circuit, network interface circuit, I/O interface circuit-bus, and circuit test interface. The processor is an onboard processor; the clock and power supply include the system main clock, USB clock, and power control circuit; the memory part includes Flash for storing the operating system and SDRAM for program operation; the network interface includes the Ethernet chip and configuration circuit; the I/O bus interface is composed of a 16-bit double word data line, read signal, write signal, address line, reset, wait, chip, and power signal; the debug and test interface includes JTAG and UART interface. The software part is composed of integrated operating system and application software [9]. The hardware composition block diagram of the communication system is shown as in Figure 1.

SINR is the quality of the signal, which is equal to the ratio of the signal strength C to the sum of the noise N and the interference strength I. If the users want to communicate data through the current network, they must obtain corresponding channel resources. When the channel C_{m} is allocated to A (x) and B user pair y, the calculation formula of the instantaneous SINR of A (x) is

$$S_{x,y}^{m} = \frac{F_x^a Z_{x,D}^m}{F_y^b Z_{y,D}^m + \theta^2} \tag{1}$$

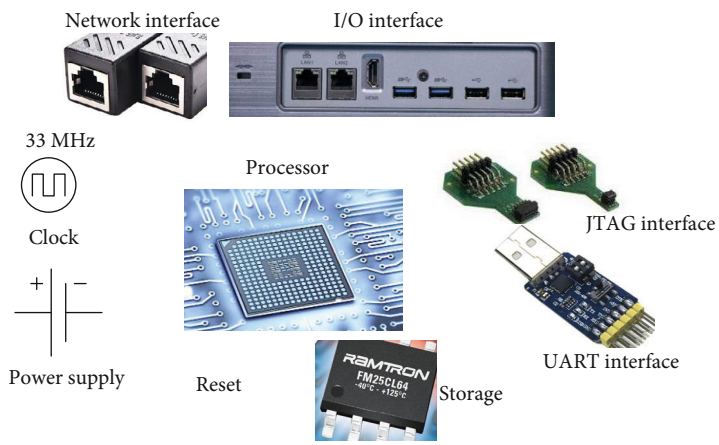


Figure 1: Frame diagram of communication system.

In the same way, the instantaneous SINR calculation formula of user B to the receiving end of y is as follows:

$$S_m^{y,x} = \frac{F_y^b Z_m^{y,y}}{F_x^a Z_m^{x,y} + \theta^2} \tag{2}$$

In the formula, the operating power of the noise can be expressed as θ^2; the transmission power of the signal can be expressed by F_x^a and F_y^b, respectively, the output of A (x) and B users to y; and the increase in channel power between A (x) and D can be expressed by $Z_{x,D}^m$. It means that the utility of signal interference between different links of y and D can be expressed by $Z_{y,D}^m$, the increase in channel power of B users to y can be expressed by $Z_m^{y,y}$, and the utility of signal interference between different links of B and y can be expressed by $Z_m^{x,y}$.

In order to achieve the maximum overall rate of the communication link, it can be calculated by solving the mathematical formula:

$$\max \sum_{m=1}^{M} \sum_{x=1}^{E+G} \sum_{y=1}^{F+H} \left\{ \gamma_{x,m} J \log_2 \left(1 + S_{x,y}^m\right) + \lambda_{y,m} J \log_2 (1 + S_m^{y,x}) \right\} \tag{3}$$

$$s.t. S_{x,y}^m \geq S_{x,\min}^m, \forall x \in P \tag{4}$$

$$S_m^{y,x} \geq S_{m,\min}^{y,x}, \forall y \in Q \tag{5}$$

$$\sum_{m=1}^{M} \gamma_{x,m} \leq 1, \sum_{m=1}^{M} \lambda_{y,m} \leq 1, \forall x \in P, \forall y \in Q \tag{6}$$

$$\sum_{x=1}^{E+G} \gamma_{x,m} \leq 1, \sum_{y=1}^{F+H} \lambda_{y,m} \leq 1, \forall x \in P, \forall y \in Q \tag{7}$$

The bandwidth is represented by J, and the minimum SINR thresholds of A (x) and y can be represented as $S_{x,\min}^m$ and $S_{m,\min}^{y,x}$. $\gamma_{x,m}$ and $\lambda_{x,m}$ are the output of resource allocation results, which are piecewise functions, which can be expressed as

$$\gamma_{x,m} = \begin{cases} 1, C_m \text{ is assigned to A(x)} \\ 0, \text{Other cases} \end{cases} \tag{8}$$

$$\gamma_{x,m} = \begin{cases} 1, C_m \text{ is assigned to user B to y} \\ 0, \text{Other cases} \end{cases} \tag{9}$$

The establishment of the above conditions meets the minimum SINR requirement, which reduces the error caused by signal interference to a certain extent [10].

We use t to represent users who are not connected to the network. With the aid of the signal-to-noise ratio, the users will uniformly evaluate the channel of the cell. The output signal-to-noise ratio of the link can be represented by $\omega_{t,m}$, and its calculation formula is

$$\omega_{t,m} = \begin{cases} \dfrac{F_t Z_{t,D}^m}{\theta^2}, \alpha_{t,m} = 1 \\ 0, \alpha_{t,m} = 0 \end{cases} \tag{10}$$

Among them, the transmitted power F_t can be expressed as t, and the power increase of the channel between D can be expressed as $Z_{t,D}^m$. Binary is a number system widely used in computing technology. The binary data is a number represented by two digits, 0 and 1. The characteristic value of the user can be represented by a binary indicator function $\alpha_{t,m}$. According to the above conditions, whether$\alpha_{t,m}$ can be used for data communication by users can be preliminarily determined. $\alpha_{t,m} = 1$, it means the result is definite, otherwise $\alpha_{t,m} = 0$ [11]. The specific situation is shown in Table 1.

Table 1: $\alpha_{t,m}$ Assignment in different situations.

	A (x)	B user
Already reused by users A (x) and B	0	0
Only occupied by A (x)	0	1
Only occupied by B	1	0
Unoccupied	1	1

Here, the variable Q_t is used to rank the preference of user t for each channel, which is called the feature value list of user t. The specific expression is

$$Q_t = \left\{ \left(\alpha_{t,m_1}, m_1, \varepsilon_{t,m_1}\right), \left(\alpha_{t,m_2}, m_2, \varepsilon_{t,m_2}\right), \cdots, \left(\alpha_{t,m_3}, m_3, \varepsilon_{t,m_m}\right) \right\} \tag{11}$$

Among them, $\varepsilon_{t,m_1} \geq \varepsilon_{t,m_2} \geq \cdots \geq \varepsilon_{t,m_m}$; m refers to the index value. Suppose that the first channel index value in Q_t is the channel with the greatest user interest.

According to the characteristic values of users, a user alliance is formed on the channel with the highest similarity, and L_m refers to the alliance of channel m.

Let I_m be the initial rate of the alliance, which is provided by users with stable communication, which can be expressed as

$$V_0^m = \sum_{x=1}^{N} \gamma_{x,m} D \log_2\left(1 + S_{x,y}^m\right) + \sum_{y=1}^{M} \lambda_{y,m} D \log_2\left(1 + S_m^{y,x}\right) \tag{12}$$

$$\forall x \in G_m, \forall y \in H_m \tag{13}$$

Alliance V_m is expressed as the final rate of the alliance. It can be represented by a weighted bipartite graph. Its two sets of unrelated endpoints, and the A (x) and B user distributions represent two sets of unrelated endpoints. Establish an alliance of users waiting to access the network. The channel may have four different states (C_m has been multiplexed by two users, C_m is only occupied by A (x), C_m is only occupied by B users, and C_m is not occupied) [12], as shown in Figures 2 and 3, which can be represented as four different forms of it.

As shown in Figure 2(a), channel C_m is multiplexed by two groups of users at the beginning, so its final rate is determined, which can be expressed as

$$V_{\text{sum}}^{\text{m}} = \sum_{x=1}^{N} \gamma_{x,m} D \log_2\left(1 + S_{x,y}^m\right) + \sum_{y=1}^{M} \lambda_{y,m} D \log_2\left(1 + S_m^{y,x}\right) \tag{14}$$

As shown in Figure 2(b), C_m a group of data is occupied by A (x), and another group of data is to be allocated by user B. At this time, the final rate is determined by two sets of weights, which can be expressed as

$$V_{\text{sum}}^{\text{m}} = D \log_2\left(1 + S_{x,y'}^m\right) + \sum_{y=1}^{F+H} \lambda_{y',m} D \log_2\left(1 + S_m^{y',x}\right) \tag{15}$$

As shown in Figure 3(a), C_m one group of data is occupied by user B, and the other data is to be allocated by A (x). At this time, the final rate can be expressed as follows:

$$V_{\text{sum}}^{\text{m}} = \sum_{x=1}^{E+G} \gamma_{x',m} D \log_2\left(1 + S_{x',y}^m\right) + D \log_2\left(1 + S_m^{y,x}\right) \tag{16}$$

As shown in Figure 3(b), when C_m is not occupied by any user at the beginning, the new user provides its final rate. The calculation formula is as follows:

$$V_{\text{sum}}^{\text{m}} = \sum_{x=1}^{E+G} \gamma_{x',m} D \log_2\left(1 + S_{x',y'}^m\right) + \sum_{y=1}^{F+H} \lambda_{y,m} D \log_2\left(1 + S_m^{y',x'}\right) \tag{17}$$

In Formulas (13)–(16),

$$\forall x \in G_m, \forall y \in H_m \tag{18}$$

$$\forall x' \in G_p, \forall y' \in H_q \tag{19}$$

Constraints (4) and (5) are the prerequisites for all SINR calculations. The calculation of the increase in the rate of users waiting to access the network can be expressed by the following formula:

$$V_{x'',y''}^m = V_{sum}^m - V_0^m \tag{20}$$

When calculating indicators such as averages in statistics, the values that have a weighted effect on each variable value are called weights. When calculating, if a number has the same number, multiply the number by the number, and this number is called the right. Among them, $\forall x'' \in G$, $\forall y'' \in H$. Suppose the weight value of the connection between $A(x'')$ and y'' is denoted by $V_{x'',y''}^m$. The users with the maximum rate increase of C_m data communication are grouped into a set, denoted as $A(x^*)$. Delete the first element of the C_m user characteristics that has not been obtained so that it can connect to the new alliance to transmit data. The appeal operation is repeated until all users meet the requirements or $Q_t = \varphi$.

2.2. Interaction Design. Interaction can understand the process of a certain experience produced by the interaction between people and things and between people. In this process, there must be conscious and feeling exchanges and changes [13]. The interaction design establishes effective communication between designing products and users. In this process, it can help users use products quickly and achieve the goal of a pleasant experience. Therefore, whether

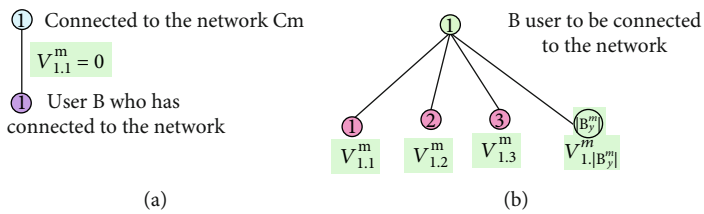


Figure 2: The alliance channel is multiplexed by two users and occupied only by A (x).

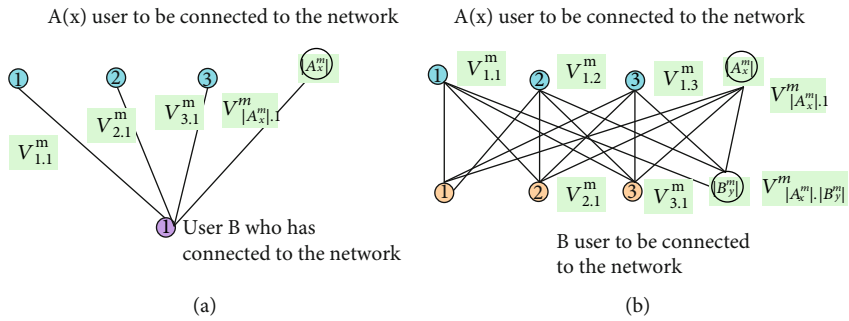


Figure 3: Schematic diagram of manifestations of only occupied and unoccupied by user B.

interactive design is classified into behavioral design or product design, it is based on user experience. It requires designers to understand the characteristics of users' psychological needs, conduct analysis and research based on users' daily operating habits, and combine the design concepts of interactive design. Produce products that meet user requirements. Because interaction design involves many disciplines, such as design psychology, semiotics, human function, and consumer behavior [14, 15]. So excellent interaction design not only enables users to operate the product conveniently and quickly, but also uses various effective interaction methods to make users have a pleasant interactive experience. With the development of network technology, more and more merchants choose online interactive methods for product promotion, forming a new form of human-computer interaction in marketing, and enhancing people's attention to interactive experience.

2.3. Internet Media Attributes and Content Strategy. When the Internet was just emerging, it was used for hypertext production, allowing people to create and share documents with each other and then became a way for research teams to communicate in a small area. When this technology expands the scope of dissemination and is widely used for resource sharing and material reference, the Internet has shown its excellent media function. By adding advanced technologies and new features to browsers and servers, the Internet has gradually evolved from static information collection to database-driven dynamic sites [16].

The Internet can not only convey information, but also collect and control this information. At the same time, it supports real-time response with users and provides multiple possibilities for interaction with audiences. And it has therefore become a hybrid with the dual functions of media (traditional characteristics) and instant response [17]. On the one hand, we can take the website as the carrier of information release and retrieval, and think about the function and meaning of the website; on the other hand, we can regard the website itself as a "localized" software product, pay attention to its use process, and think about how users achieve the set goals. The two perspectives complement each other and interact to form a complete Internet media. The content of the Internet is an important element of its role as a media. The content refers to what users read, learn, and see on the site. It is generated by a series of complex information linked through the network. It is a comprehensive experience and is by no means a certain part of the specific content understood in a narrow sense. The content has the functions of persuasion, notification, verification, instructions, and entertainment, which is the core strength of the website to attract and retain users. Therefore, we need a good content strategy to provide guidelines for the use of website content, so as to meet website goals and user needs and at the same time lead the content life cycle (from creation to death) and provide a basis and criteria for iterative design and development. The "quadruple diagram" constructs the content strategy of the website, as shown in Figure 4; it can determine the ideas and methods of using the website content to achieve its own goals and meet the needs of users [18].

(1) The Core content strategy sets a long-term development trend for website content. The core strategy is based on product demand decision-making based on ease-of-use standards. In an agile user experience project, the core content strategy of Internet media is based on product requirements, and it is designed to be changed flexibly with each iteration in the product life cycle

(2) The substance part needs to clarify what type of content the website needs and what information they

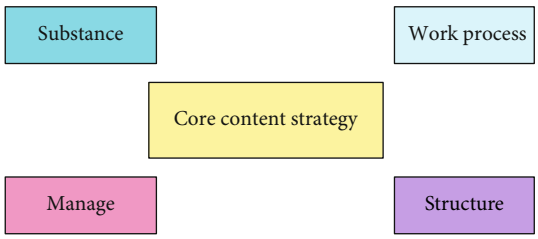


Figure 4: Four elements of the Internet media content strategy.

need to convey in order to achieve communication with users

(3) The structure refers to the form, priority, organization, and display method of content arrangement (including information architecture, hyperlinks, and metadata).

(4) The workflow refers to the work process, methods, and personnel arrangements of the entire life cycle of the content (release, maintenance, expansion, update, cancellation, etc.).

(5) The management (governance) covers content and the rules of content strategy continuation and change

2.4. Interesting. What is fun is understood from the literal meaning of the psychological effect that makes people feel pleasant and arouses interest. It can be explained in two levels. Affection refers to sentiment and affection, and fun refers to interest and pleasure [19]. Human emotions include emotions, anger, sorrow, and joy. Interesting design will make people feel interesting, will move people's heart and attract people's attention, and can make people happy and beautiful aftertaste the moment. It is difficult for everyone to resist interesting things. A fun life is rich and wonderful. People oriented is the basic principle in modern product design, because the relationship with people will produce emotional connections. We can find that many products in life use humor and humor to convey emotions, which allows users to relax and enjoy. The interesting design of the product is the embodiment of people-oriented.

The interesting design can be explained from two aspects. One aspect is that the designer incorporates relevant interesting elements at the beginning of the product design to make the product interesting. On the other hand, the consumer design chooses products according to their own aesthetic taste. In fact, there is no essential difference between the two aspects, and both are to satisfy the user's aesthetic taste as the core. Fun design itself is the aesthetic taste of the public. Only by designing products that meet the emotional needs of the public can the emotional interaction between the product and the user be harmonious and consistent [20].

2.4.1. From the perspective of the product itself, the high user satisfaction is very important for the product. According to the results of the American Customer Satisfaction Index (ACSI) model, user satisfaction is the degree to which users' desires are satisfied. If the product and service far exceed user expectations, then the user satisfaction will be greatly improved, and vice versa. Generally speaking, each satisfaction score is multiplied by the corresponding weight to obtain a weighted value, and all the weighted values are multiplied together to obtain the total average value.

For interactive products, the product provides services to match user needs. In this process of enjoying the service, the users will classify their experience levels according to their own experience. Whether you can achieve your set goals is the basis for measuring satisfaction, and user experience during use is an important condition for measuring satisfaction. After completing the set tasks, whether there are other gains is the key to improving satisfaction, as shown in Figure 5.

In the design of interactive products, the user experience is a bottom-up mapping, which puts forward bottom-up requirements for the design of interactive products. Whether the functional structure of interactive products is efficient is the prerequisite for product design. Evaluating the rationality of product structure is the basis of product design. The product can achieve the target demand that is the direction of product manufacturing. The product meets the psychological needs of users and is original and interesting. The details are shown in Figure 6.

As shown in Figure 7, the use of a product is an indispensable condition for product design. However, the usability of the product is only the basis. For users, the level of satisfaction lies in whether it can provide higher demand. User experience and emotional needs are the breakthrough points to improve user satisfaction. Interest, as a basic element to meet human emotional needs, is essential for enhancing product originality and emotional needs. It meets the deep-seated needs of users, makes the product usable and easy to use, and further enhances user satisfaction. According to the relevant knowledge content of human psychology and communication, people tend to tend to things that bring them pleasure. Interestingness is reflected in the design, which can bring pleasant emotions to the viewers.

2.4.2. Better optimize software functions. The fusion of interesting elements and software operations can have an amplification effect. The users feel the interesting elements in the process of using the software. While achieving their goals, satisfying emotional needs is a direction of fun design.

Generally speaking, we can see the pattern of the alarm clock and the specific time display. As time passes to the set time, the alarm sounds, which is in line with the user's habit. However, if you increase the sound of running water or the ticking sound of the clock when setting the alarm, the user will be visually and audibly moved and feel the passage of time. The whole process is lively and interesting. While playing the function of the alarm clock, it makes the function of the alarm clock more three-dimensional and meets the needs of users. For example, the effect of the dynamic logo of the video website iQIYI is to use the point element on the original logo to make it dynamic and jump to its position one by one join in.

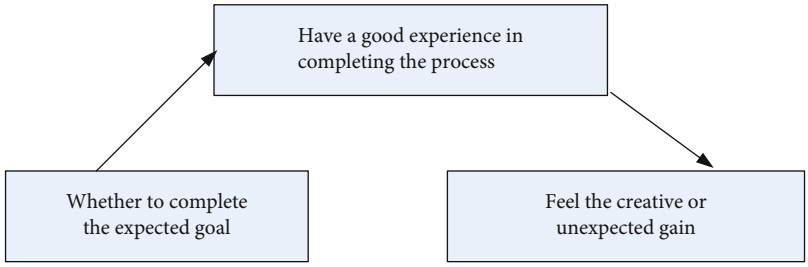


Figure 5: Interactive product experience structure from high to low.

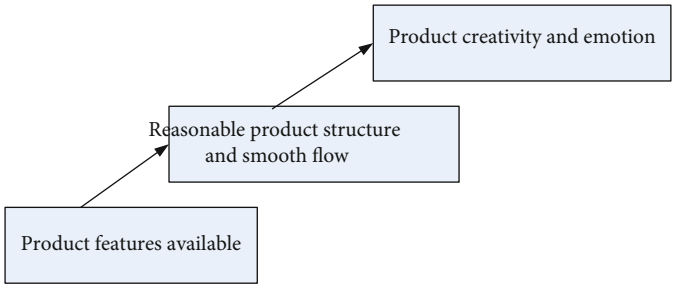


Figure 6: The experience structure corresponds to the design requirements of interactive products from high to low.

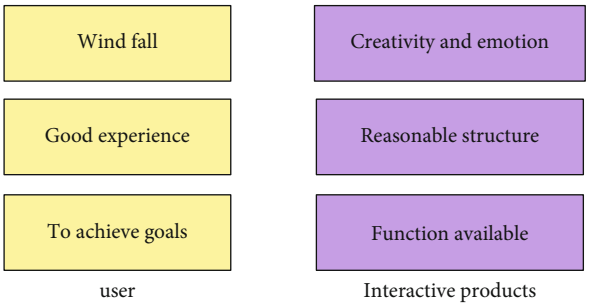


Figure 7: The relationship between user experience and interactive product design requirements.

Table 2: Introduction to the age structure of the questionnaire.

	18-25	25-35	35-45
Internet age	5-8	9-15	9-15
Total	36	48	21
Male	17	20	8
Female	19	28	13

Table 3: Satisfaction evaluation standards for interaction-related elements.

	Image	Word	Action	Audio	Video
Total	70%	76%	71%	65%	58%
Familiarity	73%	85%	67%	56%	39%
Satisfaction	68%	72%	75%	73%	77%

2.4.3. Guide users to use the software efficiently. When it comes to unlocking software, I believe that many people are familiar with it, especially when almost everyone has a mobile phone, it is widely used by users. Sliding patterns, rich pictures, and interesting action designs make unlocking operations not only easier, but also more interesting. Simply slide your finger to enter the system interface smoothly. Although the change is only a small gesture, it is also a sublimation of the user's concept, taking the user as the center to participate in product design at a deeper level, so as to maximize the function of the software.

2.4.4. Optimize software based on human psychology. In real life, it is a natural habit of people to open the closet and search for clothes one by one. This habit is translated into computer operation, and along with the movement of the mouse, the clothes move in sequence for the user to choose. This design conforms to the physiological habits of the human body and also breaks the boundary between reality and the network, allowing the network to further touch reality. Throughout the process, the products are continuously available and fully displayed, giving people an immersive shopping experience. The details are even more surprising. When users interact with the product, their psychological and material needs are met in both directions, making software optimization more reasonable and more valuable.

2.4.5. Bring emotional adjustment and comfort. At present, smart phones are used as information communication channels to connect various resources. With the help of smart phones, people can communicate with the world, watch film, television, and audio programs and enjoy better spiritual civilization needs. More and more people spend their time on the Internet and on mobile phones. Smartphones also rely

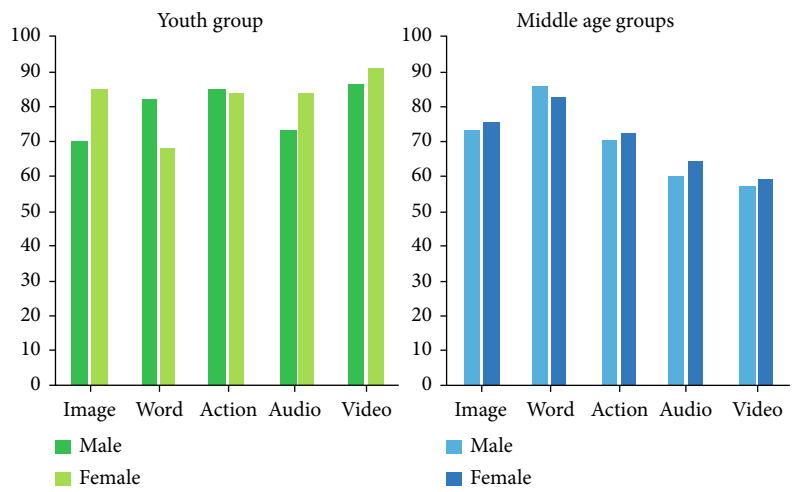


Figure 8: Evaluation of different interaction elements by different young groups.

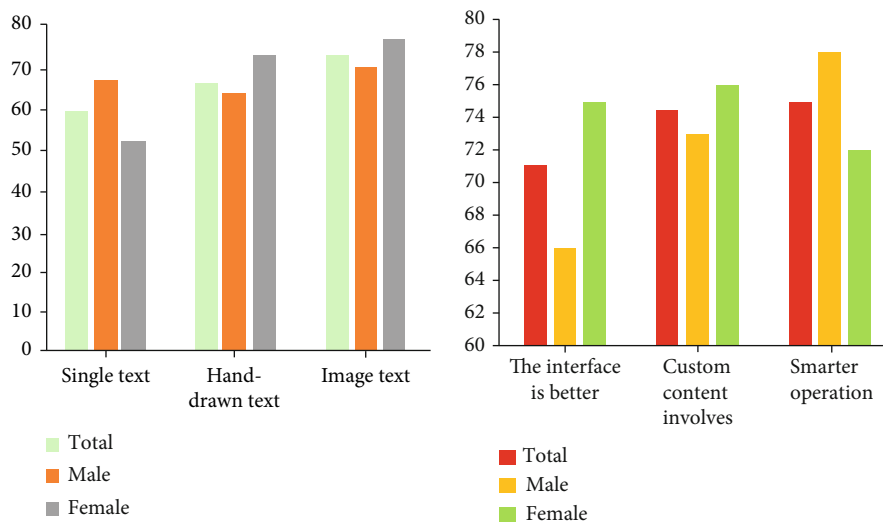


Figure 9: User experience of common font interaction design and areas for improvement.

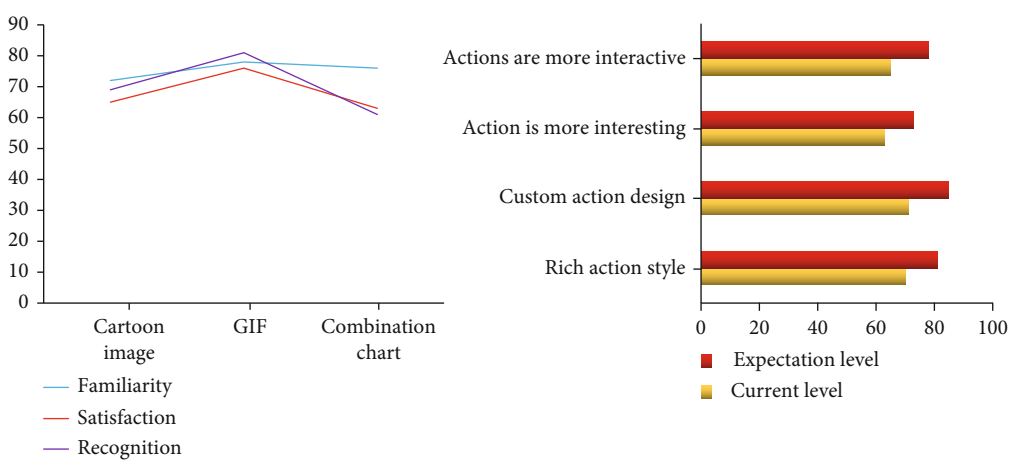


Figure 10: People's evaluation of images and expectations for action improvement.

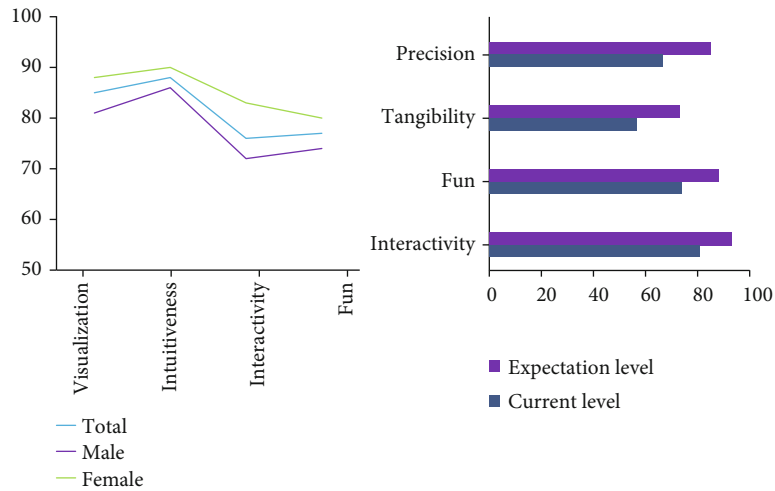


Figure 11: People's perception of audio-visual interaction and future expectations.

on their resource advantages to touch the users' emotional needs; impress the users; give the users a shock, anger, or sad, or joy experience; promote the users' thinking, and let them stay away from the boring life and enjoy the beauty and beauty of life [21].

3. Experiment and Analysis

3.1. Experimental Design. As of June 2021, the number of Chinese Internet users has exceeded to 1 billion, and the scale is expanding rapidly. Research data shows that the Internet usage rate of 18-45 years old accounts for more than 81%. Therefore, people in this age group are regarded as the core group of Internet surveys. In addition, people in this age group are more accepting of interaction design. Therefore, netizens of this age group are taken as the main survey subjects, and surveys about the needs and satisfaction of netizens in Internet interactive activities are collected through questionnaire surveys [22].

The questionnaire survey was distributed through e-mail, QQ, WeChat, and other online channels. A total of 105 people participated in the questionnaire, 99 of which were valid questionnaires, with a completion rate of 95%. The average age of the people participating in the survey was 30.2 years old. The average Internet age of the respondents is more than 3 years, as shown in Table 2. The purpose of this questionnaire is to explore the needs of users' Internet interaction design.

3.2. Data Analysis. The questionnaire asks questions about the netizens' feelings, evaluations, opinions, and suggestions on Internet interaction design, and the results are as follows [23].

For the interactive design of the Internet, we showed the investigators the forms of images, text, actions, audio, video, etc. In comprehensive comparison, they are more familiar with text, actions, and images, and the evaluation is more acceptable, and they have expectations for audio and video. Recognizing the application of video and audio in the field of Internet interaction in the future, the specific results are shown in Table 3.

Specifically segmenting the population, we found that 18 to 35-year-olds prefer the interactive operation of audio and video. They feel that they have a stronger sense of interaction, richer pictures, and more attractive. People aged 35-45 prefer words and images. For them, simple descriptions are more moving and conform to their general aesthetics. For the action interaction, the general acceptance is good, they think that the action is in line with their usage habits, and it is more convenient to use. The details are shown in Figure 8.

As shown in Figure 9, for text, a single text is often too dull and not highly recognized by users. Generally speaking, most interesting texts on the Internet are creative designs, or hand-drawn designs, with cute tones, and diverse fonts. When asked about the elements that users want to add, the interface is more "good-looking," the customized content design, and the operation is more "smart" are the choices of most people.

As shown in Figure 10, for images, a single image is sometimes too rigid. Cartoon images, animated images, and combined images are more in line with people's understanding of Internet images. These types of image combinations meet people's needs for nonstationary and transformable pictures. This is also the direction for future image improvement. For actions, the existing actions are limited and not interactive. In the future, action styles should be enriched, and custom action designs should be provided to make actions more interesting.

As shown in Figure 11, video and audio are the direction of future image interaction, which includes the advantages of images, text, and actions. People have higher expectations for video and audio. Most people have a good opinion of the visualization, intuitiveness, and interactivity of video and audio. For the future development of video and audio, people have put forward new requirements, looking forward to the future; video and audio will make breakthroughs in interactivity, visualization, and interactivity and achieve deep-level "human-computer interaction" [24, 25].

4. Discussion

This article investigates the young people's feelings and evaluations of the interactivity of Internet products. Most people's evaluations of various interactive products on the Internet are positive. After an in-depth discussion, they also found the shortcomings of interactive products. Interaction design is still limited to a single interactive product, which is not interesting enough and has limited user appeal. Interaction design is currently only at the initial stage of development, but the attention of users on the Internet is limited. It is impossible to capture the user's attention in a short period of time, and it is difficult to realize the next series of strategies. When the technology cannot be improved quickly, there is a lot to do in satisfying the interest of users. In the future, the design of Internet products should investigate user needs, use rich expressions, and combine different design elements to enhance the interest of products, attract users' attention, and achieve a greater degree of interactivity. Through the research on the interactive interface theory, this paper organizes and summarizes the interface design and interaction design of Internet products, summarizes the methods of interface design and interaction design of Internet products, and uses specific cases for design analysis. At the same time, it studies the users of Internet products, summarizes the characteristics of the user experience of Internet products, and proposes a dynamic division method for product user groups on the basis of traditional user group division.

5. Conclusion

This paper conducts research on the interaction of Internet fun design resources, combined with the in-depth application of the network communication system, to investigate users' feelings and evaluations of the interaction of Internet products. After surveying people of different age groups, we have roughly understood that the current users' evaluations of Internet interactive products are positive. For specific interactive products with different elements, modern video elements are more attractive than traditional graphic actions. Interesting has a powerful advantage, which can help the product enhance the attractiveness and increase the added value. This provides a direction for future interactivity improvements. However, there are also some shortcomings: It has not been able to conduct a more comprehensive analysis and research on the interactive design of hardware; when studying the interestingness of interactive product design, the research on the reasons for the interestingness of people is not comprehensive enough, and I hope that follow-up research. There are professional and scientific research methods and equipment to make more in-depth and detailed discussions on the reasons for human interest and to conduct more comprehensive and in-depth learning of other related disciplines, so as to develop a more comprehensive interesting design in the Internet interactive design research on the methods applied in. The existence of the design is to solve the problem, the fun is the flavoring agent in the process, and its expression is not limited to certain categories, and there is a huge development space and prospects worth exploring.

Data Availability

No data were used to support this study.

Conflicts of Interest

The author declares that there is no conflict of interest with any financial organizations regarding the material reported in this manuscript.

References

[1] R. Woodbury, A. Mohiuddin, M. Cichy, and V. Mueller, "Interactive design galleries: a general approach to interacting with design alternatives," *Design Studies*, vol. 52, pp. 40–72, 2017.

[2] G. Sequenzia, G. Fatuzzo, S. M. Oliveri, and R. Barbagallo, "Interactive re-design of a novel variable geometry bicycle saddle to prevent neurological pathologies," *International Journal on Interactive Design & Manufacturing*, vol. 10, no. 2, pp. 165–172, 2016.

[3] S. Rizzuti and L. D. Napoli, "An interactive robust design approach in product development," *International Journal on Interactive Design & Manufacturing*, vol. 12, no. 2, pp. 677–688, 2018.

[4] W. T. Abbood, O. I. Abdullah, and E. A. Khalid, "A real-time automated sorting of robotic vision system based on the interactive design approach," *International Journal on Interactive Design and Manufacturing (IJIDeM)*, vol. 14, no. 1, pp. 201–209, 2020.

[5] K. H. Kim, J. C. Park, Y. S. Suh, and B. H. Koo, "Interactive robust optimal design of plastic injection products with minimum weldlines," *International Journal of Advanced Manufacturing Technology*, vol. 88, no. 5-8, pp. 1–12, 2016.

[6] L. F. Gutierrez, S. Vargas, F. Cervantes, G. Torres, and F. Ramos, "Simplifying the design of interactive simulations of deformable objects," *IEEE Latin America Transactions*, vol. 14, no. 1, pp. 391–397, 2016.

[7] A. Amory, "Activity theory and interactive design for the African storybook initiative," *Journal of Educational Multimedia & Hypermedia*, vol. 27, no. 3, pp. 293–308, 2018.

[8] H. Jim, "Art game: an early interactive design from the office of Charles and Ray Eames," *Interactions*, vol. 24, no. 3, pp. 27–35, 2016.

[9] H. S. Yoo and Y. J. Da, "Gesture interaction design based on user preference for the elastic handheld device," *Journal of the Ergonomics Society of Korea*, vol. 35, no. 6, pp. 519–533, 2016.

[10] C. K. Andersson, "Google is not fun : an investigation of how Swedish teenagers frame online searching," *Journal of Documentation*, vol. 73, no. 6, pp. 1244–1260, 2017.

[11] L. C. Vieira and C. D. S. F. Soares, "Assessment of fun in interactive systems: a survey," *Cognitive Systems Research*, vol. 41, pp. 130–143, 2017.

[12] I. M. Verner, A. Polishuk, and N. Krayner, "Science class with RoboThespian: using a robot teacher to make science fun and engage students," *IEEE Robotics & Automation Magazine*, vol. 23, no. 2, pp. 74–80, 2016.

[13] N. Tegn, M. Abdelnoor, L. Aaberge et al., "Invasive versus conservative strategy in patients aged 80 years or older with non-ST-elevation myocardial infarction or unstable angina pectoris (after eighty study): an open-label randomised controlled trial," *The Lancet*, vol. 387, no. 10023, pp. 1057–1065, 2016.

[14] A. Greenhill, K. Holmes, J. Woodcock et al., "Playing with science," *Aslib Journal of Information Management*, vol. 68, no. 3, pp. 306–325, 2016.

[15] C. Grange and H. Barki, "The nature and role of user beliefs regarding a website's design quality," *Journal of Organizational and End User Computing (JOEUC)*, vol. 32, no. 1, pp. 75–96, 2020.

[16] Z. Han, Y. Ma, G. Gong, R. Huang, L. Song, and Y. Bi, "White matter pathway supporting phonological encoding in speech production: a multi-modal imaging study of brain damage patients," *Brain Structure & Function*, vol. 221, no. 1, pp. 577–589, 2016.

[17] A. Fiebich, "Narratives, culture, and folk psychology," *Phenomenology and the Cognitive Sciences*, vol. 15, no. 1, pp. 135–149, 2016.

[18] J. Shen, Y. Liu, R. P. Thornburgh, A. R. Kreshock, and M. L. Wilbur, "Design and optimisation of an aerofoil with active continuous trailing-edge flap," *Aeronautical Journal*, vol. 120, no. 1231, pp. 1468–1486, 2016.

[19] S. Lee, S. Kim, M. Seo, and D. Har, "Synchronization of frequency hopping by LSTM network for satellite communication system," *IEEE Communications Letters*, vol. 23, no. 11, pp. 2054–2058, 2019.

[20] H. Futaki, J. Lee, and K. Hamabe, "Radio communication system, communication apparatus, radio communication network system and method therefor," *Mediterranean Journal of Social Sciences*, vol. 3, no. 2, pp. 145–150, 2017.

[21] S. K. Viswanath, C. Yuen, W. Tushar et al., "System design of internet-of-things for residential smart grid," *IEEE Wireless Communications*, vol. 23, no. 5, pp. 90–98, 2016.

[22] R. van Renesse, H. Weatherspoon, Z. Shen, and W. Song, "The Supercloud: applying internet design principles to interconnecting clouds," *IEEE Internet Computing*, vol. 22, no. 1, pp. 82–86, 2018.

[23] J. J. Zakelj and K. M. Passino, "Feedback controller design over the internet of things," *Engineering Applications of Artificial Intelligence*, vol. 82, pp. 96–101, 2019.

[24] K. Shilton, "Engaging values despite Neutrality," *Science Technology & Human Values*, vol. 43, no. 2, pp. 247–269, 2018.

[25] Z. Lv, X. Li, and W. Li, "Virtual reality geographical interactive scene semantics research for immersive geography learning," *Neurocomputing*, vol. 254, pp. 71–78, 2017.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 6960343, 9 pages
https://doi.org/10.1155/2022/6960343

Research Article

Recognition and Prediction of Badminton Attitude Based on Video Image Analysis

Songsong He[1] and Ruyao Gong[2]

[1]*School of Electronic Information Engineering, Ningbo Polytechnic, Ningbo, 315800 Zhejiang, China*
[2]*College of Sports and Health, Linyi University, Linyi, 276000 Shandong, China*

Correspondence should be addressed to Ruyao Gong; gongruyao@lyu.edu.cn

Received 7 February 2022; Revised 26 February 2022; Accepted 7 March 2022; Published 26 March 2022

Academic Editor: Hoon Ko

Among many sports, badminton is one of the most popular events and it is deeply loved by people. However, there is relatively little research on pose recognition and prediction of badminton, so this paper uses video image analysis to perform badminton action recognition prediction and action classification. In order to better realize the lower limb movement pattern recognition and motion posture prediction of badminton players, this paper chooses BP neural network algorithm to establish a badminton motion posture recognition and prediction model based on video image analysis. The simulation results of the models are compared, and it is found that the recognition of the motion pose and the prediction model established by the neural network are accurate, and they almost coincide with the actual motion trend. In this paper, a face detection and recognition framework is established. The algorithm is used to filter the pictures of the input recognition network. Through calculation, the average accuracy rate of the face detection algorithm realized in this paper can reach 92.6%. This shows that the face detection algorithm implemented in this paper basically meets the standards. In this paper, a single inertial sensor is used to classify and recognize badminton movements using sensors located on the right wrist, left wrist, waist, and right ankle. This shows that the right wrist is the best position and achieves different strokes.

1. Introduction

Since the 1980s, the behavior recognition of badminton players has attracted a large number of researchers. With the emergence and development of various advanced technologies, the content of badminton sports recognition has been continuously enriched. The current research areas mainly include badminton serve recognition, badminton shot recognition and analysis, badminton posture detection, and various action recognition and prediction. The above research on behavior recognition can provide theoretical and basic support for a variety of practical applications.

In order to better realize the lower limb movement pattern recognition and motion posture prediction of badminton players, this paper chooses BP neural network algorithm to establish a badminton motion posture recognition and prediction model based on video image analysis. In this paper, a face detection and recognition framework is established, and the algorithm is used to filter the pictures input to the recognition network. In this paper, a single inertial sensor is used to classify and recognize badminton movements using sensors located on the right wrist, left wrist, waist, and right ankle. This shows that the right wrist is the best position and achieves different strokes.

The Section 1 of this article introduces the research background of the article and explains the main work of the article. The Section 2 is a review of references at home and abroad on recognition and prediction of motion poses. The Section 3 is the algorithm model and specific implementation steps constructed by the article. The Section 4 is the experimental data and set parameters of the article, and the evaluation index is set. The Section 5 is the analysis and research of the experimental results, and the main conclusions of the full text are summarized in the Section 6.

2. Related Work

In view of the importance of current motion recognition and prediction, many research teams at home and abroad have conducted in-depth research. In [1], the author introduces a new method for single-shot gesture recognition, processing motion pictures associated with video to obtain a PCA model. In [2], the authors developed a new method for gesture recognition using jointly calibrated jumping motions. The correct recognition rates were reported as 94.2%, 95.1%, and 90.2%, respectively. In [3], the authors propose a novel system for measuring the 3D motion of multiple independently moving objects at a macroscopic distance. The author also developed a method for measuring 3D micro-motion histograms of multiple independent moving objects without tracking a single motion trajectory. In [4, 5], the authors propose an interactive diagnosis system based on augmented reality for preoperative coronary heart disease gestures. All gesture recognition experiments have shown the discriminative ability and generalization ability of the algorithm. In [6], the authors proposed a method of motion segmentation and classification based on sequence alignment, which reconstructs the template sequence by estimating the average motion of the category.

However, there are many researches on badminton at home and abroad, but mainly focus on the analysis of the video images to study the relationship between racket and ball speed, the relationship between arm movement and racket speed and ball speed, and the study of badminton trajectory. However, few studies have focused on the identification and prediction of badminton players' movements.

Super-resolution reconstruction of video images is a research hotspot in the field of image processing [7]. In [8], in order to overcome the influence of stray light and impurities on the video laryngoscope image, the author eliminated the stray light and image impurities and improved the image quality. In [9], this paper proposes a one-dimensional (1-D) coding framework for images and videos based on deep learning neural networks and image patch clustering. The proposed method can achieve higher compression ratio and peak signal-to-noise ratio at the same time than the existing methods [10]. In [11], the author proposed a deep learning method. Experimental results show that the author's method is superior to the latest method in pairing drone video image patches. In [12], the author introduced a study of a new type of traffic flow measurement system by processing regularly acquired video images, which can determine the size of each vehicle and measure its two-dimensional motion. In order to reduce bandwidth requirements, the redundancy present in multimedia signals must be removed [13–15].

3. Method

3.1. Three-Dimensional Trajectory Tracking and Prediction Model Based on Kinematics Equation of Badminton

3.1.1. Badminton Trajectory Prediction. There are slight differences in the position of the points on the shuttlecock in space, because these slight differences have little effect on the overall flight trajectory prediction. In this paper, in the prediction of the trajectory, the shuttlecock is regarded as the centroid, and the median of the disparity map is used as the parallax of the centroid. Calculate three-dimensional space coordinates, the calculation formula is as follows:

$$\begin{bmatrix} X \\ Y \\ Z \\ W \end{bmatrix} = Q \begin{bmatrix} x \\ y \\ d \\ 1 \end{bmatrix}, Q = \begin{bmatrix} 1 & 0 & 0 & -c_x \\ 0 & 1 & 0 & -c_y \\ 0 & 0 & 0 & f \\ 0 & 0 & -1/T_x & (c_{xl} - c_{xr}) \end{bmatrix} \tag{1}$$

Where x and y are the pixel coordinates of the badminton centroid in the left image, d is the parallax, and the three-dimensional coordinates of this point are (X/W, Y/W, Z/W). To simplify the model, this article assumes that C, ρ, and S are constant during the flight of badminton, so k is a constant. To simplify the calculation, the force of badminton is decomposed:

$$\begin{cases} a_x m = f_x = f \sin\varphi \cos\theta = -kv_x^2 \\ a_y m = f_y = f \sin\varphi \sin\theta = -kv_y^2 \\ a_z m = f_z = f \cos\varphi - mg = -kv_z^2 - mg \end{cases} \tag{2}$$

The estimation accuracy of the least squares method depends on the accuracy of the measurement. If the obtained data contains large noise, the results obtained by using the least squares method will be far from the true value. Badminton is an irregular object. There is an error in the estimation of the center of mass. Due to the limitation of the camera's resolving ability at a distance from the camera, there is also an error in the calculation of parallax. In order to improve the accuracy of solving the kinematics equation of badminton, it is necessary to pre-process the obtained space track position of the empty wool ball to reduce the noise in the data.

3.1.2. Kalman Filter. The following badminton takes the x-axis movement as an example to construct an extended Kalman filter. The extended Kalman filter state equation is:

$$x_k = f(x_{k-1}, u_{k-1}, w_{k-1}), z_k = h(x_k, v_k) \tag{3}$$

Expand the Taylor series of $f(x_{k-1}, u_{k-1}, w_{k-1})$ and $h(x_k, v_k)$, and take a term to get:

$$x_k = \tilde{x}_k + A(x_{k-1} - \hat{x}_{k-1}) + Ww_{k-1}, z_k \approx \hat{z}_k + H(x_k - \tilde{x}_k)Vv_k \tag{4}$$

For movement $w_{k-1} = 0$ in the x and y directions, and movement $W = [0\,0\,1]^T$ and $w_{k-1} = -g$ in the z direction. Using extended Kalman filter to filter the badminton flight trajectory can get the badminton flight trajectory closer to the real value.

3.2. Construction of Face Detection Model Based on Video Images. The extraction of face samples, non-face samples and some face samples is accomplished in the following steps:

Step 1: Randomly select candidate frames from Wider_face and calculate the IOU value from the labeled data. If it is greater than 0.65, then the candidate area is a face. If it is greater than 0.4 and less than 0.65, then it is part of the face data, and less than 0.4 is non-face data;

Step 2: There will be border data for face data and some face data, but there is no border data for non-face data.

After everything is ready, the real face detection of the video begins [16–18]. The first step is to generate an image pyramid. Enter the picture and calculate the transformed image size using the following formula:

$$\min_L = org\_L(12/minsize)factor^n, n = [0, 1, 2, 3, \cdots, N] \quad (5)$$

The number of n is the number of pictures that can be scaled out. It should be noted here that the zoom size cannot be less than 12, when the length or width of the image is about to be less than 12, the zoom stops; minsize is the smallest detectable image, and factor is the zoom factor when forming the image pyramid.

Step 3: Enter the image into the first layer network P_{net}. Put each picture in the image pyramid obtained in the previous step into the neural network of this layer to train, and get the results of face score and border regression. After obtaining the face score, it is filtered according to the score, and the sliding box with the score below the threshold (0.6) is excluded. Then, the candidate boxes retained after the screening are merged by the NMS method.

Step 4: Use the R_{net} layer to perform more detailed processing on the image, and score and filter the face frame obtained by P_{net} again. In order to complete the task, first cut the face frame output from the P_{net} network from the original image, and shrink it to a size of $24*24*3$, and input it into the R_{net} layer.

3.3. Motion Pose Recognition and Prediction Algorithm Based on Neural Network

3.3.1. BP Neural Network Algorithm Based on L-M Back Propagation Algorithm. In the above formula, $y'(k)$ is the expected output of the neural network, and $y(k)$ represents the actual output of the network.

$$\Delta w = \left(J^T J + \mu I\right)^{-1} J^T e \quad (6)$$

In the above formula, e is the error vector generated by the learning and training of the neural network, J is the Jacobian matrix of the error versus weight differentiation, and μ is a scalar. With the change of μ, the L-M back propagation algorithm will produce different training effects. When μ increases, the training effect of the L-M algorithm gradually approaches the gradient descent algorithm. When μ decreases to 0, the training effect of the algorithm is equivalent to Newton's algorithm. When the final sum of the squared errors is less than a certain target error, the algorithm can be considered to converge.

3.3.2. Self-Organizing Competition Neural Network Algorithm. The input vector of a self-organizing competitive neural network is usually a binary vector, that is, all the elements in the input vector are 0 or 1, the relationship between the neuron j in the competition layer and the i in the input layer and the connection weight w_{ij} are as follows:$s_j = \sum_{i=1}^{N} w_{ij} x_i$. Where x_i is the ith element of the input sample vector. After the competition is over, the connection weights of the winning neurons need to be adjusted and optimized, as shown in the following formula.

$$w_{ij} = w_{ij} + \propto \left(\frac{x_i}{m} - w_{ij}\right) \quad (7)$$

Among them, $\propto$ is the learning parameter of the neural network, which usually satisfies the condition $0 < \propto << 1$, and generally takes a value between 0.01 and 0.03 in specific experiments; m represents the number of input layer neurons with an output value of 1, and x_i/m represents that when x_i is 1, the corresponding connection weight increases, otherwise the connection weight decreases. As mentioned above, all connection weights satisfy the condition that "sum is 1", so if the connection weight w_{ij} of a certain input layer neuron changes in some way, the remaining connection weights may change in the opposite direction.

4. Experiment

4.1. Data Sources

4.1.1. Equipment and Environment. The wireless node signal acquisition software used in this research was developed under the Windows platform, with Visual Studio 2010 as the development environment, and written in C # language. The algorithm used in this study for data preprocessing and action classification is implemented under the Microsoft Windows platform with Matlab as the development platform and M language. The wireless inertial sensor node is mainly composed of an accelerometer and a gyroscope. The accelerometer model is ADXL326, and its measurement range is ±16 g; the gyroscope models are LPR550 and LY550, with a measurement range of ±500dps, which fully meets the needs of high-speed sports monitoring such as badminton. The wireless inertial sensor node transmits data to the personal computer terminal through the wireless receiving node.

4.1.2. Experimental Data. A total of 12 volunteers participated in the experiment. The volunteers were members of the XX University badminton school team and the badminton association. The volunteers were between 18-25 years old. Before the experiment, all volunteers were informed of the purpose of the experiment and the course of the experiment. Table 1 shows the types of actions performed for this experiment. A1-A12 are batting actions and A13-A14 are non-batting actions. The four wireless sensor nodes used in the experiment were worn on the left wrist, right wrist, right waist and right ankle of volunteers. The sampling frequency

Table 1: Types of human shots performed in the experiment.

	Action number	Action type	Action number	Action type
Batting action	A1	Forehand serve	A7	Backhand net rubbing
	A2	Backhand serve	A8	Forehand
	A3	Forehand high far ball	A9	Backhand
	A4	Backhand high far ball	A10	Forehand push ball
	A5	Bash the ball in front of the forehand	A11	Forehand corner kick
	A6	Rub the ball in front of the forehand	A12	Backhand corner kick
Non-batting	A13	Pick up the ball while walking		
	A14	Prepare your posture while shaking your wrists up and down		

Table 2: One-to-one correspondence between the results of motion pose recognition and motion patterns based on neural networks.

Recognition result	Specific exercise mode	Recognition result	Specific exercise mode
0 0 0 0 0 1	A1 forehand serve	0 0 1 0 0 0	A7 backhand net rubbing
0 0 0 0 1 0	A3 forehand high far ball	0 1 0 0 0 0	A10 forehand push ball
0 0 0 1 0 0	A5 bash the ball in front of the forehand	1 0 0 0 0 0	A13 pick up the ball while walking

of the wireless node in the experiment is 100 Hz, and each action is repeated 55 times.

4.2. Experimental Parameter Settings. This article randomly selects 6 movements in 14 sets of data, and collects 100 sets of data, respectively, for a total of 600 sets of data. Then randomly select 50 groups from each action sample as the training sample set of the neural network classifier, and then randomly select 50 sets of feature data as the test sample set to test the accuracy of the classifier. For the training samples of various actions, the corresponding output nodes are set to 1, and the other output nodes are set to 0, which are used as the desired output for training.

According to the approximate relationship between the number of neurons in the hidden layer and the number of neurons in the input layer, and multiple experiments, this paper chooses a hidden layer with 13 neurons; in the end, there are 6 types of motion patterns that need to be identified from the 14 sets of data, so the number of neurons in the output layer is 6. In addition, in order to improve the learning rate, the learning rate of the BP neural network model is set to 0.5; the target error is 0.001, that is, the training is stopped when the number of trainings exceeds 1000 times or the training error is less than 0.001. According to the T vector representing the expected output, we can see that the one-to-one correspondence between the recognition results of the motion pose recognition and prediction model based on the neural network and the specific motion mode is shown in Table 2 below.

4.3. Badminton Attitude Recognition Process. This article uses BP neural network to recognize badminton strokes. The specific steps are as follows:

(1) Collect acceleration and angular velocity data of batting action and non-batting action, and extract feature values to obtain feature vectors. Each feature vector is a sample

(2) This paper uses neural network pose recognition and prediction algorithm for posture recognition. Firstly, the algorithm model is used to distinguish the two categories of hitting action and non-hitting action, and then the model is used to identify the type of hitting action. The attitude recognition model is trained based on data including batting and non-batting movements; the attitude prediction model is trained based on data on 12 batting movements. In both models, algorithms using neural networks are used for classification and recognition. The specific action recognition process is shown in Figure 1

4.4. Evaluation Index. Some indicators commonly used for the recognition and detection of motion gestures are: Accuracy, Precision, Recall, and so on. The simplest of these is accuracy, and the most commonly used are accuracy and recall. This article will test the algorithm through these two indicators and evaluate the performance of the algorithm.

True positive (TP): a subset of samples that are detected as true and actually also true;

True negative (TN): a subset of samples that are detected as negative and actually not;

False positive (FP): a subset of samples that are detected as true but not actually;

False negative (FN): A subset of samples that are detected as negative but actually true.

The calculation formula for accuracy is:

$$Accuracy = \frac{TP + FN}{P + N} \tag{8}$$

The calculation formula of accuracy is:

$$Precision = \frac{TP}{TP + FP} \tag{9}$$

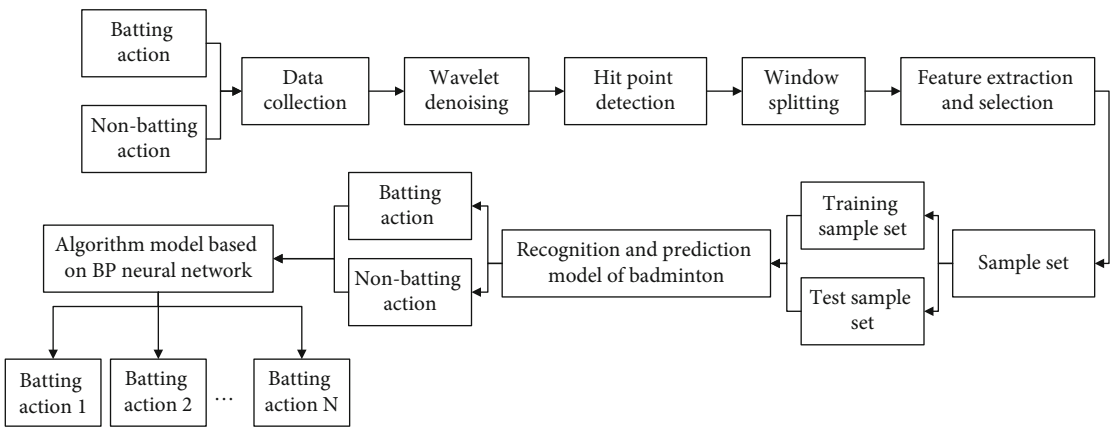


Figure 1: Flowchart for recognition and prediction of badminton movement based on neural network.

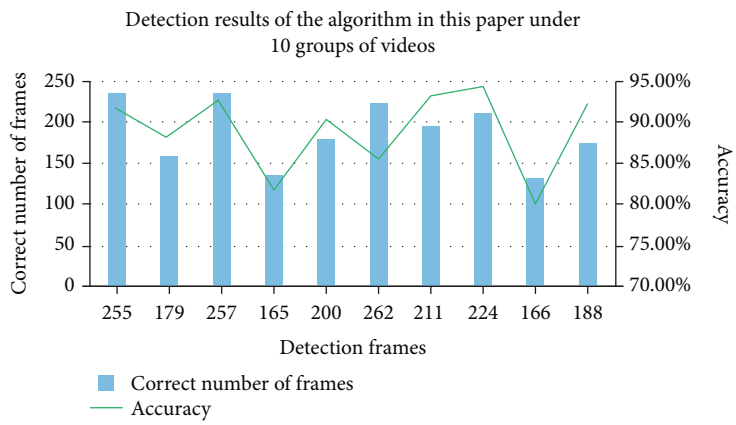


Figure 2: Detection results of the algorithm in this paper under 10 groups of videos.

The formula for calculating the recall rate is:

$$Recall = \frac{TP}{P} \tag{10}$$

5. Results and Discussions

5.1. Performance Evaluation of Face Detection Algorithms Based on Video Images

5.1.1. Test Accuracy. Next, the performance of the algorithm is tested for accuracy. A total of 100 surveillance video clips were collected in this article. Each clip is about 20 seconds. The resolution of the video is 1280×720 and the frame rate is 29 frames/second. The detected lighting conditions include strong light, low light and light changes. The detection scenarios are indoor and outdoor. When testing the accuracy, a total of 10 videos were randomly selected in each time period and scene. In order to ensure the quality of the test and the real-time detection, the video is detected every three frames. In this way, about 196 times will be detected for a video of about 20 seconds. Next, the results are shown in a chart.

Through the above-mentioned Figure 2, it can be intuitively felt that the face detection algorithm implemented in this paper has a high accuracy rate and has a good effect. By calculation, the average accuracy of the face detection algorithm realized in this paper can reach 92.6%. This shows that the face detection algorithm implemented in this paper basically meets the standards. When the light is weak or the face is severely occluded, the detection effect is poor, but for the case where the light is sufficient and the face is full, the detection effect is good.

5.1.2. Evaluation of Experimental Performance. By comparing the face detection and tracking algorithm implemented in this paper with the best face extraction algorithm and the test results without the algorithm, the following Table 3 can be obtained.

After obtaining the comparison data results, a histogram is used to visually show the comparison of the index results in the two cases, as shown in Figure 3.

Through the above test results, we can analyze:

(1) Compared with the data in rows 2 and 3 under column 3, the addition of the face extraction algorithm only reduces the recognition rate by 0.8%, and the recognition rate does not decrease significantly. It is proved that the extraction algorithm does filter out

Table 3: Performance test of face extraction algorithm.

	Data compression ratio	Screening image recognition ratio	Full face recognition rate
Add face extraction algorithm	9.3 (%)	83.1 (%)	78 (%)
No face extraction algorithm added	100 (%)	83.9 (%)	72 (%)

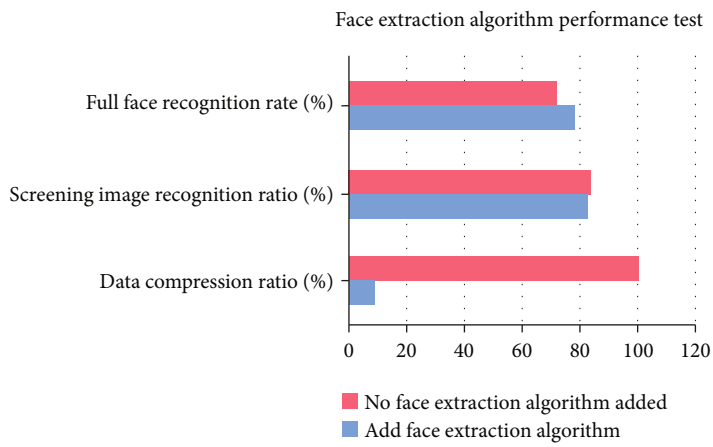


Figure 3: Performance test chart of face extraction algorithm.

some low-quality video images and does not affect the overall recognition rate

(2) Analyze the data in rows 2 and 3 under column 4 and add the face retrieval algorithm to increase the recall rate by 6% compared to the case without adding recognition. This shows that the face extraction algorithm proposed in this paper does filter out some low-quality face images, so the recognition rate of the push video image is improved

(3) Analyze the data in rows 2 and 3 under the first column, and the compression rate of the data extracted by adding the face extraction algorithm is greatly reduced compared with the data extracted without adding the face. It shows that the face extraction algorithm presented in this paper reduces a lot of redundant data calculations

Through the above experimental analysis, we can get that the face extraction algorithm given in this paper can extract high-quality video images, meet the real-time needs of the system, and improve the accuracy and efficiency of subsequent recognition.

5.2. Analysis of Posture Recognition Results in Badminton. In the actual game, you will encounter non-hitting actions such as walking and picking up the ball. The acceleration generated by these actions is greater than 1.5 g, which cannot be filtered out during window division. Therefore, it is necessary to filter out non-hitting actions before the hitting action recognition. This paper uses a model to distinguish between hitting and non-hitting actions. The specific classification and recognition results are shown in Table 4.

As can be seen from the table, the algorithm constructed in this paper can almost completely distinguish between the hitting action and the non-hitting action, further reducing the impact of non-hitting information on the classification and recognition of hitting action. In order to verify the feasibility of the above indicators, the following explores the feasibility of hierarchical classification using only angular velocity data and both acceleration data and angular velocity data. The classification result is shown in Figure 4:

Table 4: Recognition results of batting action and non-batting action.

	Batting action	Non-batting action
Average recognition rate	99.46%	98.97%

It can be known from Figure 4 that when only the angular velocity data is used for classification, the recognition rate can reach 98.32%, and the feasibility of the above indicators is verified from another aspect. This paper also did an experiment of using the acceleration and angular velocity data to classify the classification. The recognition rate can reach 98.84%, which again shows that acceleration and angular velocity can be used to classify the classification. It shows that electronic equipment can be used to help coaches perform rating evaluations, making the rating evaluation more scientific. The following explores the impact of different body parts' motion data on the classification of levels. The motion data of the four body parts of the right wrist, right side of the waist, right ankle, and left ankle are used to classify the classification. The recognition results are shown in the table:

It can be seen from Figure 5 that no matter whether only acceleration data or angular velocity data is used or both types of data are used at the same time, the classification recognition rate obtained by the right wrist is the highest. This shows that the right wrist is the best place for classification,

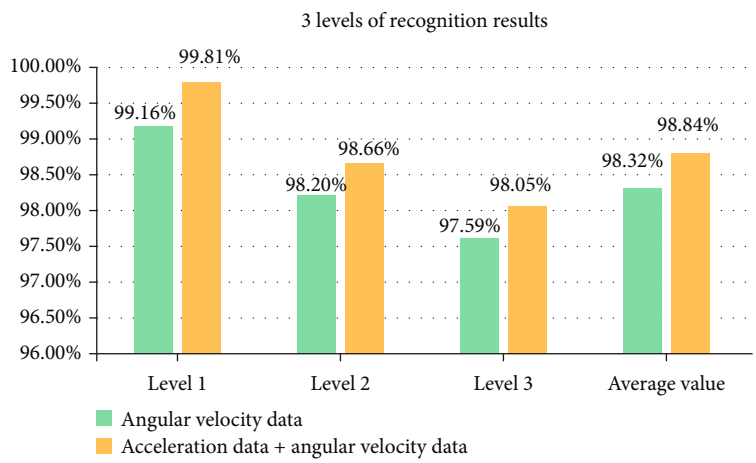


Figure 4: 3 levels of recognition results.

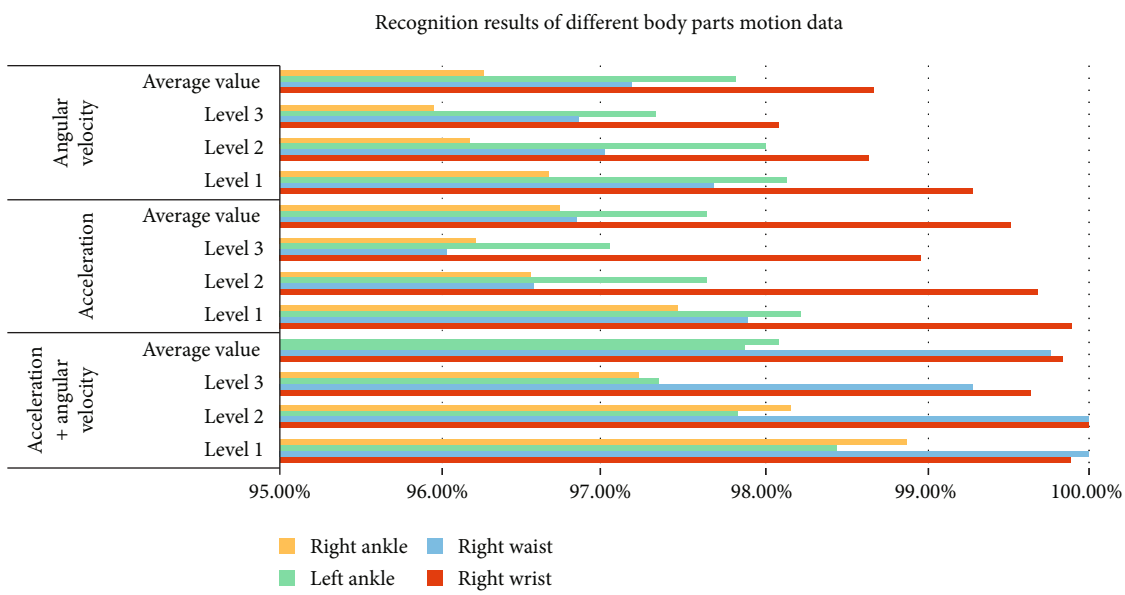


Figure 5: Recognition results of different body parts motion data.

and it is the waist, left ankle, and right ankle. By comparing the data of Figures 4 and 5, it can be known that when a single sensor is used for classification, the recognition result using acceleration and angular velocity data at the same time is the best, followed by acceleration data and finally angular velocity data. However, the recognition rate obtained from the data of multiple body parts is higher than the recognition rate of a single body part, indicating that badminton is a sports that requires the coordination of different parts of the human body.

5.3. Test of Badminton Attitude Recognition System Based on Video Image Analysis

5.3.1. Analysis of System Runtime. Under a master node and a 7-node system cluster, the processing time of each module is recorded by processing 50G video data. The module running time distribution is shown in Figure 6:

It can be obtained from the data of the graph that the first few process video key frame acquisition, image preprocessing and motion gesture recognition all need to perform massive video processing, resulting in the need to consume most of the time, accounting for 74.04% of the entire application time. After the first few processes are screened and optimized, the amount of data in the subsequent kinematic attitude prediction process is greatly reduced, so the time consumed in these processes is also greatly reduced, accounting for 7.05% of the entire system application time.

5.3.2. Analysis of System Cluster Size on Data Processing Efficiency. The system tests the impact of cluster size on data processing efficiency by processing the time required for 50G data in different system cluster sizes. First, configure the size of the cluster by configuring the number of worker nodes, and the number of worker nodes will be configured to 1, 3, 5, and 7.

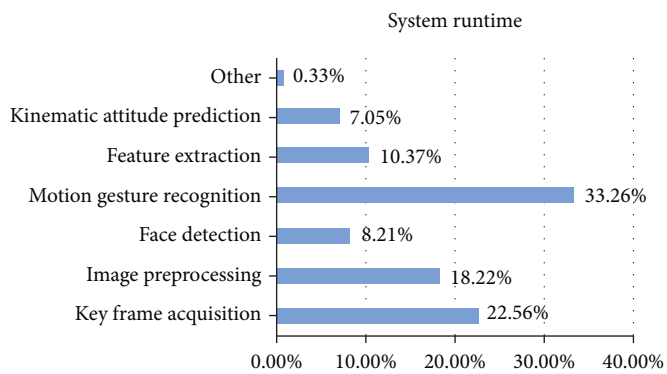


Figure 6: System module running time.

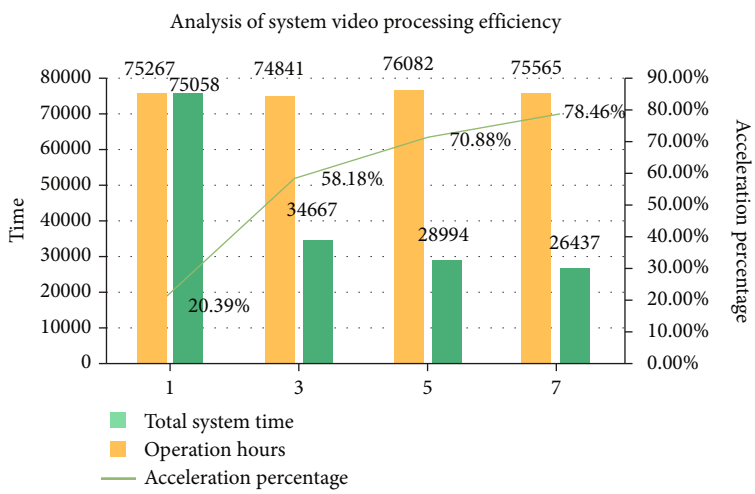


Figure 7: Analysis of cluster size versus data processing efficiency.

The experimental results are shown in Figure 7. When there are 7 worker nodes, the system cluster processing time for processing 50G data is 75058 seconds, and the total time for the system cluster machine to run 7S is reduced by 207 seconds compared to 1 worker node. More importantly, the application runtime 7 T is 26437 seconds with 7 Worker nodes, which is 48621 shorter than the 78539 seconds of a single Worker node. It can also be seen from the experimental data iA that 7 Worker nodes can accelerate 78.46% compared to a single Worker node. It can be concluded that the larger the system's cluster (that is, the more processing nodes), the less time the system needs for data processing and the faster the efficiency.

6. Conclusions

In order to better recognize the lower limb movement pattern and predict the posture of the badminton player, this paper converts the knee joint angle signal into the knee joint angle characteristic value, and performs simple normalization processing on the knee joint angle signal. In this paper, the more mature BP neural network algorithm and self-organizing competitive neural network are used to establish a badminton sport posture recognition model, and the model is trained and simulated separately. The simulation results of the models are compared, and it is found that the recognition of the motion pose and the prediction model established by the neural network are accurate, and they almost coincide with the actual motion trend.

In order to improve the efficiency and accuracy of face recognition, this paper establishes a face detection and recognition framework, and presents a face recognition and feature extraction algorithm. The algorithm is used to filter the pictures of the input recognition network, thereby improving the performance of the overall recognition process. By calculation, the average accuracy of the face detection algorithm realized in this paper can reach 92.6%. This shows that the face detection algorithm implemented in this paper basically meets the standards. When the light is weak or the face is seriously blocked, the detection effect is poor, but when the light is sufficient and the face is full, the detection effect will be very good.

This article uses a single inertial sensor to classify badminton movements, and uses sensors located on the right wrist, left wrist, waist, and right ankle to classify badminton movements. The recognition rates are 99.84%, 99.76%, 97.87%, and 98.08%. The right wrist has the highest recognition rate, indicating that the right-handed athletes mainly

control the racket through the right wrist and achieve different strokes, indicating that the right wrist is the best position.

Data Availability

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

Conflicts of Interest

The author(s) declared no potential conflicts of interest with respect to the research, author- ship, and/or publication of this article.

Acknowledgments

This work was supported partly by Zhejiang Provincial Natural Science Foundation of China under Grants No. LGG20F020013.

References

[1] H. J. Escalante, I. Guyon, V. Athitsos, P. Jangyodsuk, and J. Wan, "Principal motion components for one-shot gesture recognition," *Pattern Analysis & Applications*, vol. 20, no. 1, pp. 167–182, 2017.

[2] W. Zeng, C. Wang, and Q. Wang, "Hand gesture recognition using leap motion via deterministic learning," *Multimedia tools and applications*, vol. 77, no. 21, pp. 28185–28206, 2018.

[3] B. M. Smith, P. Desai, V. Agarwal, and M. Gupta, "Colux: multi-object 3d micro-motion analysis using speckle imaging," *ACM Transactions on Graphics*, vol. 36, no. 4, pp. 1–12, 2017.

[4] Y.-b. Zou, Y.-m. Chen, M.-k. Gao, Q. Liu, and D.-h. Zhang, "Coronary heart disease preoperative gesture interactive diagnostic system based on augmented reality," *Journal of Medical Systems*, vol. 41, no. 8, p. 126, 2017.

[5] X.-H. Hu, W.-B. Li, and J.-M. Kan, "Gesture control technology based on surface electromyography for automatic pruning machine," *Journal of Beijing Forestry University*, vol. 39, no. 2, pp. 117–124, 2017.

[6] M. Liu, X. Zhang, and X. Guiyun, "Continuous motion classification and segmentation based on improved dynamic time warping algorithm," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 32, no. 2, article 1850002, 2017.

[7] X. Yang and Y. L. Tian, "Super normal vector for human activity recognition with depth cameras," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 39, no. 5, pp. 1028–1039, 2017.

[8] X. Mao and C. Li, "Combining point context and dynamic time warping for online gesture recognition," *Journal of Electronic Imaging*, vol. 26, no. 3, article 033023, 2017.

[9] Y. Li, Q. Miao, K. Tian, Y. Fan, and J. Song, "Large-scale gesture recognition with a fusion of rgb-d data based on saliency theory and c3d model," *IEEE Transactions on Circuits & Systems for Video Technology*, vol. 28, no. 10, pp. 2956–2964, 2017.

[10] P. Głomb, M. Romaszewski, A. Sochan, and S. Opozda, "Unsupervised parameter selection for gesture recognition with vector quantization and hidden markov models," *Proceedings of the American Mathematical Society*, vol. 72, no. 2, pp. 422–424, 2017.

[11] Y. M. Lui, "Human gesture recognition on product manifolds," *Journal of Machine Learning Research*, vol. 13, no. 1, pp. 3297–3321, 2017.

[12] Z.-Q. Feng, X.-W. Yang, T. Xu, H. Liu, and Z.-P. Xu, "Gesture recognition based on combining gesture binary descriptor and hausdorff-like distance," *Acta Electronica Sinica*, vol. 45, no. 9, pp. 2281–2291, 2017.

[13] H. Zhu, H. Wei, B. Li, X. Yuan, and N. Kehtarnavaz, "Real-time moving object detection in High-Resolution video sensing," *Sensors*, vol. 20, no. 12, p. 3591, 2020.

[14] O. I. Khalaf, C. A. T. Romero, A. A. J. Pazhani, and G. Vinuja, "VLSI Implementation of a High-Performance Nonlinear Image Scaling Algorithm," *Journal of Healthcare Engineering*, vol. 2021, Article ID 6297856, 2021.

[15] M. Abdolmaleky, M. Naseri, J. Batle, A. Farouk, and L. H. Gong, "Red-green-blue multi-channel quantum representation of digital images," *Optik*, vol. 128, pp. 121–132, 2017.

[16] S. Ding, S. Qu, Y. Xi, and S. Wan, "Stimulus-driven and concept-driven analysis for image caption generation," *Neurocomputing*, vol. 398, pp. 520–530, 2019.

[17] Z. Wan, Y. Dong, Z. Yu, H. Lv, and Z. Lv, "Semi-supervised support vector machine for digital twins based brain image fusion," *Frontiers in Neuroscience*, vol. 15, article 705323, 2021.

[18] H. Song and M. Brandt-Pearce, "A 2-D Discrete-Time Model of Physical Impairments in Wavelength-Division Multiplexing Systems," *Journal of Lightwave Technology*, vol. 30, no. 5, pp. 713–726, 2012.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 3960120, 16 pages
https://doi.org/10.1155/2022/3960120

Research Article

Cyclic Anaerobic Exercise Performance and Neuromuscular Activity Based on Artificial Intelligence Genetic Algorithm

Lun Wan and Jiannan Liu

Ministry of Sports, Chongqing Jiaotong University, Chongqing, 400074 Chongqing, China

Correspondence should be addressed to Jiannan Liu; 980200900013@cqjtu.edu.cn

Received 17 January 2022; Revised 21 February 2022; Accepted 7 March 2022; Published 26 March 2022

Academic Editor: Hoon Ko

Anaerobic exercise capacity and neuromuscular activity research is an important content and an emerging research field in sports training research. At present, the understanding of anaerobic exercise ability in academia is still at the general and overall cognitive level, and the understanding and application of anaerobic exercise ability cannot fully meet the needs of competitive sports practice. In order to solve these problems, this paper proposes a cyclic anaerobic exercise performance and neuromuscular activity based on artificial intelligence genetic algorithm, aimed at studying the theory and application mechanism of anaerobic exercise capacity and neuromuscular activity and its application characteristics in competitive sports practice. The approach in this paper is to design genetic operators, compare artificial intelligence genetic algorithms, and test neuromuscular movements. The purpose of these methods is to provide exercisers with a feasible and more effective new method of daily training and to investigate whether this new training method can optimize anaerobic exercise in humans. In this paper, by studying the kinematic basis of anaerobic exercise capacity and the mechanism of neuromuscular regulation, a model of muscle neuron population in anaerobic exercise is established. The results showed that the CMC of the beta band was significantly higher than that of the alpha band at the same strength level, with a difference of 0.03.

1. Introduction

In the continuous development of sports training, the research on human exercise capacity, especially the research on anaerobic metabolism and aerobic metabolism, the basic pathways of material metabolism and energy metabolism in the body's exercise process, has always been the focus of debate and hot issue. Under normal circumstances, aerobic metabolism cannot meet the needs of the body at this time, so sugar undergoes anaerobic metabolism to quickly generate a large amount of energy. Exercise in this state is anaerobic exercise. From the point of view of material metabolism and energy metabolism, the exercise performance of the human body consists of the corresponding aerobic and anaerobic exercise performance. The performance of the human body in various sports or technical movements is nothing more than aerobic, anaerobic, or a combination of aerobic and anaerobic exercise. A detailed and systematic study of the theoretical mechanisms, basic characteristics, and application characteristics of human aerobic and anaerobic exercise capacity will be the core of future exercise training research. In this paper, an example of finger movement is selected to study the characteristics of anaerobic exercise and neuromuscular activity. The realization of human finger movement function is closely related to the control of cranial nervous system movement. From hand projection areas to motor cortex, this paper found that the brain's control mechanism of finger movement is very complex. At present, research on the neural regulation mechanism of the human hand mainly focuses on cortical neural activity, electromyography, and muscle strength.

At the same time, this research can reveal the improvement of the anaerobic exercise ability of the athletes by the optimal training method, provide certain theoretical guidance for the coaches and athletes to improve the training methods in training and competition, and promote the better development of sports. Different types of anaerobic fitness have their own basic characteristics and practical

significance, and different types of anaerobic fitness have sports advantages, technical movements, unique training methods and other characteristics and means, and the basic requirements of unique training. This paper redefines the concept of anaerobic exercise capacity in the theoretical system of sports training by studying the theory and application mechanism of anaerobic exercise capacity and the actual effect of sports training and discusses it from the perspective of biology and training. This paper studies neuromuscular activity in depth and can propose effective measures to prevent and treat neuromuscular junction, peripheral nerve, and muscle diseases. Anaerobic exercise performance of various types and forms of the human body in competitive sports practice has important theoretical significance and practical value for in-depth understanding of the theory and application mechanism of the body's anaerobic exercise ability in academics and disciplines, as well as the effective grasp and application of anaerobic exercise ability in sports training practice.

In this study, the anaerobic training method was applied to the stage training to improve the athletes' special anaerobic exercise ability, and a better training effect was obtained. This study is the first to design and analyze the characteristics of different anaerobic exercise capacity requirements and the relationship between anaerobic exercise capacities of the same athletes, so as to provide a reference for the evaluation of anaerobic exercise capacity. The implementation of the training tasks provided in this study is simple and convenient, and the requirements for the venue are not high. During the training process, various index data of athletes can be automatically recorded directly through software and training tools. By observing the surface EMG of lower limb muscles, this paper confirmed the effect of 30% 1RM weight-bearing additional vibration stimulation on the activation of lower limb muscles during finger movement, perfected the theory of weight-bearing vibration training, and provided a basis for the frequency selection of weight-bearing vibration training.

2. Related Work

Paying attention to the neuromuscular changes in anaerobic exercise is of great significance for improving human movement patterns and exercise outcomes. Tsujimura believes that genetic algorithms are one of the most powerful tools for solving job shop scheduling problems, especially for large-scale real-world problems. He demonstrated its performance through a standard benchmark for job shop scheduling problems with two different fuzzy subset ranking methods, where he fused a genetic algorithm, where the processing time is represented by a fuzzy number [1]. Qiang studied a global optimization method combining genetic algorithm and Hooke-Jeeves method to solve a class of constrained optimization problems. The method he proposed significantly improves the accuracy and convergence speed of genetic algorithms. He has studied many classic test problems through precise penalty functions [2]. Yoshitomi proposed a modified genetic algorithm called Genetic Algorithm in Uncertain Environments, in which objective function constraints fluctuate according to the distribution functions of their random variables. The algorithm is applied to the stochastic optimal assignment problem, the stochastic knapsack problem, and the newly formulated stochastic image compression problem [3]. Tayebi designed to investigate the acute response of athletes to three types of aerobic, anaerobic, and wrestling sports. He randomly divided 24 volunteers into three groups and gave them three nonconsecutive exercise programs. To study acute responses, blood sampling was performed after the first and third exercise sessions; study shows exercise type has no effect on serum iron [4]. The purpose of Kang was to measure the muscle function and anaerobic exercise capacity of the knee joint according to three exercise programs. He selected 21 athletes as research subjects. The results demonstrated that the tested athletes performed better than their peers on almost all metrics, including body composition, physical fitness factors, isokinetic muscle function, muscle strength, and anaerobic capacity [5]. Park proposed that the need for lactate cycling after aerobic exercise eliminates lactate produced during exercise, a process that requires energy expenditure. Supplementation with D-ribose increases muscle cell energy, promotes PRPP synthesis in cardiac and skeletal muscle, and eliminates pentose phosphate within the lower limit of glucose 6-phosphate dehydrogenase activity [6]. Morana designed to detect and determine the ability of RQA-sensitive fatigue components. He analyzed muscle activity with RQA to obtain percent certainty, and at the onset of exercise, a significant increase in Pt was observed in both groups [7].

3. Artificial Intelligence Genetic Algorithm

3.1. Design of Genetic Operators. Genetic operations include three genetic operators: selection, crossover, and mutation. Each operator in the group first performs a selection operation. The process of selecting excellent individuals from a population and excluding inferior individuals is called selection and replication. The selection operation is an operation used to determine how to select which individuals are inherited from the parent group to the next-generation group according to a specific method, and its main purpose is to avoid the loss of useful genetic information and improve global convergence and computational efficiency. The quality of the selection operator directly affects the calculation result of the genetic algorithm. If the operator makes an improper decision, it will lead to the increase of individuals with similar similarity values in the group, making the offspring individuals similar to the parent individuals, and the evolutionary stagnation. Or the value of the fitness function increases, which leads to the loss of the formation of genetic diversity and the problem of precocious puberty in the process of group development. The selection operation is based on an assessment of the suitability of the individuals in the group. Commonly used selection operators are fitness ratio method, optimal personal preservation method, sorting selection method, league selection method, etc. In addition, there are other selection options, such as exclusion method and expected value method. Each method has a different

impact on the search efficiency of the genetic algorithm. Among many selection operations, the fitness proportional selection method is the most basic and most commonly used method, but it is prone to large selection errors. The optimal individual preservation strategy ensures that the final result of the iteration is the individual with the highest fitness function value in the past generations. Then, after choosing the best individual preservation strategy, the fitness ratio method is used for the modern population [8]. Assuming a population size of n, the probability of individual i being selected is as follows.

$$P_i = \frac{f_i}{\sum_{j=1}^{n} f_i}. \tag{1}$$

In the formula, f_i represents the fitness function value of the ith chromosome. In this method, each chromosome is selected with probability P_i; it reflects the ratio of the individual's fitness to the individual's overall fitness. Arithmetic crossover refers to the creation of a new individual from a linear combination of two individuals. Assuming an arithmetic crossover between two objects x_1 and x_2, the two new individuals after the crossover are

$$\begin{cases} X_1' = ax_1 + (1-a)x_2, \\ X_2' = ax_2 + (1-a)x_1. \end{cases} \tag{2}$$

Among them, $a \in [0, 1]$, a is a real number.

As one of the control parameters of genetic algorithm, population size also affects the performance of genetic algorithm. This paper needs to increase the population size so that genetic engineering is more likely to process more patterns, generate meaningful gene blocks, and gradually evolve into optimal solutions. But the size of each group is too large, which increases the computational complexity of the algorithm and increases the convergence time [9]. The size of the group can be selected from 10 to 200 according to the actual situation.

During the optimization process, the crossover probability P_c always controls the dominant crossover operator. Inappropriate crossover probabilities can lead to unintended consequences. A high probability of crossover results in complete crossover every generation, but too high P_c is more likely to destroy good individuals and randomize search trends. Also, if P_c is too small, more individuals may go directly to the next generation and the search may stall. In general, it is recommended that the value of P_c be in the range of $[0.5, 1]$. The mutation probability P_m controls how often the mutation operation is used. A larger P_m can generate more individuals and increase the diversity of the group, but it can also destroy good individuals, making the performance of the genetic algorithm approximate to random search.

In order to avoid the tedious work of determining the crossover probability P_c and the mutation probability P_m by repeated experiments, this paper adopts an adaptive genetic algorithm in which P_c and P_m can change with the fitness value, as shown in the formula.

$$\begin{aligned} P_c &= \begin{cases} \left(f_{\max} - f'\right)/(f_{\max} - f_v), & \left(f' \geq f_v\right), \\ 1, & \left(f' < f_v\right), \end{cases} \\ P_m &= \begin{cases} \frac{1}{2} * (f_{\max} - f)/(f_{\max} - f_v), & \left(f' \geq f_v\right), \\ (f_v - f)/(f_v - f_{\min}), & \left(f' < f_v\right). \end{cases} \end{aligned} \tag{3}$$

In the formula, $f_{\max}$ is the maximum fitness value of the previous group, $f_{\min}$ is the minimum fitness value of the previous group, f_v is the average fitness value of the current group, and f' is the fitness value of the two chromosomes participating in the crossover and is the fitness value of the chromosome with a high-order value. Figure 1 shows the-training target curve of the pure genetic algorithm.

As can be seen from Figure 1, the genetic algorithm can converge to the error target value of 0.002 at step 815. After initializing the coding, calculating the objective function, transcoding the binary numbers and programming the decimal numbers, and then passing through inheritance, crossover, and mutation, the maximum fitness value and the optimal individual in the population are finally obtained, as shown in Figure 2.

It can be seen from Figure 2 that the genetic algorithm has found the global maximum point of the nonlinear function. The importance of each node in the network depends not only on its own objective function value but also on the number of adjacent nodes and the ranking of the objective function value in adjacent nodes [10]. To assess the importance of nodes, this chapter provides the following definition of the node consistency NF of an exponential network.

$$\mathrm{NF} = \left(\frac{1}{k+1}\right)^{R-1}, \tag{4}$$

where R is the rank of the objective function value of nodes near k. $R = 1$ means that the objective function value of the node is lower than that of all neighboring nodes. Therefore, the goodness-of-fit value of this network node is 1. With the increase of R, the rank of the node's objective function value on the adjacent nodes also increases so that the node's network node fitting degree decreases. In this paper, a fitness calculation method is proposed, which is defined as follows.

$$\mathrm{NF} = \frac{k - R + 1}{k}. \tag{5}$$

There are still some problems with this definition; the fitness value of some nodes is equal to 0, which cannot reflect the difference between nodes. Figure 3 shows how the fitness of network nodes changes with the change of topology. When the connecting edges of nodes change, the fitness of network nodes also changes [11].

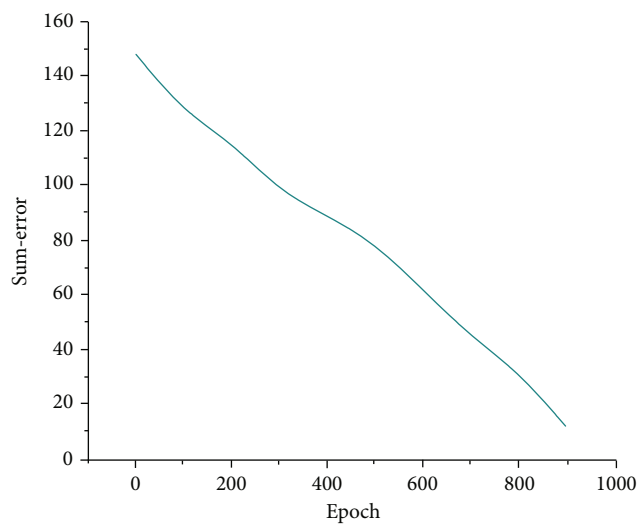


Figure 1: Target curve for standard algorithm training.

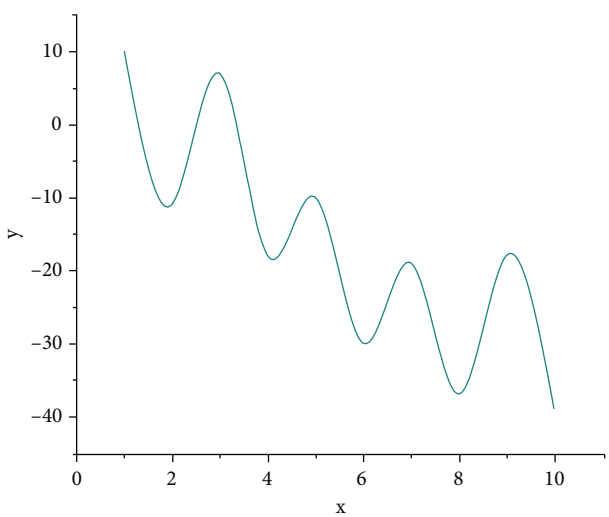


Figure 2: Simulation diagram.

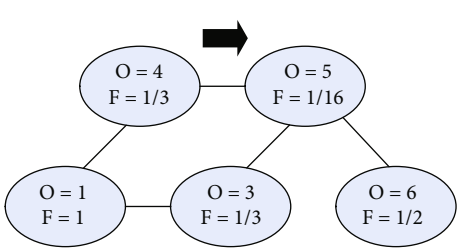


Figure 3: Network node fitness change graph.

The network in Figure 3 consists of 5 nodes and edges, where O is the objective function value of the node and F is the fitness value of the network node. The arrows on each graph point to nodes with lower network node fitness; if it is a horizontal line, the network node fitness of the top two nodes is equal.

The five genetic algorithm test functions used in this chapter are as follows.

Rosenbrock function:

$$f_1(x) = \sum_{i=1}^{30}\left[100\left(x_{i+1} - x_i^2\right)^2 + \left(x_i - 1\right)^2\right]. \tag{6}$$

Ackley function:

$$f_2(x) = -20\exp\left(-0.2\sqrt{0.5\sum_{i=1}^{2} x_i^2}\right) - \exp\left(0.5\sum_{i=1}^{2}\cos(2\pi x_i)\right) + 20 + e. \tag{7}$$

Schwefel's function:

$$f_3(x) = \sum_{i=1}^{30}|x_i| + \prod_{i=1}^{30}|x_i|. \tag{8}$$

Rastrigin function:

$$f_4(x) = \sum_{i=1}^{30}\left[x_i^2 - 10\cos\left(2\pi x_i\right) + 10\right]. \tag{9}$$

Sphere function:

$$f_5(x) = \sum_{i=1}^{30} x_i^2. \tag{10}$$

3.2. Comparison of Artificial Intelligence Genetic Algorithms. This section compares SWGA (small-world genetic algorithm) with SODNGA (self-organizing dynamic network genetic algorithm) and SGA (standard genetic algorithm) [12]. The SODNGA parameters were set as follows: population size $M = 50$, each algorithm was independently run 10 times and run for 500 generations, crossover probability $P_c = 0.7$, and mutation probability $P_m = 0.06$. The parameters of the SWGA and SGA algorithms are the same as those of the SODNGA algorithm. First calculate the Hamming distance between individuals as follows:

$$H(i,j) = \sum_{g=1}^{W}\left(1 - \beta(x_i, y_i)\right). \tag{11}$$

The population diversity is then calculated as follows:

$$D = \frac{\sum_{i=1}^{M}\sum_{j=1}^{M} H(i,j)}{M(M-1)(W/2)}. \tag{12}$$

Among them, W refers to the gene string length of the individual, and x_i and y_i represent the ith gene of individual x and y, respectively. If $x_i = y_i$, the function $\beta(x_i, y_i)$ is 1;

otherwise, it is 0, and M is the population size. The results of the population diversity experiment are shown in Figure 4.

In the early days, the population diversity of the three algorithms was similar. With the increase of evolutionary generation, it is difficult for standard genetic algorithms to maintain population diversity. Compared with the standard genetic algorithm, the small-world genetic algorithm can better maintain the population diversity. Genetic algorithms that organize dynamic networks are ideal for maintaining population diversity [13].

When the three functions take different crossover probabilities to reach the iteration termination condition, the test results of the average running time and the average running algebra of the traditional evolutionary strategy genetic algorithm and the improved evolutionary strategy genetic algorithm are shown in Table 1.

As can be seen from Table 1, for the 3 functions, when the crossover probability is taken as 1, the average running time and the average running algebra are the least for any test function. And the average running algebra and average running time gradually decrease with the increase of the crossover probability. Self-organizing topology not only has the above advantages for genetic algorithm but also can significantly improve the convergence speed and accuracy of genetic algorithm. Figure 5 depicts the experimental results of the algorithm's convergence performance, each algorithm is run independently 10 times and run for 500 generations, the population size $M = 100$, and the network node fitness is defined by formula (4).

As can be seen in Figure 5, compared with the other two algorithms, the genetic algorithm based on self-organizing dynamic network is not accurate; the standard genetic algorithm can easily perform locally optimal classification. At the same time, the convergence performance of small-world genetic algorithm is better than that of standard genetic algorithm, but the convergence speed is slow [14].

3.3. Neuromuscular Motor Testing. Two test modes with isokinetic motion at 60°/s were used: isokinetic centrifuge/centrifuge and isokinetic centrifuge/centrifuge. At weekly intervals between the two test modes, the order of the two muscle fatigue tests and the two ankle joints was randomly determined. The experiments in this study were single-blind, and subjects were required to do their best while testing all of the above indicators and were often verbally encouraged. After the test, subjects were asked to perform stretching and relaxation exercises [15]. The criterion for neuromuscular fatigue is to change the test process to the proprioceptive test mode after the ankle plantar flexion and dorsiflexion moments are lower than 50% MVC for 3 consecutive times and perform a postfatigue proprioceptive test.

In the proprioceptive test, the three actual measurements are a_1, a_2, and a_3, and the target is a. The three errors are calculated as constant error (CE):

$$\mathrm{CE} = \frac{\sum_{i=1}^{3}(a_i - a)}{3}. \tag{13}$$

It is represented by CEJPS in position sense and CEFS in force sense. In this study, the relative value (RC) of the constant error is used to express

$$\mathrm{RC} = \frac{CE}{0.25 * M}, \tag{14}$$

where M is the maximum isometric force and the standard deviation of three repeated measurements of proprioception is expressed as variable error (VE):

$$\mathrm{VE} = \sqrt{\frac{\sum_{i=1}^{3} a_i^2 - \left(\left(\left(\sum_{i=1}^{3} a_i^2\right)^2\right)/3\right)}{2}}. \tag{15}$$

It is represented by VEJPS in position sense, VEFS in force sense, and in this study by the relative value of variable error (RV):

$$\mathrm{RV} = \frac{\mathrm{VE}}{0.25 * M}. \tag{16}$$

Calculating the mean of the absolute values of the differences between the three target and test values, the absolute error (AE)

$$\mathrm{AE} = \frac{\sum_{i=1}^{3} |a_i - a|}{3}. \tag{17}$$

It is expressed as AEJPS in position sense, AEFS in force sense, and in this study the relative value of absolute error (RA):

$$\mathrm{RA} = \frac{\mathrm{AE}}{0.25 * M}. \tag{18}$$

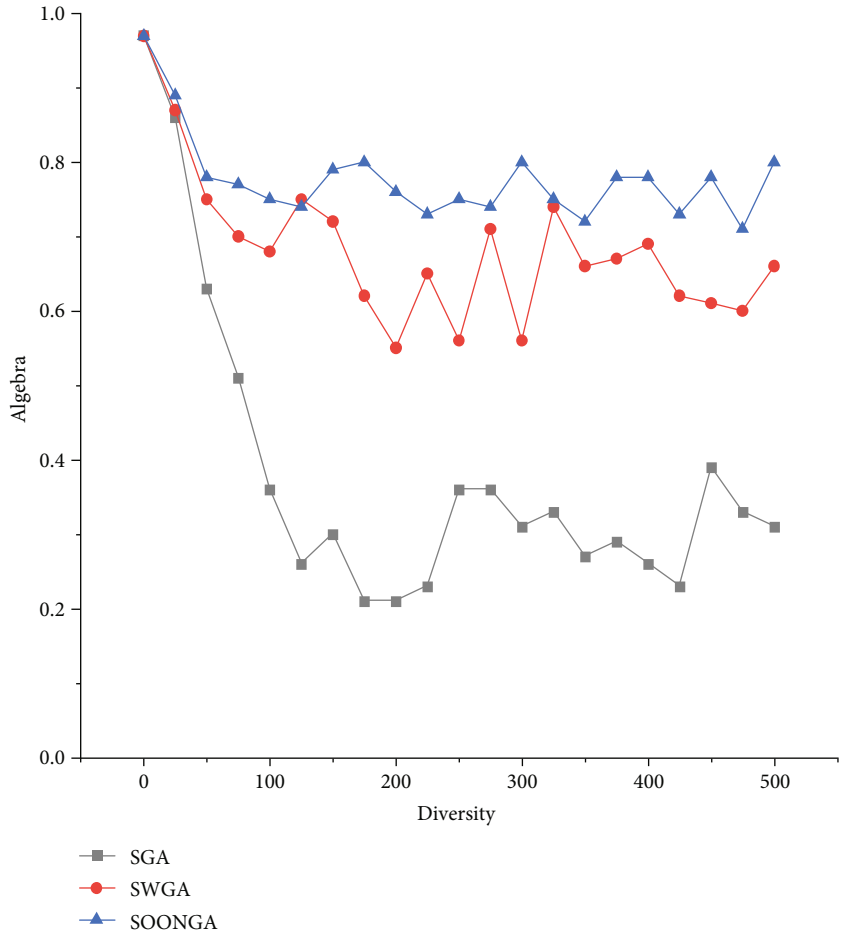
In this paper, test-and-retest reliability and standard deviation of measurements were calculated, and the intraclass correlation coefficient (ICC) was used to measure test-retest reliability.

$$\mathrm{ICC} = \frac{\mathrm{BM} - \mathrm{EM}}{\mathrm{BM} + (k-1)\mathrm{EM} + k[(\mathrm{TM} - \mathrm{EM})/n]}. \tag{19}$$

Among them, k is the number of tests, n is the sample size, TM is the mean square of the test, EM is the mean square error, and BM is the mean square between groups.

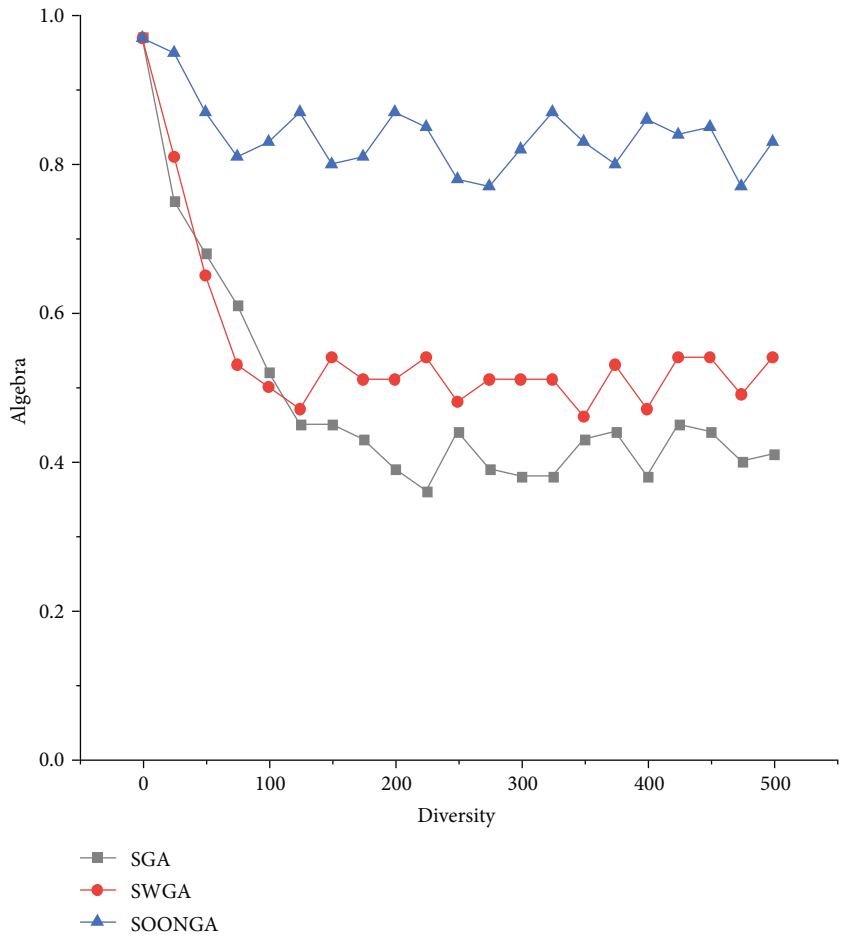
4. Anaerobic Exercise and Neuromuscular Activity Experiment and Analysis

4.1. The Kinematic Basis of Anaerobic Exercise Capacity. Exercise capacity is the concentrated expression of various functional activities of the body. The metabolism of substances and energy is the basis for the functional activities of various tissues and organs. In the practice of sports training, understanding and mastering the laws of material and energy metabolism in sports to arrange and adjust training, mastering and following the body's recovery process, and



(a) Rosenbrock

Figure 4: Continued.



(b) Rastrigin

Figure 4: Population diversity.

Table 1: Average runtime.

Function	Method	Crossover probability				
		1	0.9	0.8	0.7	0.6
SOONGA	Traditional GA	59.9	53.52	59.58	51.06	56.72
	Improved GA	6.93	7.28	7.01	7.43	7.01
SWGA	Traditional GA	15.93	13.72	15.51	15.48	16.27
	Improved GA	7.78	6.87	6.39	7.21	6.87
SGA	Traditional GA	24.94	19.22	17.01	24.89	21.91
	Improved GA	13.72	13.19	10.8	10.99	10.65

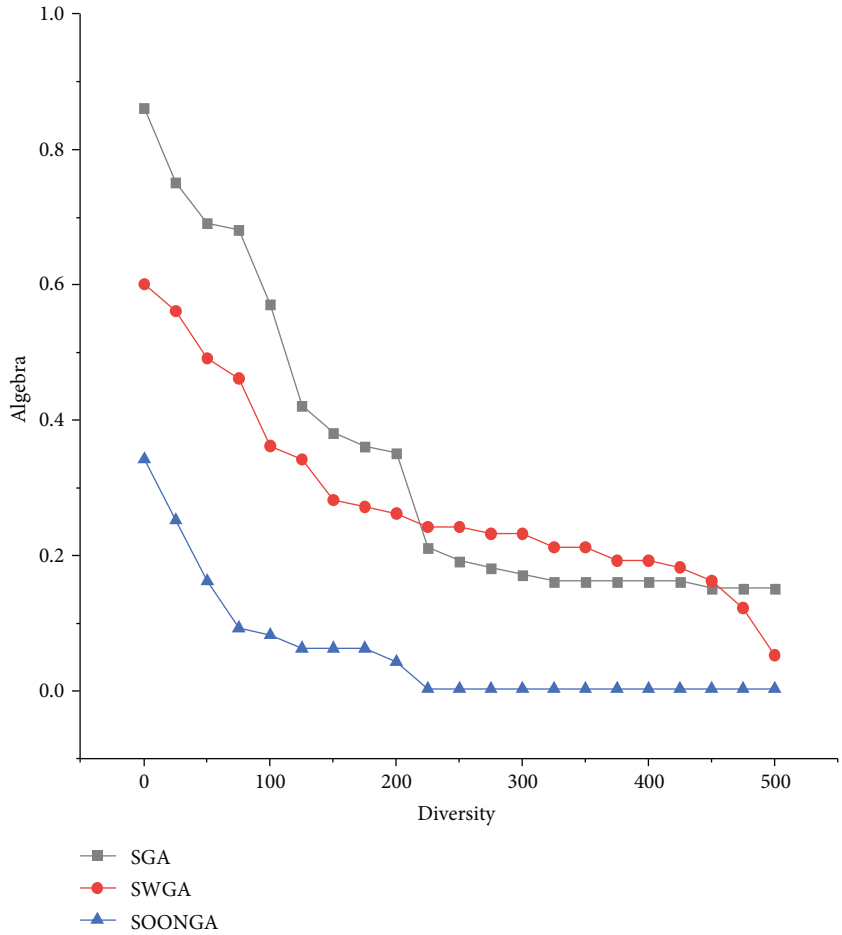
rationally adjusting nutrition are effective ways to scientifically and safely improve athletes' athletic ability [16].

The basic process of energy output during human exercise is anaerobic metabolism process and aerobic metabolism process. The relationship between metabolic processes and exercise performance conforms to the rules of anaerobic exercise. In short, different sports require different metabolic processes as the main basis for energy supply, and different metabolic capacities directly affect different sports performance.

Aerobic and anaerobic metabolism of athletes depends on three aspects. One is the storage of energy substances, that is, the amount of ATP and CP muscle glycogen in skeletal muscle. The second is the ability to regulate metabolic processes, including changes in enzyme activity during metabolic processes under the influence of training, regulation of metabolism by nerves and hormones, acid-base balance, and regulatory changes between various organs in the internal environment. Third is the metabolic capacity of the recovery process after exercise. The postexercise recovery process is not a simple adverse reaction to the consumption process during exercise. Scientific and rational arrangements for rest, nutrition, and various means of relieving fatigue can speed up the recovery process. The comprehensive performance of human skeletal muscle metabolic capacity is shown in Table 2.

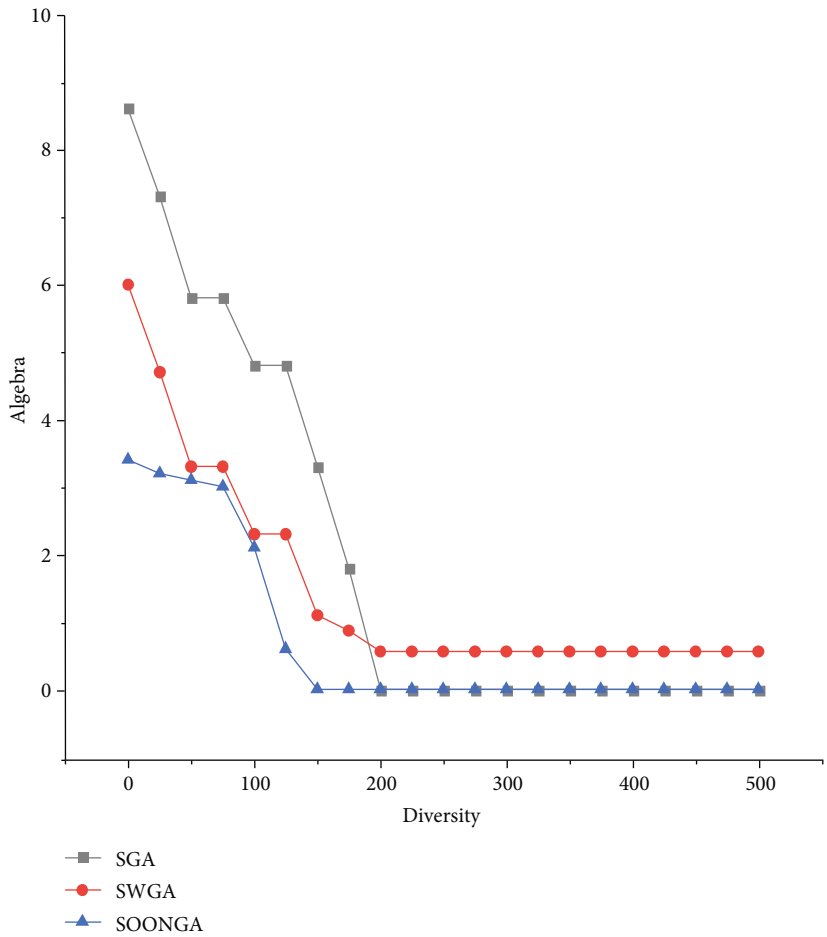
In various sports, after ATP is depleted, the body needs to restore the amount of ATP as soon as possible to maintain sports performance. Enzymes and hormones, such as creatine kinase, phosphofructokinase, and epinephrine, which are required to restore ATP to accelerate the rate of ATP resynthesis and metabolic regulation, are also very important. Table 3 shows the corresponding percentages of muscle energy requirements.

It can be seen that the energy system relies on the glycolysis system in addition to phosphate during the 200 m



(a) Schwefel's

Figure 5: Continued.



(b) Ackley

Figure 5: Convergence performance graph.

Table 2: Metabolic capacity of human skeletal muscle during exercise.

Metabolic process	Reserve	Amount of ATP synthesized	Available time for exercise	
			Extreme exercise	70%VO_{2max} intensity (min)
Anaerobic metabolism				
ATP	29	261	6 s-9 s	0.21
CP	67	198	6 s-10 s	0.05
Glycolysis system				
Muscle glycogen	425	1024	3 min-5 min	0.84
Aerobic metabolism				
Muscle glycogen	463	13502	1 h-1.5 h	0.93
Fat	65	65214	1 h-2 h	1.62

running. During the 400 m running and the 100 m swimming, the anaerobic metabolic systems of phosphate and glycolysis interact with each other. But in the 800 m and 1500 m run and most swimming events, the glycolysis and aerobic metabolic systems are clearly dominant, and the aerobic metabolic system is dominant in the 1500 m swimming and marathon running [17].

Sustained anaerobic exercise ability reflects the energy supply capacity required by the body in short-term extreme sports, is the embodiment of the body's ability to maintain extreme exercise, and reflects the ability of the body to maintain speed and maintain load intensity. The ability to maintain anaerobic exercise mainly reflects the ability of the body to complete extreme or submaximal exercise within 30-180 seconds in competitive competitions and training. The energy supply of the maintenance anaerobic exercise ability reflects a kind of maintenance ability of the body's limit energy supply and the body's ATP-CP energy supply

Table 3: Energy requirements and available quantities and rates for different distances.

Project	Energy requirements		Available energy and speed	
	Rate	Quantity	Maximum rate	Quantity
100 m run	1.84	1.48	2.44	0.67
400 m run	1.48	2.89	3.51	5.26
800 m run	1.75	3.1	3.08	1.5
1500 m run	3.12	3.25	2.75	86
Marathon	1.52	1.89	1.67	1250

ability and glycolysis energy supply ability. If the neurological anaerobic exercise ability and the primary anaerobic exercise ability are both critically restricted by the genetic factors of the exerciser, then the maintenance anaerobic exercise ability is mainly obtained through the acquired scientific training of the exerciser. Therefore, this study believes that acquired scientific exercise training has a decisive role and significance in improving the maintenance anaerobic exercise ability of exercise individuals [18].

According to the characteristics of the energy supply basis of continuous anaerobic exercise ability defined in this paper, as well as the theoretical mechanism of continuous anaerobic exercise ability and the characteristics in sports training practice, sustained anaerobic exercise capacity corresponds to extreme sports within 15 minutes. The sports items corresponding to the sustained anaerobic exercise capacity defined in this study include 3000-meter running, 5000-meter running, 5000-meter speed skating, 10,000-meter speed skating, 400-meter swimming, 800 swimming, middle distance cycling, and middle and long distance racing rowing. In the past sports training theory, there are different opinions on the sports ability corresponding to these sports items. In the past, it was believed that these sports items lasted for a long time, and the dominant sports ability of the sports individual should be aerobic exercise ability. But the characteristics of these sports today in competitive competitions are obviously contrary to aerobic fitness. Because with the increasingly fierce competition in modern competitive sports, the level is getting higher and higher, and the ability of sports individuals is getting stronger and stronger; in order to win the competition in these competitive sports, the sports body is required to exercise close to the extreme intensity from the beginning of the competition. And it is required that the sports body must be able to maintain extreme sports, high-speed, high-intensity, and high-load sports for a long time in order to win the game [19].

4.2. Mechanisms of Neuromuscular Regulation. FCMC is considered a classic and commonly used method for assessing functional connectivity between neural signals and associated somatic muscles. FCMC analysis was initially implemented based on magnetoencephalography (MEG) and EMG. It was subsequently generalized to local field potential (LFP) electroencephalography (EEG), electrocorticography (ECoG), and surface electromyography (sEMG) and validated in multiple methods and species. This paper takes finger strength as a practical example to study the neuromuscular regulation mechanism [20].

Table 4: CMC peak mean and standard deviation in alpha band at different strength levels.

Strength level	CMC peak mean ± standard deviation
I 10% MVC	2.067 ± 0.106
I 20% MVC	2.246 ± 0.113
I 30% MVC	2.282 ± 0.103
I 40% MVC	2.008 ± 0.109
I 50% MVC	2.239 ± 0.108

To explore the effect of finger strength level on CMC, one-way ANOVA was performed on 50 groups of coherence data at two strength levels, I 20% MVC and I 40% MVC. One-way ANOVA and T-test were performed on the CMC in the alpha band under the two strength levels of I 20% MVC and I 40% MVC, and it was found that the peak CMC in this band was significantly different at the two strength levels. The results are shown in Table 4.

At the I 20% MVC strength level, the CMC of the alpha and beta bands were compared and found to be significantly different ($P < 0.05$) between the two different bands. Similarly, the CMC in the alpha and beta bands at I 40% MVC strength level was compared, and the results were also significantly different ($P < 0.05$).

In order to make an intuitive comparison of the effect of the finger strength level on the alpha and beta frequency bands, a box plot of the obtained results is made, as shown in Figure 6.

As can be seen from Figure 6, both α and β band CMCs increased with increasing intensity levels. In addition, at the same strength level, the CMC of the beta band is significantly higher than that of the alpha band, with an average of 0.231.

There are three explanations for the current mechanism of the effect of increased strength levels on beta-band CMC. (1) At lower strength levels, in order to make movements with less error, a large number of corrective movements are involved, thereby reducing the CMC at lower strength levels. (2) The tuning effect of the firing rate of the motor unit to the beta band enhances the CMC at greater force levels. (3) The increase in CMC in the beta band at higher force levels compared to lower force levels indicates stronger binding between cortex and motor neurons. It is generally accepted that CMC reflects the direct connection between the brain and muscles, which is thought to be related to force control. In addition, some studies have found that the oscillatory activity in the lower beta frequency band and multiple cortical areas combine to form a network structure in visual guidance tasks, thereby effectively promoting neurophysiological processes [21].

4.3. Muscle Neuron Population Model. The essence of FCMC is the transmission of information in the corticospinal pathway between the primary motor cortex and the muscles, as shown in Figure 7.

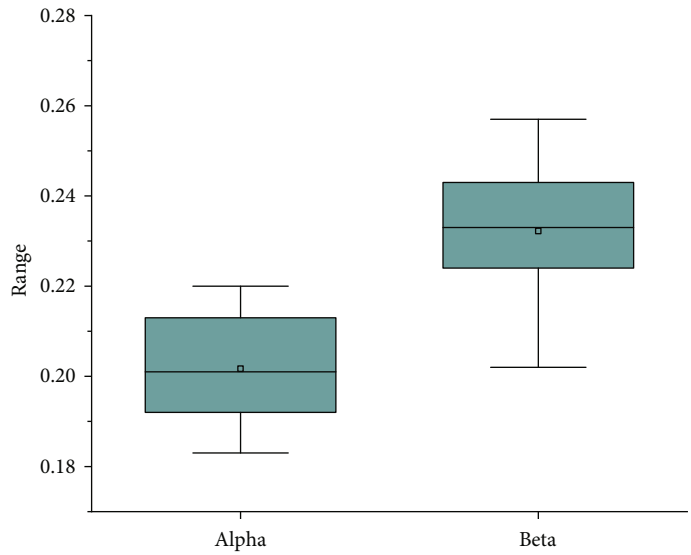


Figure 6: Box plots of alpha and beta band CMCs at different strength levels.

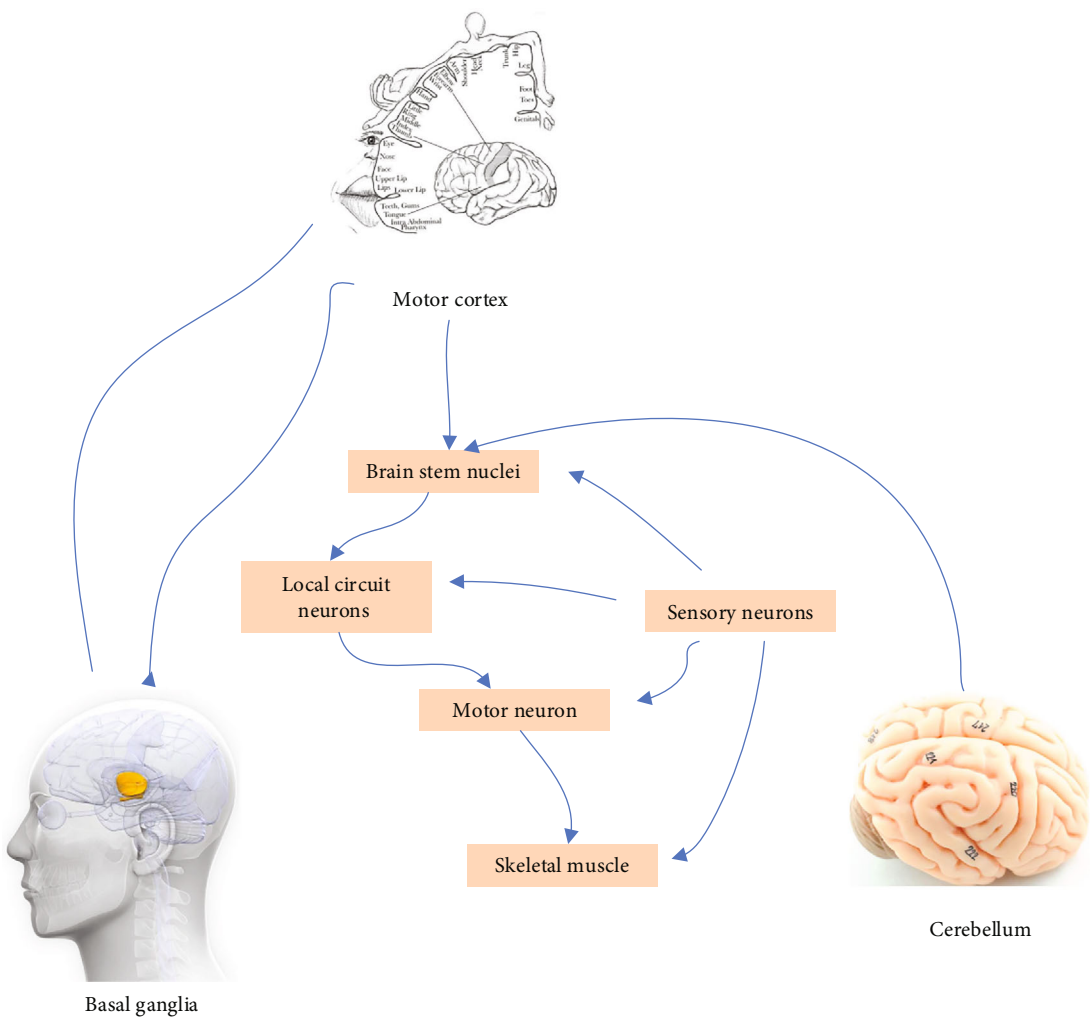


Figure 7: Neuromuscular information transmission pathway.

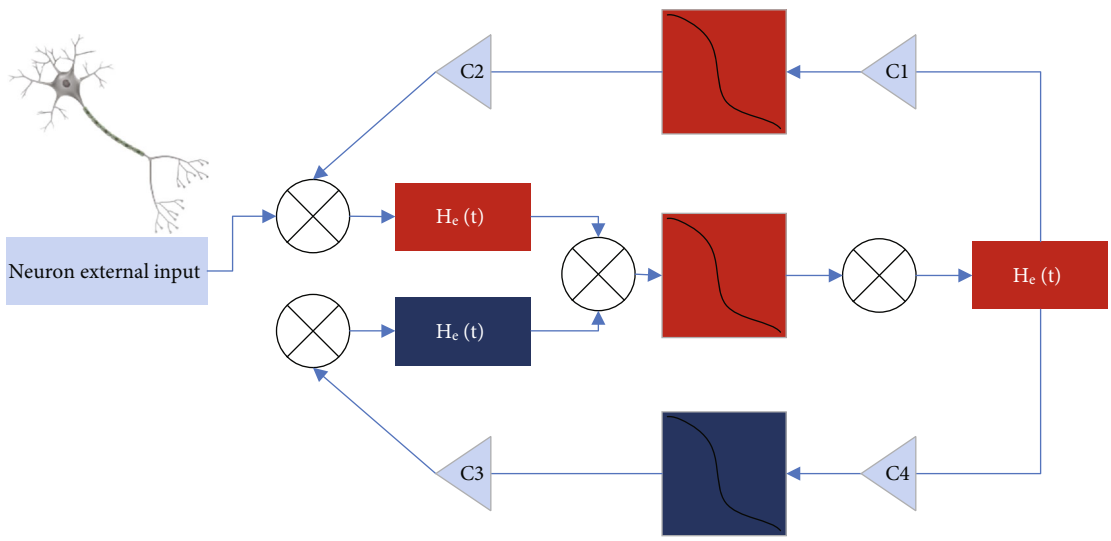


Figure 8: Schematic and equations of JR-NMM.

FCMC studies can reflect the interaction between the brain and muscles. It represents information transfer within the neuromuscular system and is related to the transmission of commands from the brain to the muscles and the upward feedback of muscle contractions [22]. As such, it contributes to understanding the mechanisms by which the brain controls muscle activity, the impact of muscle movement on brain function, and the underlying causes of certain physiological and pathological conditions.

The Jansen-Rit neuron population model (JR-NMM) can model the activity of neuron populations aggregated in local neural circuits by simulating pyramidal neuron population activity. This pyramidal neuron can receive both inhibitory and excitatory feedback from local neurons as well as excitatory input from nearby or distant neuronal populations. Figure 8 shows a schematic diagram of the overall structure of the JR-NMM used to simulate UER in this study.

JR-NMM contains three neuronal population structures, namely, pyramidal neuron population, excitatory interneuron population, and inhibitory interneuron population. They are connected by excitatory synaptic impulses in response to $\mathrm{He}(t)$ and inhibitory synaptic impulses in response to $\mathrm{Hi}(t)$ [23].

In order to confirm that the passive membrane time constant of JR-NMM can be used as an identification parameter to simulate the rhythmic characteristics of UER, in this study, different brain rhythms were simulated by adjusting the passive membrane time constant, and the relationship between the passive membrane time constant and the frequency band distribution characteristics of JR-NMM (from alpha band to ripple band) was analyzed, and Figure 9 is obtained.

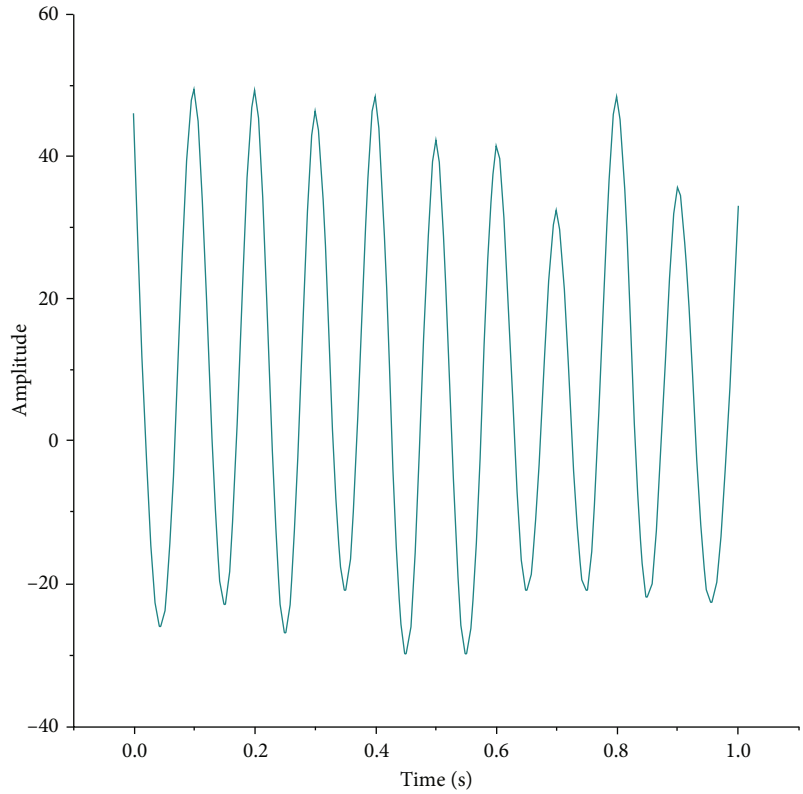
Figure 9(a) shows the simulated brain rhythm waveforms in the alpha frequency band. The results show that JR-NMM can simulate brain rhythm waveforms in different frequency bands by adjusting the excitatory and inhibitory passive membrane time constants. Figure 9(b) shows the distribution relationship between passive membrane time constants and simulated brain rhythms in JR-NMM. Among them, the x-axis is the excitatory passive membrane time constant, the y-axis is the inhibitory passive membrane time constant, and the color scheme represents the frequency peaks of the simulated brain rhythm. The results show that the frequency peaks of the simulated brain rhythms increase with increasing passive membrane time constants. Therefore, the passive membrane time constant in JR-NMM can be used as an identification parameter to simulate the brain rhythm characteristics of UER. Moreover, according to the results of isolated neurons of ultrasound neuromodulation, ultrasound-regulated neuronal firing is caused by the change of the passive membrane time constant under the action of ultrasound radiation force [24]. Therefore, in this study, the passive membrane time constant was used as a "biological effector" for the neuronal firing activity induced by ultrasonic radiation force.

First, this paper uses the short-time Fourier transform (STFT) method to estimate the time-frequency characteristics of the LFP signal and the EMG signal in the frequency range of 5~200 Hz under a single ultrasound stimulation, as shown in Figure 10.

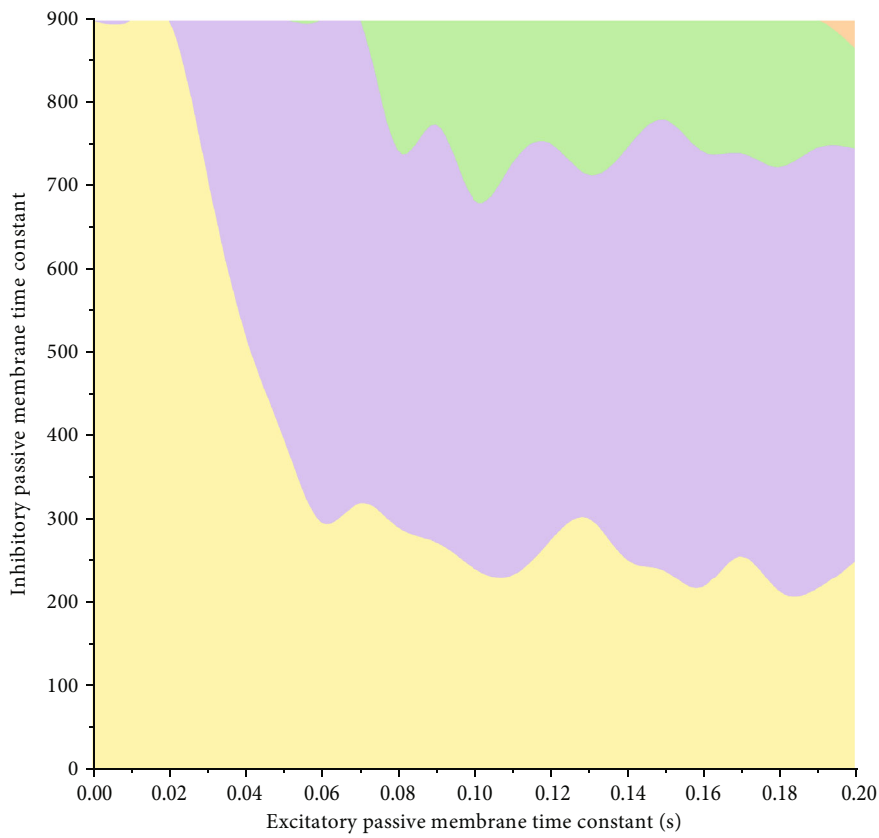
Figure 10 shows the waveforms and time spectra of the LFP and EMG signals for the next stimulation cycle with a duration of 4 s in the next stimulation cycle at this ultrasound parameter (0.20 W/cm^2, 50% DC, 400 ms). Among them, 0~2 s of the x-axis is the nonstimulation period. In the waveform diagrams of the LFP and EMG signals, it can be clearly observed that the signal amplitude during the stimulation period is higher than that in the nonstimulation period.

5. Discussion

There are still shortcomings in this paper, and it is recommended that researchers supplement the data later and increase the sample size. The rhythmic keystrokes discussed

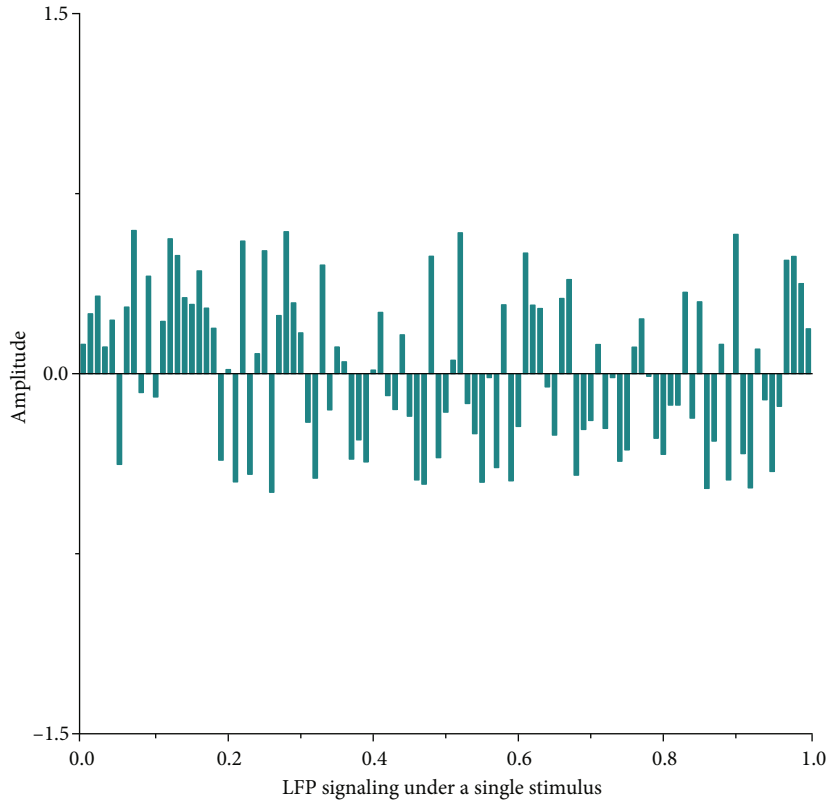


(a) Simulated membrane rhythm

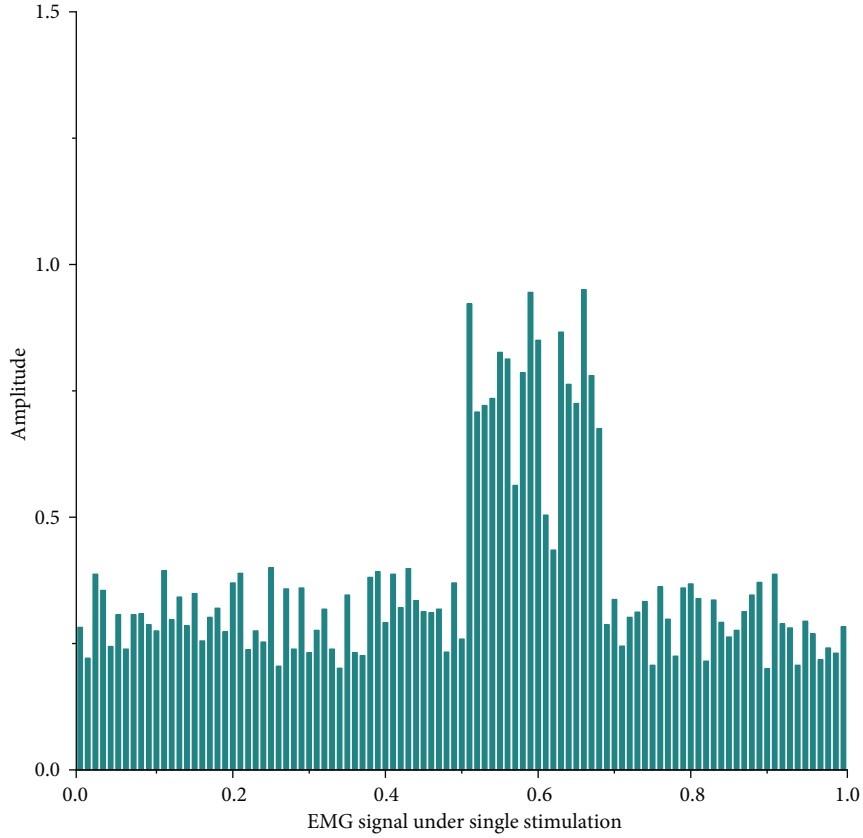


(b) Passive membrane time constant

Figure 9: Brain rhythms simulated by tuning the passive membrane time constant of JR-NMM.



(a) LFP



(b) ENG

Figure 10: LFP and EMG signals and their corresponding time-frequency plots under a single stimulus.

in this work only took the data of one healthy subject, and the MEG, SEMG, and finger strength signals of five healthy subjects were collected in the visual-guided finger force tracking task. However, relatively speaking, the sample size is still small. In order to avoid errors caused by insufficient samples in the statistical analysis of the characteristic information of cortical neural activity, the sample size needs to be increased in the future. The cause of exercise-induced fatigue is unknown, and the hydrogen ion theory is one of the generally accepted theories. Due to insufficient oxygen supply in the body during high-intensity exercise, skeletal muscle cells need to perform contraction through glycolysis, but at the same time as producing ATP, they also produce a large amount of lactic acid. In a short period of time, the body's own buffering system becomes difficult, and the lactic acid produced by the body is completely buffered, resulting in the accumulation of lactic acid and an increase in the concentration of hydrogen ions.

6. Conclusions

Based on the results of expert interviews and the basis of biology and sports training, combined with the characteristics of finger movement, this paper mainly studies three types of anaerobic exercise capacity, maintenance anaerobic exercise ability, and continuous anaerobic exercise ability. Different types of anaerobic fitness have their own basic characteristics and practical significance, and there are also different training characteristics in competitive sports practice, such as training content characteristics, methods, and basic requirements. In this paper, we investigated the effects of weight-bearing loading and vibrational stimulation on the activation of the subjects' finger muscles by analyzing electromyography and identified the vibrational frequency that caused maximal muscle activation by different frequencies of vibrational stimulation. In this paper, a specific vibration training program was developed, and through an 8-week training intervention, the effects on maximal strength, rapid strength, H-reflex, and T-reflex of the subjects' finger muscles were observed, and the differences were compared. This article analyzes the neuromuscular effects of weight-supported vibrational strength training, modulates properties, and explores its possible mechanisms. This study improves the theoretical system of weight-bearing vibration training and provides a reference for the public and athletes to formulate vibration strength training programs.

Data Availability

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Conflicts of Interest

The authors state that this article has no conflict of interest.

References

[1] Y. Tsujimura, M. Gen, and E. Kubota, "Solving job-shop scheduling problem with fuzzy processing time using genetic algorithm," *Journal of Japan Society for Fuzzy Theory & Systems*, vol. 7, no. 5, pp. 1073–1083, 1995.

[2] L. Qiang and C. Wu, "A hybrid method combining genetic algorithm and Hooke-Jeeves method for constrained global optimization," *Journal of Industrial & Management Optimization*, vol. 10, no. 4, pp. 1279–1296, 2017.

[3] Y. Yoshitomi, H. Ikenoue, T. Takeba, and S. Tomita, "Genetic algorithm in uncertain environments for solving stochastic programming problem," *Journal of the Operations Research Society of Japan*, vol. 43, no. 2, pp. 266–290, 2000.

[4] S. M. Tayebi, A. A. Mahmoudi, E. Shirazi, and M. Sangi, "Acute response of some iron indices of young elite wrestlers to three types of aerobic, anaerobic, and wrestling exercise," *Montenegrin Journal of Sports Science & Medicine*, vol. 6, no. 1, pp. 5–11, 2017.

[5] J. Kang, J. Park, and J. A. Johnson, "Comparison of isokinetic muscle function and anaerobic exercise capacity in the knee according to Kukki taekwondo training type," *Physical Activity Review*, vol. 9, no. 2, pp. 40–55, 2021.

[6] M. H. Park, M. S. Kang, and D. S. Choi, "Comparison analysis of isokinetic muscle function, aerobic and anaerobic exercise ability, and basic physical fitness according to the gender of middle school taekwondo athletes," *Journal of The Korean Society of Living Environmental System*, vol. 28, no. 3, pp. 261–270, 2021.

[7] C. Morana, S. Ramdani, S. Perrey, and A. Varray, "Recurrence quantification analysis of surface electromyographic signal: sensitivity to potentiation and neuromuscular fatigue," *Journal of Neuroscience Methods*, vol. 177, no. 1, pp. 73–79, 2009.

[8] R. Tavakkoli-Moghaddam, J. Safari, and F. Sassani, "Reliability optimization of series-parallel systems with a choice of redundancy strategies using a genetic algorithm," *Reliability Engineering & System Safety*, vol. 93, no. 4, pp. 550–556, 2008.

[9] H. Dawid and M. Kopel, "On economic applications of the genetic algorithm: a model of the cobweb type," *Journal of Evolutionary Economics*, vol. 8, no. 3, pp. 297–315, 1998.

[10] Z. Arabasadi, R. Alizadehsani, M. Roshanzamir, H. Moosaei, and A. A. Yarifard, "Computer aided decision making for heart disease detection using hybrid neural network-genetic algorithm," *Computer Methods and Programs in Biomedicine*, vol. 141, pp. 19–26, 2017.

[11] I. P. Panapakidis and A. S. Dagoumas, "Day-ahead natural gas demand forecasting based on the combination of wavelet transform and ANFIS/genetic algorithm/neural network model," *Energy*, vol. 118, no. JAN.1, pp. 231–245, 2017.

[12] M. Keshtkar, "Energy, exergy analysis and optimization by a genetic algorithm of a system based on a solar absorption chiller with a cylindrical PCM and nano-fluid," *International Journal of Heat and Technology*, vol. 35, no. 2, pp. 416–420, 2017.

[13] M. Nemati, M. Braun, and S. Tenbohlen, "Optimization of unit commitment and economic dispatch in microgrids based on genetic algorithm and mixed integer linear programming," *Applied Energy*, vol. 210, pp. 944–963, 2018.

[14] Y. Ma, G. Luo, X. Zeng, and A. Chen, "Genetic algorithm-based transfer learning for cross-company software defect prediction," *International Journal of Hybrid Information Technology*, vol. 10, no. 3, pp. 45–56, 2017.

[15] R. Johnston, M. Crowe, and K. Doma, "Effect of nicotine on repeated bouts of anaerobic exercise in nicotine naïve individuals," *European Journal of Applied Physiology*, vol. 118, no. 4, pp. 681–689, 2018.

[16] H. Yan, N. Merchant, J. Champagne, and M. Acloque, "Racial differences in cardiovascular responses following acute bouts of anaerobic exercise," *Medicine & Science in Sports & Exercise*, vol. 52, no. 7S, pp. 706–706, 2020.

[17] M. Marquart, E. Huffman, C. Zelonis, S. Brown, and J. Sanders, "The effects of aerobic vs anaerobic exercise on cognitive function in college aged individuals," *International Journal of Exercise Science: Conference Proceedings*, vol. 9, no. 6, pp. 78–78, 2018.

[18] Y. Yang and J. Yang, "Research on athlete's aerobic and anaerobic exercise ability with balanced nutrition," *Revista Brasileira de Medicina do Esporte*, vol. 27, no. 4, pp. 430–433, 2021.

[19] T. H. Chou and C. Hurr, "Effects of acute alcohol consumption on cycling anaerobic exercise performance: a randomized crossover study," *Exercise Science*, vol. 29, no. 3, pp. 264–271, 2020.

[20] S. M. Arent, J. Mckenna, and D. L. Golem, "Effects of a neuromuscular dentistry-designed mouthguard on muscular endurance and anaerobic power," *Comparative Exercise Physiology*, vol. 7, no. 2, pp. 73–79, 2010.

[21] C. Guervilly, M. Bisbal, J. M. Forel et al., "Effects of neuromuscular blockers on transpulmonary pressures in moderate to severe acute respiratory distress syndrome," *Intensive Care Medicine*, vol. 43, no. 3, pp. 408–418, 2017.

[22] J. G. Claudino, J. Cronin, B. Mezêncio et al., "The countermovement jump to monitor neuromuscular status: a meta-analysis," *Journal of Science and Medicine in Sport*, vol. 20, no. 4, pp. 397–402, 2017.

[23] M. I. Black, A. M. Jones, J. R. Blackwell et al., "Muscle metabolic and neuromuscular determinants of fatigue during cycling in different exercise intensity domains," *Journal of Applied Physiology*, vol. 122, no. 3, pp. 446–459, 2017.

[24] V. Bernecke, K. Pukenas, L. Daniuseviciute et al., "Sex-specific reliability and multidimensional stability of responses to tests assessing neuromuscular function," *HOMO-Journal of Comparative Human Biology*, vol. 68, no. 6, pp. 452–464, 2017.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 2816848, 9 pages
https://doi.org/10.1155/2022/2816848

Research Article

Study on the Application of Digital Information Technology in Music Teaching

Hao Cheng, Hao-ze Zhong, and Ke-cheng Ben

Nanchang University, Nanchang 330031, China

Correspondence should be addressed to Hao Cheng; chenghao82@ncu.edu.cn

Received 18 January 2022; Revised 26 February 2022; Accepted 3 March 2022; Published 22 March 2022

Academic Editor: Hoon Ko

In order to deal with the great challenge of the arrival of information society to the traditional music teaching concept and practice, an overview of the teaching ability structure of music normal students supported by information technology is put forward. The changes in the structure of teaching ability caused by information technology are analyzed, that is, to analyze the cultivation of teaching ability of music normal students by integrating music subject knowledge with information technology platform and promoting relevant elements in combination with teaching activities. By using Fitzgerald and our proposed long +short method to separate Beach Boys songs at the singing music ratio of -6, 0, and +6 dB, the singing signal (left column) and music signal (right column) are relative to the average values of SDR (first row), SIR (second row), and SAR (third row) of pure singing and pure music. Experiments show the effectiveness of this method.

1. Introduction

Music normal students are future teachers. They are a group that will face the task of "preaching, teaching, and solving doubts." The connection between the level of music teaching in the future and its teaching methods lies in whether they can well adapt to the development of the information society, skills, and social development to some extent, as shown in Figure 1 [1]. Secondly, as teachers of tomorrow, whether music normal students can acquire tomorrow's teaching ability and skills in today's teaching activities depends, to some extent, on the display of teaching ability and the application of teaching skills in teachers' teaching activities. In this sense, in the information age, music normal students are not only the passive bearers of information technology, but also the active communicators of information technology, so that music normal students can acquire and master solid teaching ability, so as to adapt to the mission of preaching, teaching, and puzzle-solving of tomorrow's teachers [2]. However, in the traditional teaching mode, in order to achieve standardized and efficient teaching effect faster and better, teachers often use one-way indoctrination to carry out teaching activities. Although this method improves the teaching efficiency, it cannot better promote the effective transfer and utilization of normal students' knowledge and skills because it cannot provide diversified practical situations, which hinders the professional development of music teachers. More importantly, in the current training process of music normal students, professional knowledge, subject teaching knowledge, and educational technology ability, especially educational information technology knowledge, are separated from teaching as a whole [3]. The professional knowledge of music normal students in the college is mainly taught by the professional teachers of the college, the subject teaching knowledge is mainly taught by the teachers of relevant colleges such as the college of education and the college of psychology, and the courses related to education information technology are taught by the teachers of the school of Education Information Technology. The teacher education college conducts skill training and testing of normal students, and the education administrative department uniformly arranges the practice of normal students. This training mode leads to the inability of ordinary music students to integrate professional knowledge, subject teaching knowledge, and educational technology well. Ability in the training process, let alone scientization and systematization

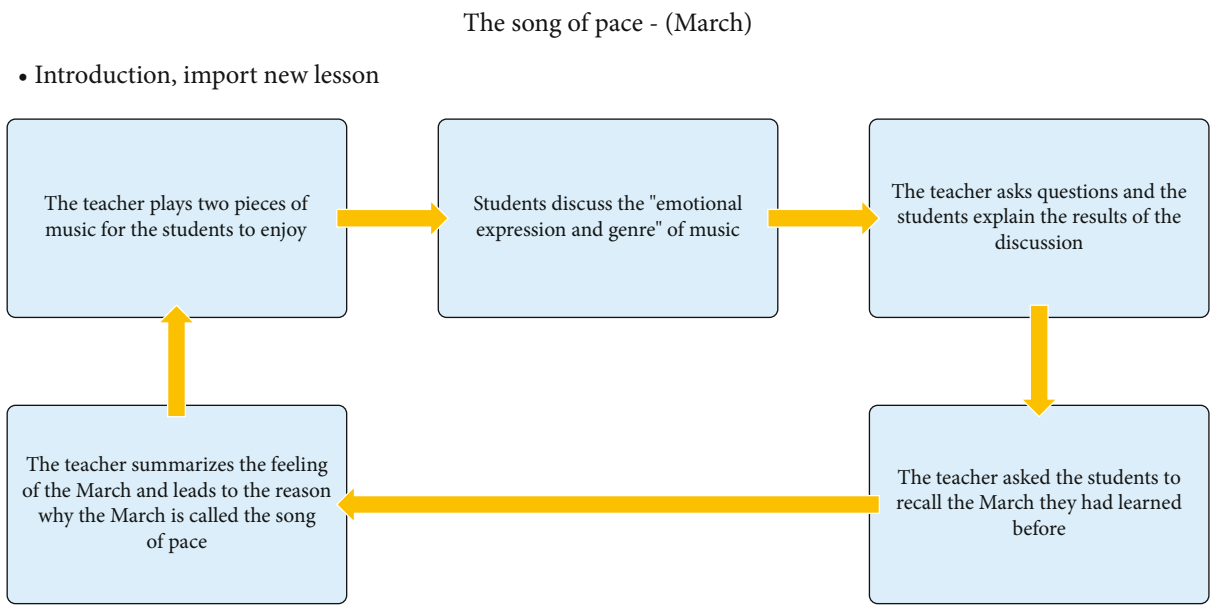


Figure 1: Application process of digital information technology in music teaching.

of music normal students [4]. In order to promote the innovative development of music normal students' education, it is necessary to further improve the teaching ability of music normal students in combination with the information environment.

2. Literature Review

By 2020, basically build an education informatization system covering all kinds of schools at all levels in urban and rural areas and promote the modernization of educational content, teaching means, and methods" [5]. Hamond et al. found that with the wide rise of new generation information technologies such as Internet, cloud computing, Internet of Things, big data, and VR technology, colleges and universities at all levels have accelerated the construction of educational informatization and gradually moved the construction direction from traditional digital campus to a higher stage of intelligent campus according to the actual teaching and management needs. It has become the theme and trend of current information development [6]. Li believes that the intelligent campus, with the people-oriented concept as the carrier and many application service systems as the media, fully integrates education and teaching, school administration management, teachers' scientific research, and the campus life of teachers and students, which is a high-level form of the development of information-based campus after digital campus [7]. Ross found that the term "smart campus" was proposed on the basis of digital campus construction in the process of educational informatization, which originated from the concept of "smart earth" proposed by IBM in 2008 [8]. Li et al. believe that the so-called "wisdom" not only represents its traditional meaning but also represents a new concept and management mode in the information age, with the extension of digitization, networking, big data, and intelligent interaction. Cerf found that the "smart earth" strategy has been generally approved by all countries, gradually penetrated into various fields related to informatization, and gave birth to many new concepts, including the term smart campus. Different Chinese scholars have interpreted the meaning of smart campus [9]. Robert believes that smart campus is an intelligent campus work, learning, and life integrated environment based on the Internet of Things. This integrated environment takes various application service systems as the carrier to fully integrate teaching, scientific research, management, and campus life [10]. Kazmina believes that smart campus is composed of smart service concept, smart environment, smart application and service (smart teaching, smart management, and smart scientific research), and smart culture and experience and expounds the theoretical basis, methods, and principles of smart design [11]. Habib et al. constructed the "understanding schema of smart education," and believed that smart campus, like smart classroom and smart terminal, is a learning space for smart education according to different scales and is an integral part of smart education, serving for wisdom education [12]. Paletta et al. proposed that smart campus refers to an open education teaching environment and a convenient and comfortable living environment with the concept of personalized service for teachers and students, which can fully perceive the physical environment, identify learners' individual characteristics and learning situations, provide seamless network communication, and effectively support teaching process analysis, evaluation, and intelligent decision-making [13]. Wang et al. believe that the smart campus is based on the Internet under the information background, makes full use of the correlation characteristics of various information, changes the interaction mode between teachers and students and between students and school through information technology, and integrates the school's teaching, scientific research, management and digital resources, and application systems,

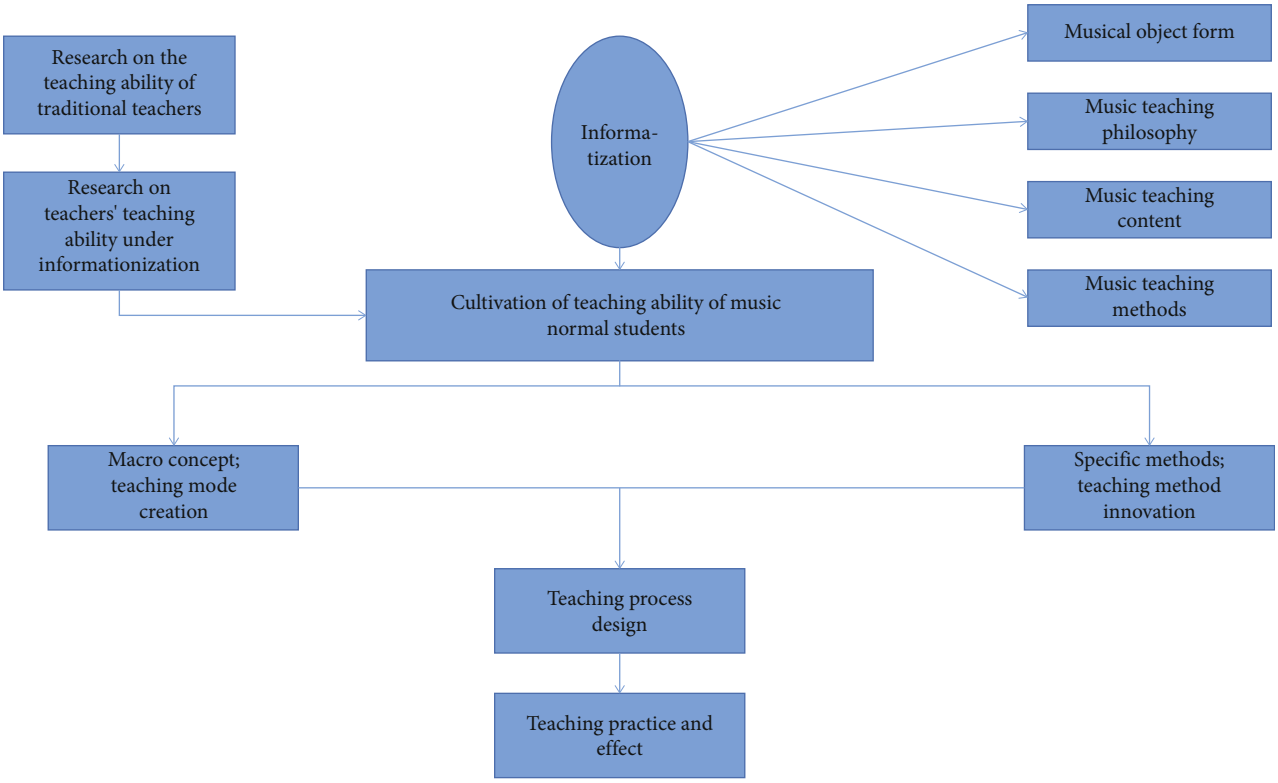


Figure 2: General diagram of research ideas.

so as to realize data sharing and instant visibility, communicate with the outside world, and provide support for teaching decision-making and analysis, to fully realize educational informatization, as shown in Figure 2 [14].

3. Method

Teaching process is a unity of students' learning process and teachers' teaching process. For quite a long time, the teaching process is basically composed of three basic elements: teachers, students, and teaching content. Teachers transmit information (content) through digital information technology, and students receive information (content) through digital information technology. Information technology has become the main way to present music teaching content [15]. The operability and applicability of information technology determine the clarity and structure of the music teaching content it transmits. At the same time, information technology itself has become an indispensable part of the music education content system, as shown in Figure 3.

Teacher's ability is the ability that teachers should have to complete certain teaching activities. The analysis here is the requirement for all professional teachers, and for specific disciplines, the teaching ability they need to master is focused [16]. On the basis of teachers' ability, this paper focuses on the ability that a music teacher should have. The training goal of music normal students is the future music teachers in primary and secondary schools. Clarifying the ability required for future teaching is of great significance to promote the training of music normal students [17]. By analyzing the teaching ability structure of middle school music teachers, it is considered that the teaching ability of middle school music teachers and music normal students is mainly reflected in four aspects: music teaching cognitive ability, music teaching operation ability, music teaching monitoring ability, and music teaching evaluation ability, as shown in Table 1.

In the information age, information literacy has become one of the important indicators for the comprehensive evaluation of talents in today's society and the basic survival ability of individuals to adapt to social life. As future teachers, the information literacy of normal students directly affects the development of educational informatization and the cultivation of high-quality and compound innovative talents. It is clearly pointed out that "adhere to emancipating the mind, adjust measures to local conditions, develop and innovate, keep pace with the times, pay attention to application, cultivate new primary and secondary school teachers with innovative spirit and practical ability, and comprehensively improve the information literacy of primary and secondary school teachers" [18]. Therefore, facing the torrent of educational informatization development, teachers must bear the brunt of comprehensively improving their information literacy, which is not only the basic path of teachers' self-lifelong learning and professional sustainable development, but also an important guarantee for the smooth realization of information education. At the same time, this is also determined by the future career of normal students [19]. This higher requirement is reflected as follows: for future teachers, normal students are the recipients,

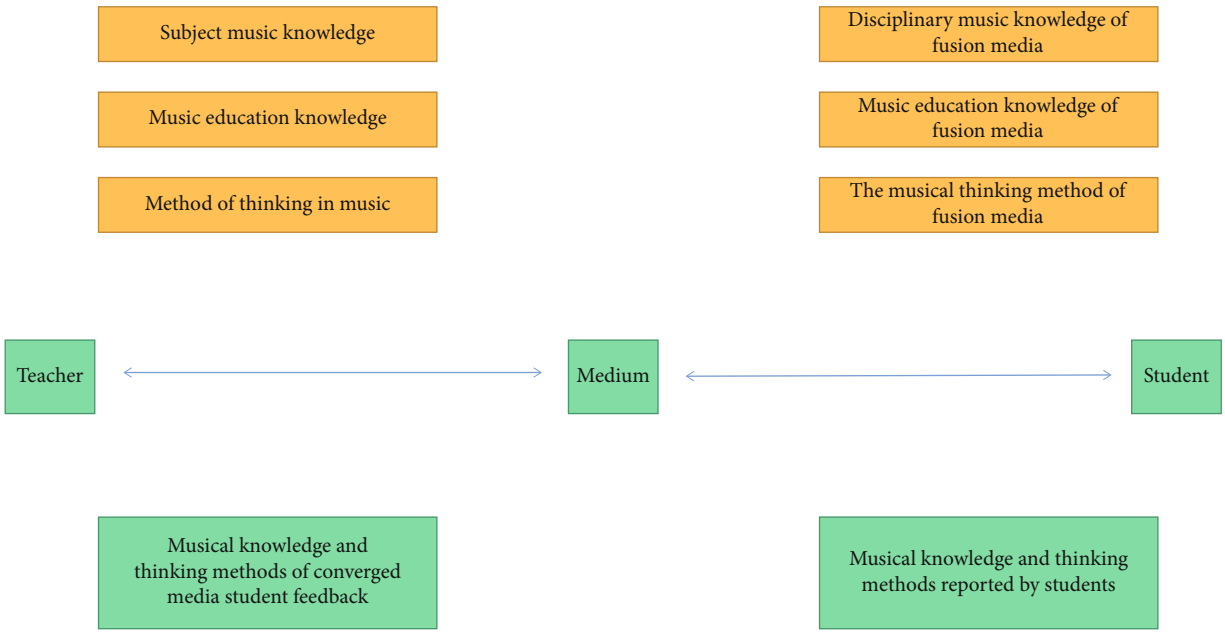


Figure 3: Music teaching content integrating information technology.

Table 1: Ability structure of middle school music teachers.

Music teaching ability	
Primary ability	Secondary ability
Music teaching cognitive ability	Ability to analyze and master curriculum standards
	Ability to understand teaching objects
	Ability to analyze and deal with teaching materials
	Teaching design ability
Music teaching operation ability	Ability to formulate (determine) teaching objectives, prepare teaching plans, arrange teaching activities, and select teaching methods and strategies
Music teaching monitoring ability	Plan and prepare music teaching activities, check, evaluate, and feed back the implementation of music teaching activities, and consciously control and adjust teaching activities
Music teaching evaluation ability	Ability to select and prepare evaluation tools, implement evaluation, and obtain feedback information

gatekeepers, and disseminators of information. As the disseminator of knowledge, the information used and disseminated by normal students must be scientific and correct. As a manufacturer of information, its information processing ability, information generation, and expression ability should have higher requirements. As a practitioner of educational informatization, he should not only have the ability to use information technology to construct effective teaching information resources, but also have the ability to use information technology for teaching design and the methods and strategies of integrating information technology in specific subject courses.

In addition, with the in-depth development of the integration of information technology and music curriculum, the information environment looks forward to the comprehensive improvement of the information literacy of music normal students [20]. It actually appears in the following three aspects: first, the awareness of the importance of information technology and information resources needs to be further strengthened. A survey of the conservatory of music of a normal university directly under the Ministry of Education is shown in Table 2. Most music normal students feel that information technology is important to them, accounting for 92% of the respondents. However, the very important proportion instrument is 39%, which is relatively low. In addition, 1% of people think that the importance of information technology is general. Although the proportion is small, there is still room for music normal students to fully adapt to the trend of the information age. In the cognition of the importance of information resources, music normal students have a higher degree of cognition, and more than 50% hold the concept of "very important." As young students, there are subjective and objective reasons for their high awareness of information technology and information resources. However, we should not ignore the problem of improving the cognitive level of some students who are still hesitant and

Table 2: Cognition of music normal students on the importance of information technology and information resources (unit: %).

Type	Very important	Importance	General	Total
Importance of information technology	39	53	7	100
Importance of information resources	53	46	0	100

Table 3: Music normal students' understanding of information knowledge (unit: %).

Type	Deep understanding	General	Lack of understanding	Total
Multimedia knowledge	84	15	0	100
Basic principles of network	67	32	0	100
Computer hardware and system maintenance	28	54	7.14	100
Network security and computer viruses	32	60	7.14	100

Table 4: Information technology utilization ability of music normal students (unit: %).

Type	Often	Occasionally	No searching	Total
Take the initiative to find information of interest	67	32	0	100
Use the Internet to find course-related materials	32	60	7	100
Use the Internet for preclass learning	46	53	0	100

lack of understanding. The relative amount of these students may be very small, but the absolute amount is still very large [21].

Secondly, the cognition of the importance of information technology and resources is only the first level. At present, the lack of information knowledge of music normal students is a more prominent problem. As shown in Table 3, multimedia teaching is an essential teaching skill in the information environment. The survey shows that 18% of music normal students still do not understand multimedia knowledge, and even fewer music normal students have a deep understanding of the basic principles of network, computer hardware and system maintenance, network security, and computer viruses. Understanding this information knowledge is the premise of using information technology ability. The lack of understanding will inevitably affect the exertion of music normal students' information skills.

Thirdly, the ability of music normal students to use information technology needs to be improved. As shown in Table 4, among the music normal students surveyed, two-thirds often take the initiative to find information of interest, which accounts for a high proportion, but more than 30% of the music normal students only occasionally, which is still insufficient. For whether to use the Internet to find course-related materials, the proportion is lower, the proportion of frequent use of Internet search is less than 50%, and a small number of students will not use the Internet for preclass learning.

In traditional teaching activities, the unequal authority between teachers and students formed by focusing on the teaching of knowledge is conducive to the mastery of knowledge to some extent, but it is not conducive to the cultivation of students' ability. The advent of information society has created conditions for teaching reform. It brings the integration of teaching design ability, teaching implementation ability, teaching evaluation, and reflection ability in the information technology platform. The reason why it is an important change is that the three have realized the new integration of teaching relations under the integration of information technology [22]. For example, in the teaching implementation stage, teachers no longer simply rely on the static knowledge provided by textbooks for teaching but may make more use of multimedia and other information-based teaching tools, build a dynamic knowledge display platform based on static book knowledge, and present the teaching content in different ways. Second, to student-centered, the relationship between teaching and learning tends to be equal. Of course, this equal relationship is not absolute equality but emphasizes two-way promotion, two-way incentive, and two-way expansion in the promotion of knowledge and ability, as shown in Figure 4.

The outstanding feature of new students is the expansion of autonomous learning ability. Of course, this is the change of learning environment created by information environment, but whether it can be realized also depends on students' own learning enthusiasm. In this sense, informatization only provides potential conditions for students' autonomous learning, and its real effectiveness is inseparable from teachers' guidance and assistance. However, on the basis of the expansion of autonomy, more importantly, students' collaborative ability has been enhanced. In traditional teaching activities, the expansion of students' autonomy is limited by many conditions. However, in the information environment, students' autonomous learning ability and opportunities have been enhanced. At the same time, in the same teaching activity, through mutual observation and reference, exchange and discussion, and mutual cooperation, the ability of students has been improved, and the learning efficiency is bound to increase [23]. At the same time, for normal students, the enhancement of learning ability based on self-experience has a great enlightenment for their future teaching activities, as shown in Figure 5.

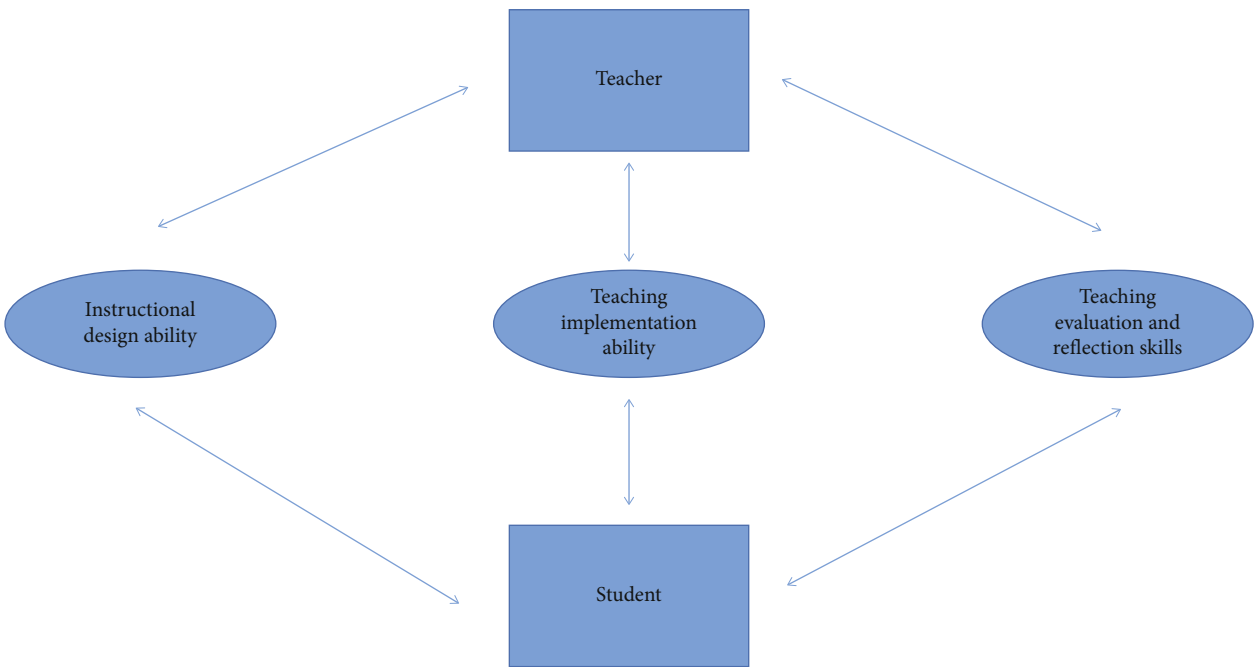


Figure 4: New teaching relationship in information environment.

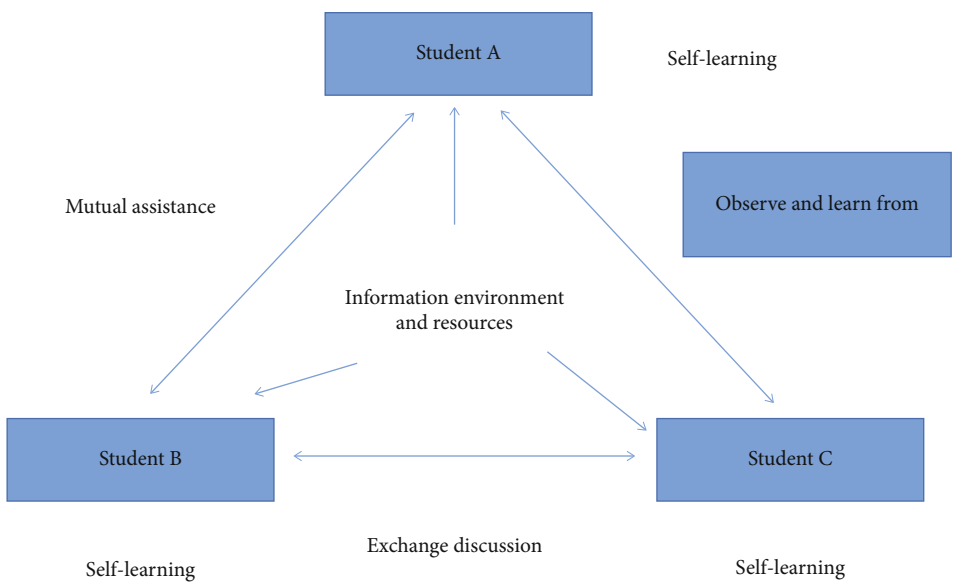


Figure 5: Development model of students' common learning ability based on informatization.

4. Experiment and Discussion

Short-time Fourier transform (STFT) is an effective tool to describe the local spectrum distribution of signal and its change with time. It is also the most commonly used time-frequency transform method in music information retrieval. Given a discrete-time signal x with sampling rate F_s, the STFT of x can be obtained by the following steps: firstly, x is divided into frames on the time axis, and there is generally a certain proportion of overlap between adjacent time frames. Then, add windows for each time frame, such as Hamming window and Hanning window, to reduce spectrum leakage. Finally, each windowed time frame is Fourier transformed, and the results are spliced frame by frame. Assuming that the frame length used for STFT is N, the overlap is 50%, and the window used is w, the STFT result of X can be written as follows:

$$X(k,t) = \sum_{n=1}^{N} w(n) x\left(n + \left(\frac{t-1N}{2}\right)\right) \times \exp\left\{-j2\pi(k-1)\frac{(n-1)}{N}\right\}, \tag{1}$$

where $k = 1$, k and $t = 1, \ldots, t$ is the index of frequency band and time frame, respectively, $k = N/2 + 1$, and t is the number of frequency band and time frame, respectively. The time corresponding to $X(k, t)$ is $(t-1)N/(2F_s)$ (in seconds), and the frequency is $(k-1)F_s/N$ (in Hertz).

The frequency scale of STFT is linear, which is inconsistent with human perception of music. At present, the popular music theory uses a logarithmic frequency scale, which defines that the distance between each frequency and twice its frequency (i.e., an octave) corresponds to an interval of the same length. Influenced by this theory, researchers began to use the time-frequency representation of logarithmic frequency scale in music information retrieval. The center frequency of each frequency band is calculated according to the following formula:

$$f_k = f_1 \cdot 2\frac{k-1}{B}, \tag{2}$$

where $k = 1$, k stands for Band Index, k represents the number of frequency bands, f_1 is the center frequency of the lowest frequency band, and B stands for the number of bands per octave. Note that the ratio of the bandwidth of each band to the center frequency is a constant:

$$Q = \frac{f_k + 1 - f_k}{f_k} = \frac{f_k + 1}{f_k} - 1 = 2\varphi - 1. \tag{3}$$

Therefore, the logarithmic frequency scale time-frequency transform is also called constant Q transform (CQT).

The time-frequency representation based on Gammatone filter bank is commonly used in computational auditory scene analysis (CASA). Given a piece of music signal, this time-frequency representation can be calculated by the following steps: (1) make the input signal pass through a gamma pass filter bank (usually 128 frequency band), and the center frequencies of each frequency band are distributed in proportion to the equivalent rectangular bandwidth (ERB). The Gammatone filter bank is the standard model of cochlear filtering, in which the impulse response of the Gammatone filter bank with center frequency f is as follows:

$$g(t) = t^{N-1} \exp\{-2\pi bt\} \cos(2\pi ft), \tag{4}$$

where N is the rank of the filter and b is the equivalent rectangular bandwidth. (2) Let the response signal of each filter band pass through a hair cell transduction model, which simulates many characteristics of auditory nerve response. (3) In each filter band, the output signal is divided into frames on the time axis, and there must be some overlap between adjacent frames.

Matrix decomposition is a technique to decompose the input matrix into the product of two or more matrices. In the field of music information retrieval, a popular matrix decomposition method in recent years, nonnegative matrix factorization (NMF) has been used. NMF is an unsupervised use of linear representations for nonnegative data. Given a nonnegative matrix X of dimension KxT and a positive integer J, NMF will find an approximate decomposition:

$$X \approx BG. \tag{5}$$

NMF imposes nonnegative restrictions, allowing only addition to the original data rather than subtraction combination, resulting in part-based representation.

In NMF decomposition, formula (5), in this process, the elements of all matrices are limited to nonnegative. For audio applications, the most frequently used cost function is Kullback-Leibler (K-L) divergence, as shown in the following formula:

$$D(X\|BG) = \sum_{K=1}^{K}\sum_{t=1}^{T} X_{k,t} \log \frac{X_{k,t}}{[BG]_{k,t}} - X_{k,t} + [BG]_{k,t}. \tag{6}$$

So it is closer to human auditory perception system. Two new mono singing and accompaniment separation algorithms will be proposed. In these two algorithms, we will also use K-L divergence as the cost function to solve NMF. In order to solve the minimization problem of K-L divergence, a simple method is proposed. This method first initializes B and G with random positive numbers and then iteratively updates B and G with the following multiplication rules.

$$B - B \otimes \frac{(X/BG)G^T}{1G^T}, \tag{7}$$

$$G - G \otimes \frac{B^T(X/BG)}{B^T 1}, \tag{8}$$

where $A \otimes B$ and A/B are multiplication and division between A and B, respectively, and 1 is a full 1 matrix of the same size as X. The CQT time spectrum is converted into gray image for image feature extraction. In order to achieve this goal, we first logarithmicize the amplitude of the time spectrum according to the following formula.

$$s(k, t) = \log |Y(k, t)|, \tag{9}$$

where Y is the spectrum at CQT and k and t are the indexes of frequency band and time frame, respectively. Compared with the original time spectrum, the logarithmicized time spectrum strengthens the components with smaller amplitude, and robust local features can be extracted from these components. After obtaining S, we then construct the time spectrum image I.

$$I(k, t) = \frac{S(k, t) - \min(S)}{\max(S) - \min(S)} \times 255, \tag{10}$$

where $\min(S)$ and $\max(S)$ are the minimum and maximum values of S, respectively.

In order to more clearly explain the above contents, we will make a more formal explanation for the relationship between pitch translation and spectral image translation

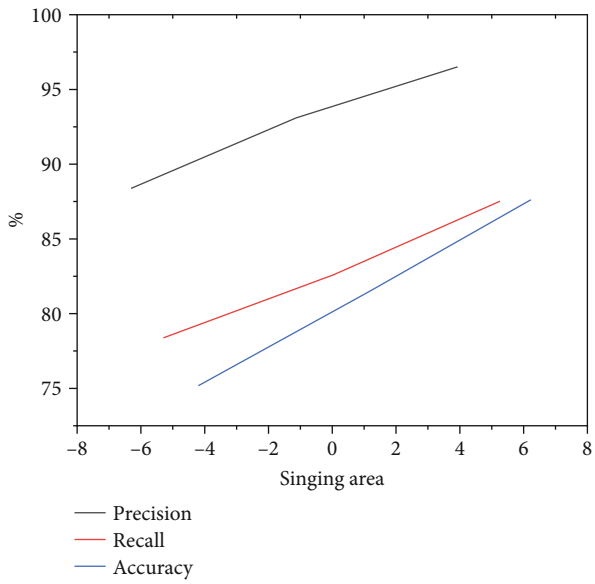


Figure 6: Mono song separation based on pitch inference improved by nonnegative matrix decomposition.

during CQT. Given a signal component with frequency f, its energy is distributed around the $y(f)$-th subband, and $y(f)$ can be obtained by transforming formula (2) into the following formula ($B = 12$).

$$y(f) = \left[12 \times \log_2 \frac{f}{f_1} + 1\right], \tag{11}$$

where $[x]$ finds the nearest integer of x. If the signal component is subjected to pitch translation with a factor of k (k is a negative number indicating pitch decrease and a positive number indicating pitch increase), its frequency becomes $(1 + k)f$ and its energy will be transferred around the $y(1 + k)$ subband. Note that the position change of the energy of the signal component on the frequency axis is independent of the value of f, as shown in the following formula:

$$y(1 + k)f - y(f) \approx [12 \times \log_2(1 + k)]. \tag{12}$$

Experimental results: as shown in Figure 6, by using the HMM-based classification method, the singing and nonsinging areas in the input song can be accurately distinguished. In particular, the accuracy of the singing area is very high for all three singing music ratios.

Figure 6 shows the precision, recall, and overall accuracy of song location for song segments in MIR-1K data set under the singing music ratio of -5, 0, and +5 dB.

Figure 7 shows the accuracy of song sound extraction for song segments in MIR-1K data set is high under the singing music ratio of -5, 0, and +5 dB.

Figure 8 shows using GNSDR for song separation of song segments in MIR-1K data set under the singing music ratio of -5, 0, and +5 dB based on the comparison among REPET proposed by Rafi, the RPCA, proposed multilayer NMF (MSNMF), the traditional pitch inference method (pitch), and improved pitch+SEG.

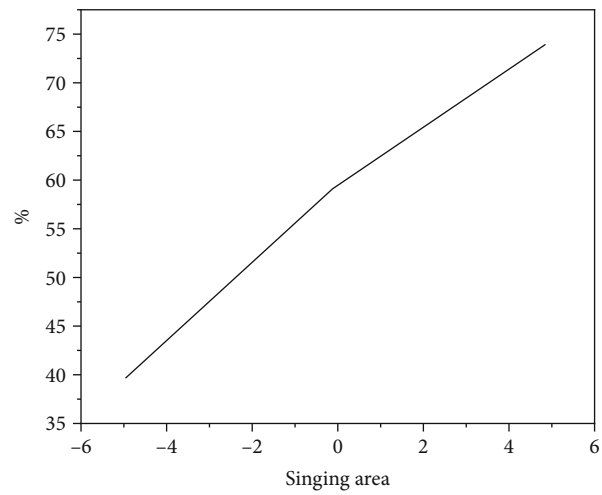


Figure 7: Mono song separation based on pitch inference improved by nonnegative matrix decomposition.

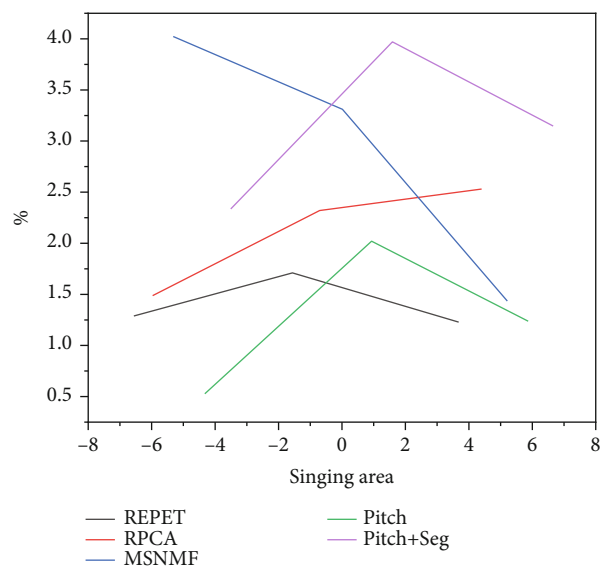


Figure 8: Mono song separation based on pitch inference improved by nonnegative matrix decomposition.

5. Conclusion

By using Fitzgerald and our proposed long+short method to separate Beach Boys songs at the singing music ratio of -6, 0, and +6 dB, the singing signal (left column) and music signal (right column) are relative to the average values of SDR (first row), sir (second row), and SAR (third row) of pure net singing and pure music. The cultivation of music normal students' teaching ability supported by information technology emphasizes the deep integration of technology, teaching method, and music subject knowledge, which is helpful to stimulate and maintain music normal students' learning motivation, understand and improve music

knowledge and skills, and master and apply teaching method knowledge and skills. Firstly, the learning motivation has been stimulated and maintained, which is mainly reflected in the fact that the music normal students have greatly improved their initiative and enthusiasm in the process of technology integrated learning with the support of information technology, especially in the increase of the interaction frequency between teachers and students and the enhancement of learners' desire for in-depth learning after class. Secondly, the mastery of the learned knowledge is more solid and reliable, which is highlighted in the achievements made by music normal students through learning, whether in music knowledge test or music creation, reflecting not only in the average score of students, but also in the number of high segments. Finally, students' mastery and application level of teaching methods have been greatly improved, which is prominently reflected in the fact that music normal students can skillfully master the knowledge of teaching methods through learning and practice, specifically reflected in the improvement of test scores. On the other hand, they can flexibly use the knowledge of teaching methods and even explore their own teaching methods in combination with their own reality. It is embodied in the improvement of the quality of its teaching design works.

Data Availability

No data were used to support this study.

Conflicts of Interest

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

References

[1] Y. Yanli and Z. Li, "Research on application of network information technology in landscape technology practice teaching," *Agro Food Industry Hi Tech*, vol. 28, no. 1, pp. 1039–1043, 2017.

[2] S. Li, Z. Hao, L. Ding, and X. Xu, "Research on the application of information technology of big data in Chinese digital library," *Library Management*, vol. 40, no. 8-9, pp. 518–531, 2019.

[3] A. D. Topal, A. K. Geer, and E. O. Budak, "An analysis of the utility of digital materials for high school students with intellectual disability and their effects on academic success," *Universal Access in the Information Society*, vol. 1, pp. 1–16, 2021.

[4] M. Liu, "Application of 3D printing technology in the production of modern complex structure sculpture," *Acta Technica CSAV (Ceskoslovensk Akademie Ved)*, vol. 62, no. 1, pp. 331–340, 2017.

[5] C. Jing, "Application and research on computer aided technology in the new music teaching method," *IPPTA: Quarterly Journal of Indian Pulp and Paper Technical Association*, vol. 30, no. 7, pp. 301–306, 2018.

[6] L. Hamond, E. Himonides, and G. Welch, "The nature of feedback in higher education studio-based piano learning and teaching with the use of digital technology," *Orfeu*, vol. 6, no. 1, pp. 01–31, 2021.

[7] J. Li, "Application of mobile information system based on internet in college physical education classroom teaching," *Mobile Information Systems*, vol. 2021, 10 pages, 2021.

[8] D. Ross, "Digital twinning [information technology virtual reality]," *Engineering & Technology*, vol. 11, no. 4, pp. 44-45, 2016.

[9] V. Cerf, "Information technology: a digital genius at play," *Nature*, vol. 547, no. 7662, pp. 159–159, 2017.

[10] I. V. Robert, "Strategic directions of development of informatization of domestic education of digital information technologies period," *Russian Digital Libraries Journal*, vol. 23, no. 1-2, pp. 145–164, 2020.

[11] I. V. Kazmina, "The impact of digital information technology on improving the competitiveness of high-tech enterprises," *Proceedings of the Voronezh State University of Engineering Technologies*, vol. 82, no. 2, pp. 174–180, 2020.

[12] M. N. Habib, W. Jamal, U. Khalil, and Z. Khan, "Transforming universities in interactive digital platform: case of city university of science and information technology," *Education and Information Technologies*, vol. 26, no. 1, pp. 517–541, 2021.

[13] F. C. Paletta and A. M. D. Silva, "Information technology management: it governance in digital library/gesto das tecnologias de informao: a sua governao na biblioteca digital," *Brazilian Journal of Development*, vol. 7, no. 1, pp. 10225–10242, 2021.

[14] H. Y. Wang, L. Sigerson, and C. Cheng, "Digital nativity and information technology addiction: age cohort versus individual difference approaches," *Computers in Human Behavior*, vol. 90, pp. 1–9, 2019.

[15] D. S. Maylawati, "Digital native on information technology acceptance in higher education," *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 9, no. 2, pp. 1201–1206, 2020.

[16] A. Khurramov, "The role and role of digital economy and information technology in the agricultural sector," *International Journal on Integrated Education*, vol. 3, no. 2, pp. 42–44, 2020.

[17] H. Khudov, "The information technology for building a test sequence to control the technical condition of digital devices," *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 9, no. 4, pp. 5987–5993, 2020.

[18] D. Beer, "Digital depression: information technology and economic crisis," *European Journal Ofunication*, vol. 32, no. 2, pp. 167-168, 2017.

[19] D. Beer, "Book review: digital depression: information technology and economic crisis," *European Journal of Communication*, vol. 32, no. 2, pp. 167-168, 2017.

[20] A. Scarano, "Digital ecology needs revamping in view of the subtle pervasiveness of information technology," *Domus*, vol. 1, p. 40, 2019.

[21] I. Yoga, P. Ni, and M. Ni, "Information technology adoption on digital marketing communication channel," *International Journal of Social Sciences and Humanities*, vol. 3, no. 2, pp. 95–104, 2019.

[22] F. Shadadeh, M. Samadbeik, F. Amiri, and A. Hajipourtalebi, "The digital gap in patients' use of health information technology and effective factors and strategies; a systematic review," *Health Research Journal*, vol. 4, no. 3, pp. 181–188, 2019.

[23] Y. G. Vasin and L. I. Lebedev, "Intellectualization of information technologies for producing digital graphic documents with weakly formalized description of objects," *Pattern Recognition & Image Analysis*, vol. 26, no. 3, pp. 638–642, 2016.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 7081286, 12 pages
https://doi.org/10.1155/2022/7081286

Research Article

A Study on the Construction and Utilization of Korean Prehistoric Remains Database

Heesoo Choi,[1] Jong Youl Hong,[2] Sangheon Kim,[1] and Kwangcheol Rim[3]

[1]*Department of Historical Contents, SangMyung University, 20 Hongjimun-2gil, Jongrogu, 03016, Seoul, Republic of Korea*
[2]*Department of Culture Contents, Korea University, 2511 Sejong-ro, Sejong 30019, Republic of Korea*
[3]*The Basic Science Institute, Chosun University, 309 Pilmundae-ro, Dong-gu, Gwangju 61452, Republic of Korea*

Correspondence should be addressed to Kwangcheol Rim; rim1201@hanmail.net

Received 13 October 2021; Revised 17 November 2021; Accepted 10 December 2021; Published 27 January 2022

Academic Editor: Ahmed Farouk

The excavation of the prehistoric sites in Korea has been on since the Japanese colonial period. However, it was only after the 1970s that it was undertaken in earnest. Many excavation research institutes, including state agencies, are still conducting excavation research. However, the excavation report, which summarizes the findings, is not serviced on an integrated platform. As a result, acquisition of integrated knowledge and research on Korea's prehistoric remains are not properly facilitated. Therefore, it is urgent to establish a database of prehistoric remains. This requires considering the characteristics of archeological excavation work and the specificity of publishing excavation reports. It is desirable to design and build database (DB) tables for excavation reports of relics and ruins, multimedia, and excavation investigations, by focusing on the DB table for prehistoric remains. Once the database is established, it will help expand cultural heritage information services through tools, such as electronic maps or the Internet of Things.

1. Introduction

1.1. The Necessity of Building a Database of Information on Prehistoric Remains in Korea. In historical research, the oldest historical sites are fascinating to explore. This is because they represent history of the present and help explore the present time's origins. In particular, the exploration of the period before the written script is entirely dependent on the achievements of archeological excavations [1]. Therefore, the systematic service of excavation reports, the result of excavations of prehistoric sites, can be of great help for historical research.

The excavation of remains of the Korean prehistoric period began in the era of Japanese occupation. The investigation by Japanese scholars strongly distorted results and denied the long history of the Korean Peninsula. After independence from the Japanese, the excavations carried out by Koreans were aimed at correcting history that was ignored or intentionally distorted by Japanese scholars. As a result, many remains of the prehistoric times, dating back to the period spanning the Paleolithic Age to the Bronze Age, have been found on the Korean Peninsula, as in other regions.

Initially, the main subjects in charge of the excavation were the national research institutes. However, as the discovery of the remains progressed, university research institutes and private research institutes in the region gradually became involved. These phenomena, and further academic development in the field of archeology and active research activities on historical sites, were perceived as a positive development. However, it is difficult to obtain comprehensive information on prehistoric remains. For instance, the excavation reports prepared by various research institutions are neither systematically organized nor accessible. Instead, they are separately managed and provided by each institution, or through commercial services offered by private companies.

Among the information on prehistoric remains, crucial information pertains to the period to which the prehistoric remains belong, the characteristics of the space it is located

in, and the other relics found in these ruins. Furthermore, it is important to study the history of the relevant era, such as that of how the ruins are distributed across the Korean peninsula, by synthesizing such crucial pieces of information. Therefore, a systematic information service for prehistoric remains needs to be urgently established.

However, this study had some limitations. First, the information related to prehistoric remains is not integrated online or offline but is scattered. Therefore, it is necessary to provide information, on a unified platform. Second, the various materials related to prehistoric sites, such as booklets, digitized files, photographs, drawings, and images, are not systematically managed. Most of them are found in excavation reports of excavating organizations, but digital files such as images and 3D objects cannot be viewed in the report. In other words, it is necessary to manage digital information regarding excavations systematically. Third, various information related to prehistoric remains are not provided in many formats.

Therefore, it is urgent to establish a database of prehistoric remains. In this paper, we review the requirements for building a database of prehistoric remains and attempt to construct a DB table for building a database that is based on the results of the review. Furthermore, we would like to explore ways to utilize the database of prehistoric remains as a "knowledge platform" [2].

2. Materials and Methods

2.1. Current Status of Korean Prehistoric Remains Excavation Report Service. Among the information on prehistoric remains in Korea, the most significant is the excavation report. The excavation report has a rich stock of information on aspects such as the various appearances of the remains in the excavated sites, the existence of the relics excavated from the ruins, and the historical significance of the remains. Therefore, the excavation report can be said to be the most basic set of data on prehistoric remains in Korea. What are the ways to obtain excavation reports? The excavation report service of the remains investigated in Korea has recently been provided through the Internet as administrative information. The representative example is the excavation investigation report service provided by the Cultural Heritage Administration [3]. (Figure 1).

This service provides a PDF file of excavation reports submitted in Korea from July 8, 2005, to December 31, 2020. A total of 4,724 excavation reports have been uploaded, and the desired information can be obtained by selecting the name of the site, publishing agency, the number of investigation attempts, and year of submission. By default, it provides reports from the latest submission date. To find the information, it is desirable to use the search results after entering the name of the site or the investigation area.

As previously mentioned , the core elements of information related to the remains are era, space, and relics. However, the previously mentioned service only provides information regarding space, that is, location-related information. For example, the methods used to check time-related information, such as the Paleolithic Era and Neolithic Era, can only be understood by reading the excavation report.

The National Research Institute of Cultural Heritage also provides excavation reports through the original text information service of the research knowledge portal on cultural heritage [4]. When searching for a given excavation report, only 87 cases appeared. If the excavation survey was, on the other hand, selected by the chronology of the excavations, a total of 15,378 excavations were found. Thus, the deviation seems to be very severe. However, it ought to be considered that excavation-related documents such as academic symposium data books and other original texts also appear in the search results. Thus, the actual number of excavation reports is much smaller.

Similar to the case of the Cultural Heritage Administration, the main information provided by the National Research Institute of Cultural Heritage is very poor. Fortunately, its service is easier to use than the Cultural Heritage Administration's Service because it provides an abstract service that helps find key information. However, it is also necessary to read the original text, in PDF, to properly comprehend information by region (space), age (time), or relic-related parameters (Figure 2).

Museums, one of the major organizations overseeing excavations, also provide information-related services. A representative example is the excavation report service of the National Museum of Korea [5]. This museum provides its own reports on excavations and investigations (Figure 3).

This online service does not conveniently provide appropriate information about the remains. The title and author of reports are the main service standards, while information on the period, region, and major relics can be viewed by downloading the report's PDF file. Only a detailed table of the contents of the report is organized and provided. Moreover, some reports are for sale in cultural goods stores. Thus, not all reports are available online.

As previously described, this study briefly reviews the major excavation report services in Korea. As a result, it was found that excavation report services have several problems. First, it is difficult to find the desired excavation report easily because there are various types of organizations and services that provide such reports. Although the services of major national institutions were examined, there are many other services given by private institutions where each of them provides their own published reports in various ways, thus, making it difficult for researchers to find the information they want.

Second, the contents of the service are concentrated only in the book called "Excavation Report." In the case of the National Museum of Korea, bibliographic information is faithfully provided, but the contents of an important excavation report cannot be known unless the original text has been read. The same applies to the rest of the services. The location-related information of the ruins can be found because it appears as a place name in the title of the report, but other information is hidden. In particular, information on the relevant period and major relics of the remains is not revealed.

Owing to these issues, it is urgent to establish a plan to integrate the excavation reports into a single service.

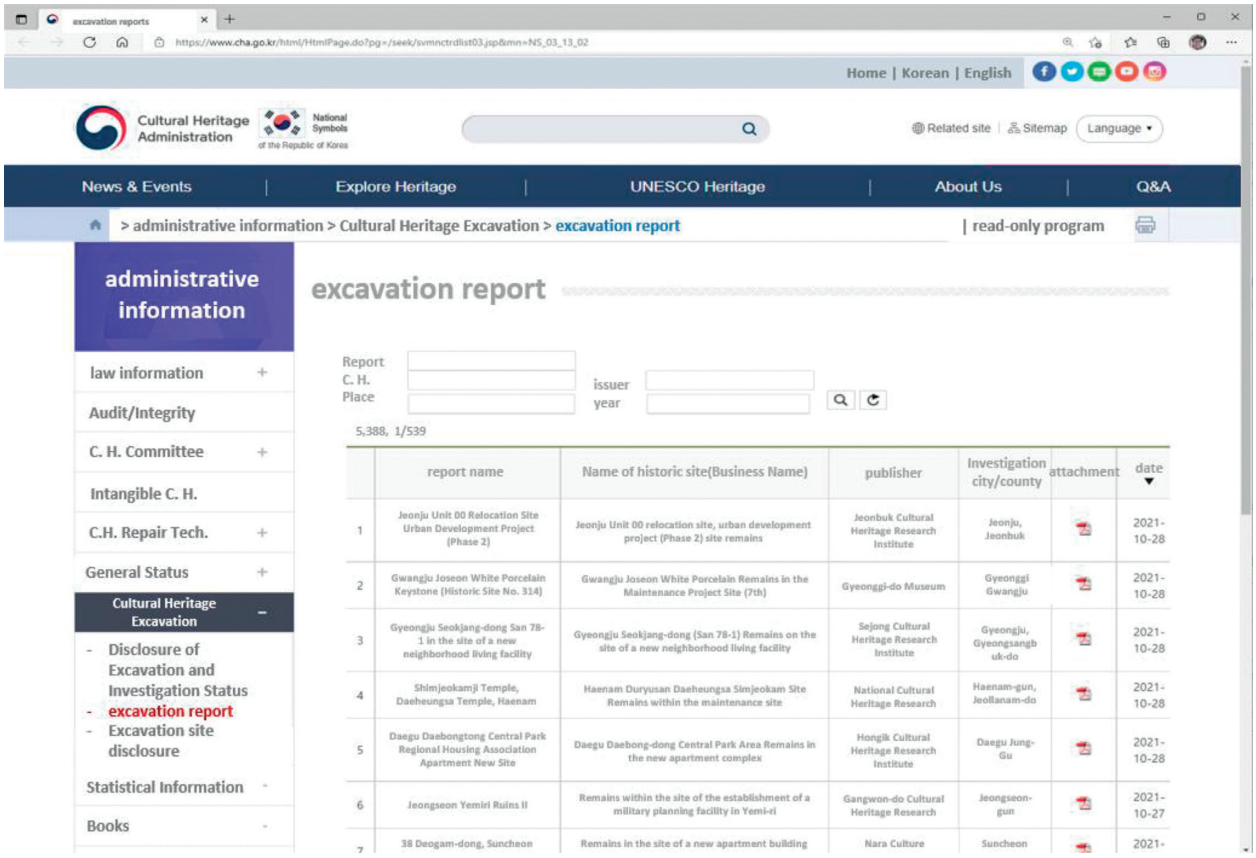


Figure 1: A screenshot of Cultural Heritage Administration's Excavation Investigation Report Service.

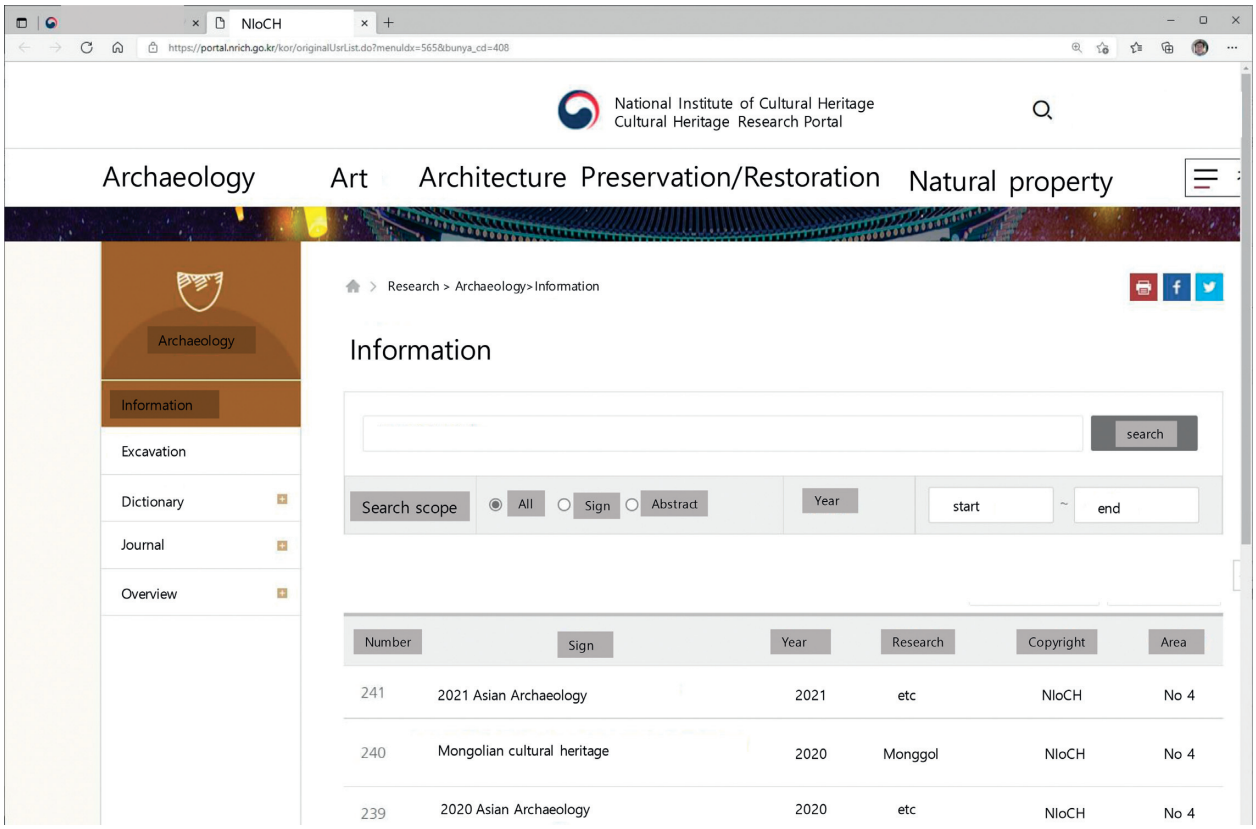


Figure 2: A screenshot of the original text information service of the National Research Institute of Cultural Heritage.

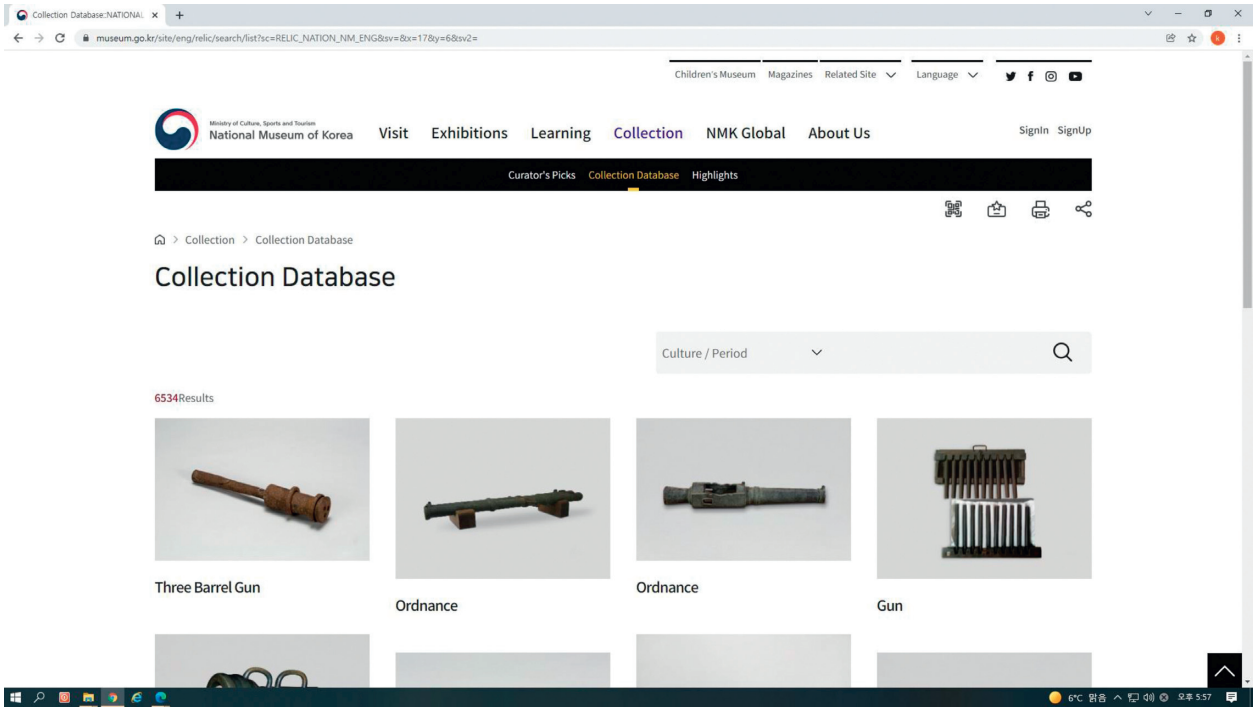


Figure 3: A screenshot of the National Museum of Korea's Excavation Report Service.

Excavation reports are composed by a wide variety of investigative agencies with a variety of contents and methods. Therefore, it is not easy to build standardized databases, such as those of academic papers. Since the specificity of the excavation report must be considered, it is necessary to review the issues of publication of excavation reports.

2.2. Characteristics of the Excavation of Korean Prehistoric Remains and the Publication of Excavation Reports. The excavation report is an academic version of the achievement of archeological research. Archeology is a study that contributes to the exploration and restoration of the history of a given time through the remains and relics left by people of that time, in addition to the history of these people, recorded as textual data. There are various objects that need to be researched in archeology, but most of them are buried. The buried objects cannot be identified before excavation; therefore, excavations are considered very important in archeology. Consequently, excavation reports containing archeological findings are very important for restoring the history and culture of a given time. To build a database of excavation reports, it is necessary to consider the characteristics, contents, and methods of archeological research [6].

Several conditions exist as prerequisites for discovering unknown archeological data. First, it should be possible to estimate that the target area will have archeological data. It is not possible to randomly excavate anywhere, so evidence is needed to gauge that there will be archeological data in the given area, through an examination of characteristics such as historical and cultural traces around it, or historical records. Second, even if there is evidence of archeological data in an area, several processes must be undertaken to determine the importance of this data. Since it is difficult to dig up an entire area, it is necessary to find data on a trial basis or investigate it through several stages [7].

Archeological excavation research involves several stages. First, excavation is carried out after a surface survey was done to collect data. A surface survey is conducted to find traces of relics or remains exposed to the surface before conducting a full-scale survey of buried materials. If an indicator survey determines that an excavation survey is necessary, an area believed to be an intensive distribution area for relics will be excavated. Excavations will also be conducted first to check the status of the buried materials before starting a full-scale excavation investigation.

Archeological excavations are conducted considering several purposes. First, there is an ordinary excavation conducted by archeologists. Then, there is research excavation that involves conducting research to verify hypotheses established for certain kinds of remains. Third, rescue excavation accounts for the largest portion of the excavation surveys. Rescue excavations should be conducted in areas at risk of destruction due to natural collapse of the ruins or manmade civil engineering structures. If a full-fledged excavation investigation is required after ordinary excavations or rescue excavations, research excavation can be conducted for decades. In other words, excavations at a site can be carried out several times.

As previously shown, the excavation process progresses, while the excavation research is summarized. Excavation is the process of collecting and recording archeological data. Through this process, the data collected are analyzed and interpreted. The data collected during the excavation survey are classified by type, their age measured, raw materials, and

functions identified, and, subsequently, they are compared with other sites. This is an analysis of archeological data. This process also involves a connection between data and cultural interpretation.

When the analysis is conducted, the intellectual process of extracting meaningful content from archeological data is carried out, which is an interpretation. There are two main interpretations of this study. One is the cultural–historical interpretation. This is used to discover the temporal, spatial, and functional elements of the site. The other is the interpretation of cultural processes. This involves the task of tracking changes in culture. The work that archeologists conduct through excavation research is to identify the nature of the remains and relics [8].

Various contents are recorded during excavation surveys. Above all, it is important to find an oil field that has traces of past structures, through the excavation of historical sites and recover and record the ruins. To begin excavating, the target area is selected and divided into several excavation sections, and excavation is carried out sequentially for each section. The data obtained during this process is recorded and collected by archeologists. The materials found are stored after systematically recording the relics in the same sequence as they were excavated, thus giving them their original location and serial number.

This is a brief overview of the characteristics and contents of archeological excavations. It is an excavation report that contains the core contents of the excavation investigation [9]. Therefore, some points from an excavation report can be summarized to build a database of excavation reports. First, the excavation report can be classified in various ways depending on the purpose and content of the excavation. Excavations are divided into categories such as ordinary excavations, research excavations, and rescue excavations. It is also confirmed that all excavations can be classified into surface surveys and excavation surveys and that excavation surveys are conducted several times depending on the characteristics of the site. Therefore, there may be several excavation reports at one site, and each excavation report may be divided according to its purpose and content.

Second, the excavation report contains various contents related to the excavation survey. Once the investigation is decided upon, an investigation team will be formed. This team comprises institutions selected as excavation agencies, and the head of the investigation team and investigators. Excavations by the investigation team usually take place over a fixed period of time. In general, one excavation report is based on a single purpose of investigation. This is closely related to the purpose of the excavation. In addition, information related to the publication of the excavation report is included after the excavation survey.

Third, the excavation report contains details of the excavation survey. In fact, these form the most important part of an excavation report. The analysis and interpretation of the archeological data previously discussed are determined using these details that largely pertain to geographical information of the site, remains excavated from the site, and relics excavated from the site.

Information on the remaining site is mainly geographical, including information such as location and geographical environment relating to the natural environment. Additionally, it also included information such as history, culture, and cultural heritage of the surrounding area related to the manmade environment. It can be accompanied by various drawings, such as geological maps, topographical maps, location maps, and distribution charts.

The remains refer to traces of past structures examined at the site. These include building sites such as houses and defensive facilities such as fortresses, traces of walls, and tombs. The shape, size, and structure of the ruins are very important because they can help identify the era to which the remains belong and various cultural phenomena. The characteristics of the site can also be defined by ruins. It can be classified as a residential site such as a house, or a village site. It can also be classified as a defensive site through identification of traces of defense facilities, or a religious site such as a Buddhist temple site, or a burial site such as a tomb, or a comprehensive site with various traces found (Figure 4).

Relics are the most carefully approached materials when examining remains. This is because they have been buried for a long time and are most vulnerable to damage when exposed. Relics can be objects people used in their daily lives. This is why all the data believed to be there can be excavated. Relics considered important in cultural history and cultural processes in archeological research so far include earthenware, stone tools, bronze weapons, and bronze ritual tools. Their excavation status is also very important for interpreting the findings of excavations.

To build a database of excavation reports on Korea's prehistoric remains, these points must be reflected upon. However, there is something to be considered here. The title is the most basic information of the excavation report. The title is determined by the purpose and scope of the excavation report. However, only after the excavation survey is completed, the value of the site established, and the site acquired as a cultural heritage site, is the name of the remains determined.

For this reason, the title of the excavation report often does not include the names of the remains known. In other words, the title makes it difficult to find information about historical sites. For example, in the case of the Jeonggwan-dong remains in Busan, the titles of excavation reports are "Cultural Remains Survey Report–Gijang-gun [10]," "Gadong, Gijang Historic Site in Jeonggwan New Town Development Zone in Busan [11]," and "Cultural Remains Research Institute in Yongsu-ri, Jeonggwan-eup, Gijang-gun [12]." The remains of Jeonggwan-dong are called Yongsu-ri Gadong Historic Site and Gijang Bangok-ri Historic Site. Thus, it is not easy to find information on the Jeonggwan-dong ruins in these titles.

It is desirable to use titles, of historical sites, that are known to the public, but most of the remains do not have such titles. In particular, in the case of excavation surveys conducted for the purpose of rescue excavation, detailed and effective policies are needed because the local names, addresses, and numbers of the areas are subject to the titles of excavation reports.

We examined the features and contents of the excavation report in Korea and found some problems. For Korea's

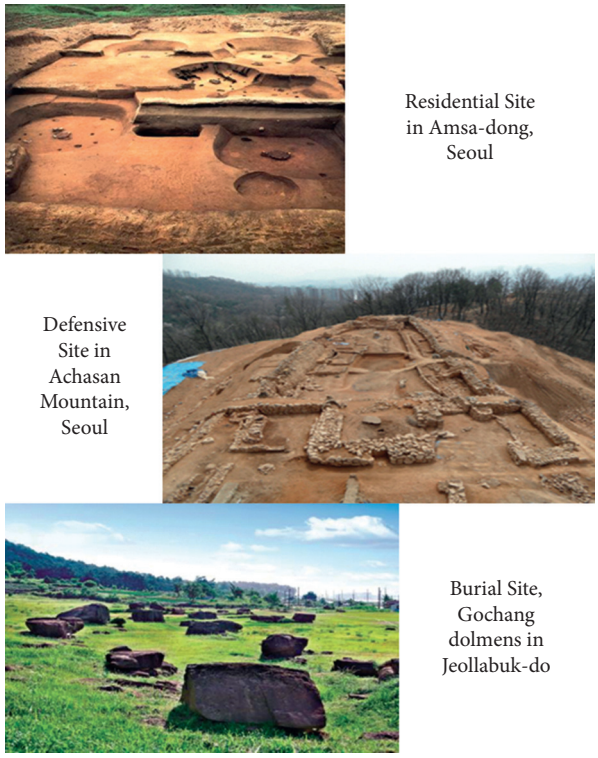


Figure 4: Case images of different archeological sites.

prehistoric remains database to be properly built and maintained, it must be first designed. In conjunction with the prehistoric heritage DB, the excavation reports DB, ruins DB, and relics DB need to be designed separately. In addition, various multimedia data related to excavation research should be connected to the DB. In addition, the establishment of the excavation investigation agency DB will facilitate the establishment of a database that covers various types of information in the archeological field.

3. Results and Discussion

3.1. Design of the Prehistoric Remains Database in Korea. As shown in the previous chapter, Korea's prehistoric remains database can be designed with six DB tables. Let us consider what is included in the database field for each table by beginning with the DB table for prehistoric remains.

The table on prehistoric remains contains four main types of information. First, it contains information regarding the physical environment of the remains. These include the site's name, corresponding era, type, and location and information on the excavation report and literature data. In addition, there is information related to excavation surveys, including the investigation agency, composition of the investigation team (team leader, investigator), and the investigation's period, purpose, scope, and method involved.

Information related to the site's investigation is included. It can be largely divided into information, and multimedia, on ruins and relics, including their names and descriptions containing the era, use, structure, and characteristics of the two classes of remains, respectively.

Multimedia information on the remains includes maps such as geological and topographical ones, distribution of historical sites, actual measurements, and photographic, video, and 3D information (to provide stereoscopic information). Based on this, the prehistoric remains DB table is as follows (Table 1).

Much of what is included in the prehistoric remains information table is extracted from the excavation report. This is why information on excavation reports is important for building the prehistoric remains information DB. As previously mentioned, several excavation reports may exist on one site. Therefore, it is desirable to configure a separate DB because it is difficult to include all excavation reports in the prehistoric remains information DB.

Since the excavation report is a booklet, it is common to include bibliographical information in it. It also includes the title of the excavation report, the author, the issuing agency, the date of issue, the size (including the number of faces), type of survey, format, language, survey period, table of contents, and overview. The DB table for the excavation report is as follows (Table 2).

Excavations of prehistoric remains should also consist of separate DB tables. In some cases, a single set of ruins is found in prehistoric sites, but most of them are found as a group of several or various types of ruins. The ruins DB table includes the name, era, purpose, structure, characteristics, area, information on excavated relics, excavation reports, and multimedia information. The ruins DB table is organized as follows (Table 3).

Similar to the ruins DB table, a relics DB table should be constructed. This is because most artifacts excavated are of many different types. The relics DB table includes the name, era, purpose, structure, characteristics, area, relevant ruins information, excavation report, multimedia information, and so on. The relics DB table is organized as follows (Table 4).

Multimedia information related to prehistoric remains includes all the images contained in the excavation reports of the past. However, owing to recent advances in digital technology, various forms of media, such as maps, drawings, photographs, images, and 3D objects, are recorded and preserved separately. Therefore, these are very important for prehistoric remains databases. Since a number of different multimedia are produced in relation to one relic or one ruins field, it is necessary to organize a multimedia information DB table separately. The DB table for multimedia information is organized as follows (Table 5).

Finally, DB tables related to the investigation agency must be constructed. In fact, in addition to state institutions, museums at each university and private research institute related to cultural assets are conducting excavations in Korea. Thus, information on these institutes is also needed. The DB table of investigation institutions includes the name, site, investigation's subject, period, and purpose, composition of investigation team, excavation report, and so on. The DB table of the investigative agency is organized as follows (Table 6).

The six tables previously mentioned are interlinked with the tables that comprise Korea's prehistoric remains

Table 1: Prehistoric remains DB table.

Elements	References	Elements	References
ID	PK	Keywords	Pl
Name		Excavation report information	Pl, excavation report ID, FK
Location		Investigation agency information	Pl, investigation agency ID, FK
Era		Ruins information	Pl, ruins ID, FK
Types		Relics information	Pl, relics ID, FK
Management		Multimedia information	Pl, multimedia ID, FK
Description			

Table 2: Excavation report DB table.

Elements	References	Elements	References
ID	PK	Language	
Title		Format	
Author		Index	
Publisher		Summary	
Date		Prehistoric remains information	Prehistoric remains ID, FK
Types		Investigation agency information	Investigation agency ID, FK
Plate type		Investigation team	
Pages		Related information	URL

Table 3: Ruins DB table.

Elements	References	Elements	References
ID	PK	Characteristic	
Name		Excavation report information	Pl, excavation report ID, FK
Location		Prehistoric remains information	Prehistoric remains ID, FK
Era		Relics information	Pl, relics ID, FK
Description		Multimedia information	Pl, multimedia ID, FK
Purpose		Keywords	Pl
Structure		Management	

Table 4: Relics DB table.

Elements	References	Elements	References
ID	PK	Characteristic	
Name		Excavation report information	Pl, excavation report ID, FK
Location		Prehistoric remains information	Prehistoric remains ID, FK
Era		Ruins information	Pl, ruins ID, FK
Description		Multimedia information	Pl, multimedia ID, FK
Purpose		Keywords	Pl
Structure		Management	

Table 5: Multimedia DB table.

Elements	References	Elements	References
ID	PK	Date	
Title		Prehistoric remains information	Prehistoric remains ID, FK
Type		Ruins information	Ruins ID, FK
Description		Relics information	Relics ID, FK
Format		Excavation report information	Excavation report ID, FK
Producer	Investigation agency ID, FK	Management	

Table 6: Investigation agency DB table.

Elements	References	Elements	References
ID	PK	Contents of investigation	Pl, prehistoric remains ID, FK
Name		Excavation report information	Pl, excavation report ID, FK
Description		Reference	

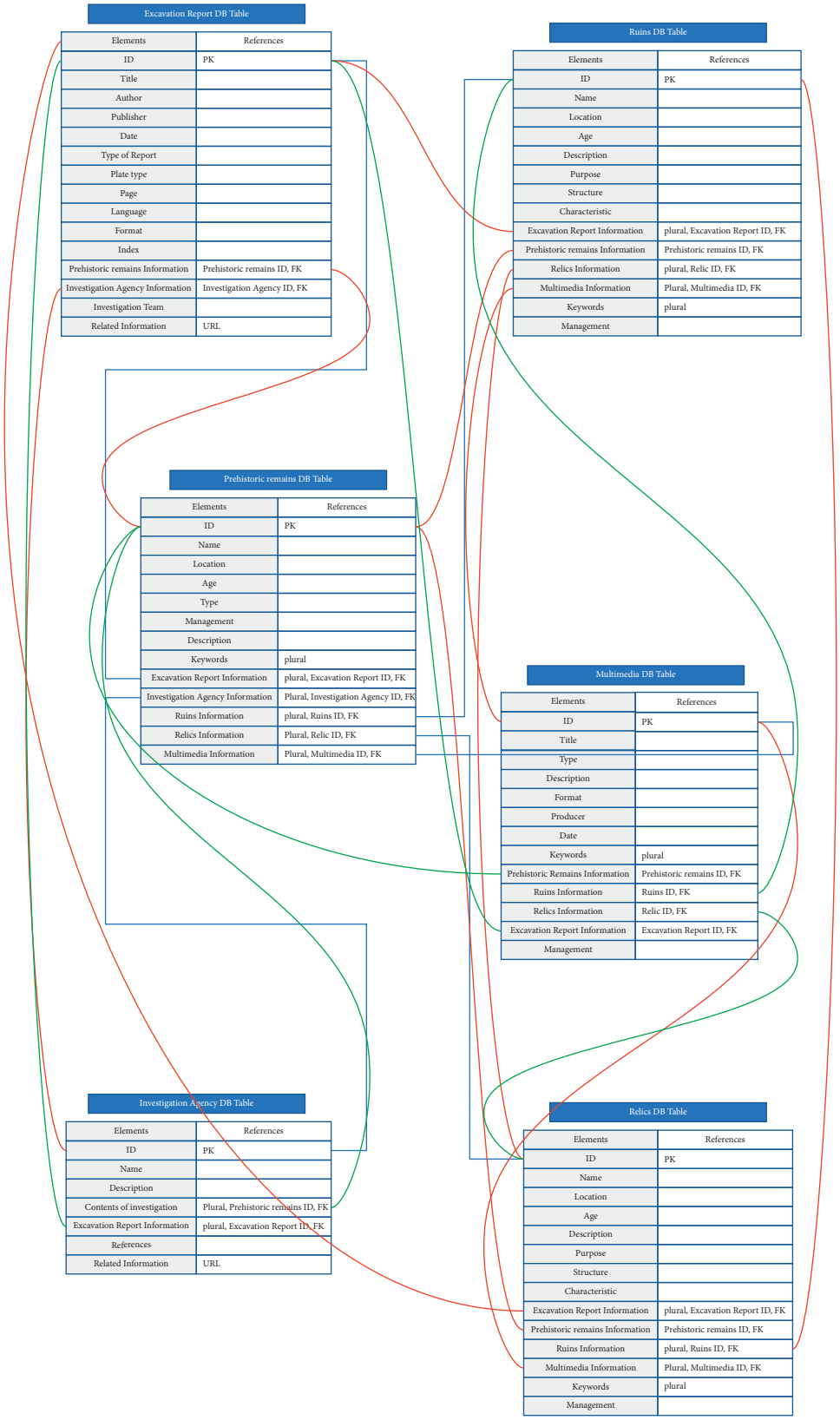
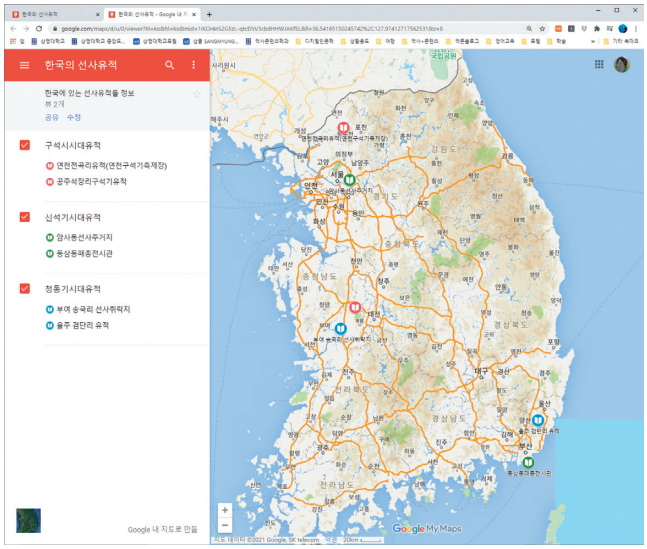
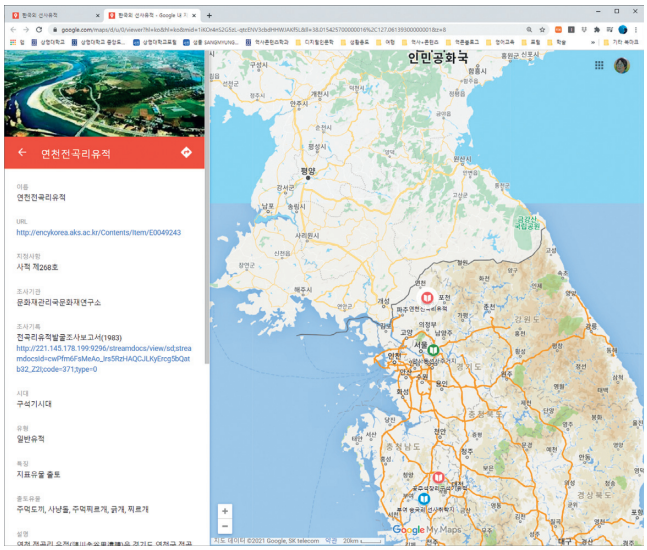


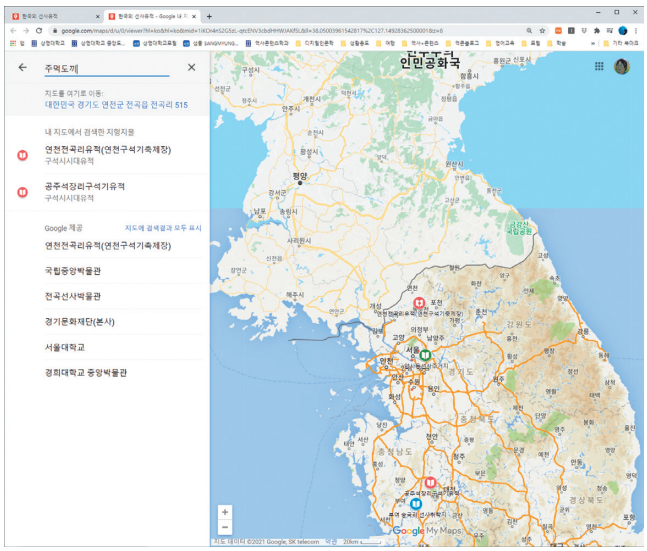
Figure 5: A conceptual map of Korean prehistoric remains database.



An example of the main screen of Korea's prehistoric sites map



An example of reading information about prehistoric sites in Korea



An example of a search in a map of prehistoric sites in Korea

Figure 6: Screenshot samples of electronic maps.

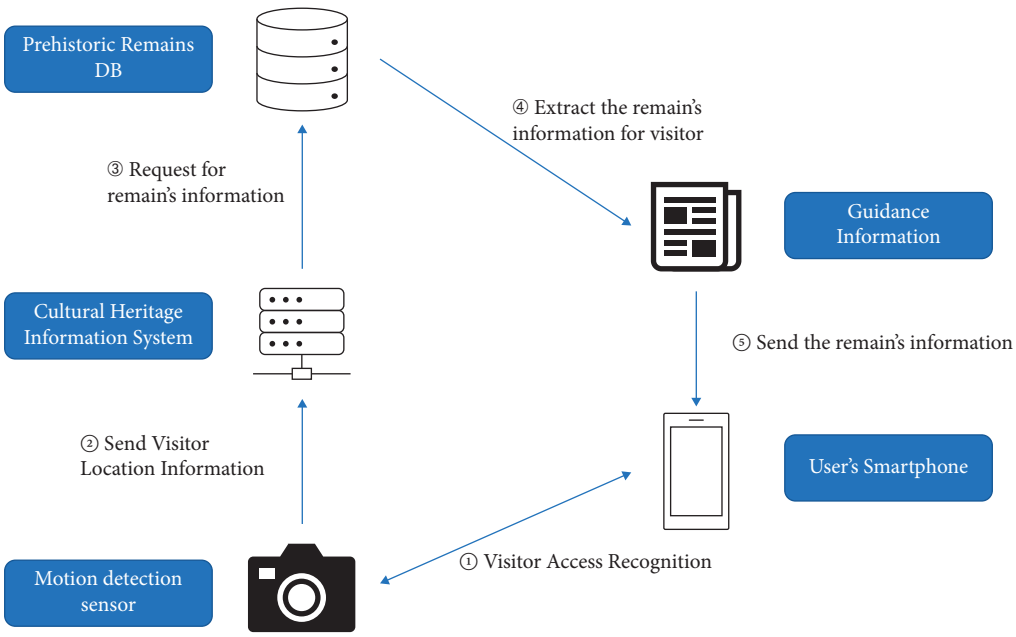


Figure 7: Remains information service based on Internet of Things conceptual map.

database. The interconnecting table is shown in the following diagram. (Figure 5).

4. Conclusions

There are many ways to collect information about prehistoric remains. It is possible to build a platform of collective intelligence [13] using the popular trend of content curation [14, 15] and supplement it using electronic maps as a platform [16]. These methods can be easily processed when various pieces of information related to prehistoric remains are converted into digital data, and the materials are available.

It is an excavation report that contains accurate and abundant information on prehistoric remains. However, it is not easy to obtain information from excavation reports related to prehistoric remains by publishing them in various ways. It may be more difficult for a single entity to collect an excavation report's elements and database it. Therefore, it is necessary to pursue projects at a national level.

A total of 15,378 excavated remains are listed in the excavation timeline of the National Research Institute of Cultural Heritage's research knowledge portal service. This includes not only prehistoric ruins but also those from the early Iron Age, the period of the Three Kingdoms, the Goryeo Dynasty, and the Joseon Dynasty. If you search for the Paleolithic ruins, 409 items are found, as are 351 Neolithic items, 1920 items from the Bronze Age, and 296 items from the Iron Age [17]. Of course, all relics from various eras are included in the search results, but, in general, approximately 2,000 relics can be targeted.

If the Korean prehistoric remains database is promoted in the previously mentioned manner, it can be used in various ways. First, as it is possible to search for prehistoric remains, it is possible to have a macroscopic approach to view prehistoric remains throughout Korea. For example, collective access to the Paleolithic sites in Korea will be available. As a result, the living environment, lifestyle, and cultural characteristics of the Paleolithic Age can be clearly identified. In other words, access from a cultural and historical perspective becomes easy.

In addition, the approach from cultural and process-related perspectives is also clear. For example, pottery is a representative artifact of the Neolithic and Bronze Ages, which has developed with different characteristics depending on the time and region. Today, only researchers who have studied earthenware for decades are capable of understanding the system or the developmental process. However, when this database is established, they will be able to easily understand the contents as well.

This will provide many other contents. In particular, new remains may be discovered or new cultural implications found through local distribution or the migration of systems. For example, if there is a site of iron production from the early Iron Age, identifying how iron was sourced may also be considered, and in the case of dolmens, which are typical relics of the Bronze Age, efforts can be made to find new relics in tombs and currently unknown residential areas.

Electronic maps are the most useful [18]. Electronic maps have many software tools that are easily accessible to the public. These include Google Maps, Google Earth, and QGIS. If you input the information regarding Korean prehistoric remains into the electronic map platform, the following picture appears: Figure 6.

As with electronic maps, a field where prehistoric remains databases can be used is a cultural heritage information service using the Internet of Things [19]. This service can be divided into two main categories. First, it uses motion-detection sensors. The sensor is attached to the cultural heritage information board and the milestone on the

cultural heritage viewing route such that it provides the corresponding information through visitors' mobile phones when they stay put or move. The system automatically guides visitors without any additional manipulation, so they can conveniently see the remains. The Internet of Things is provided through the connection of visitors' mobile phone sensors and motion sensors on the cultural heritage information board [20] (Figure 7).

Second, eye-tracking cameras can be used [21]. It is a method of installing a visual tracking camera on the cultural heritage information board to track visitors' reading of information on cultural heritage and automatically guide them with more detailed or additional information. This method has certain limitations. Currently, eye-tracking cameras have problems related to their size and cost. A company recently developed a technology for eye-tracking cameras that can be installed on PC frames, and it is expected to take time to commercialize them. However, these services will be activated and soon be used widely.

These Internet of Things-based services on cultural heritage-related information enable the accumulation of big data to improve cultural heritage management and guidance services by monitoring visitors' viewing patterns and stay time-related information [22]. In addition, more advanced guidance services can be promoted through the use of more advanced sensors. For example, it is possible to check, through sensors on the cultural heritage information board, the height of, estimate the age range of, and provide customized guidance for visitors of each age group [23].

These contents can further enhance the utilization of prehistoric remains information databases. Until now, information related to prehistoric remains has only been gathered separately and individually during the production, distribution, and consumption processes, resulting in many inconveniences. In addition, academically, the analysis of the overall view and relationship of prehistoric remains has been dependent on the individual abilities of researchers. The construction of prehistoric remains databases first contributes to systematically connecting scattered information in a certain context. It is also meaningful to create a foundation for officials and related agencies to collaborate in the process of its establishment. When a database of Korean prehistoric remains information is built and serviced through a single platform, advanced research results will enable experts to overcome these limitations. Moreover, cultural heritage information services, such as electronic map services and the Internet of Things-based information services, will be provided more effectively.

In this way, if an information database on Korea's prehistoric remains is established, it will function as a real digital archive. Individuals or institutions that individually hold information related to prehistoric remains provide archives, which can be structured and contextualized for reuse as resources for distributed databases on digital networks [24]. Information related to prehistoric remains is distributed in various ways, including those through museums, research institutes, government offices, and libraries. By combining them into a single service, knowledge about Korea's prehistoric times is expected to be systematized.

Data Availability

The data supporting the conclusions of the study are available from the corresponding author.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

This work was supported by the Ministry of Education of the Republic of Korea, the National Research Foundation of Korea (NRF-2019S1A5B5A02041334), and the National Research Foundation of Korea grant funded by the Korea government (no. 2021S1A5A2A0106408911).

References

[1] S. M. Nelson, *The Archaeology of Korea*, Cambridge University Press, Cambridge, UK, 1993.

[2] R. Shao, Y. Q. Guo, X. H. Fu, and S. M. Zhong, "Practice research on the subject knowledge platform based on the knowledge architecture,"vol. 61–66, pp. 2352–5401, in *Proceedings of the 2nd Information Technology and Mechatronics Engineering Conference (ITOEC 2016TOEC 2016)*, vol. 61–66, ITOEC, Chongqing, China, May 2016.

[3] Cultural Heritage Administration, "Excavation reports web service," 2021, https://www.cha.go.kr/html/HtmlPage.do?pg=/seek/svmnctrdlist03.jsp&mn=NS_03_13_02.

[4] National Research Institute of Cultural Heritage, "Excavation reports service," 2021, https://portal.nrich.go.kr/kor/originalUsrList.do?menuIdx=565&bunya_cd=408.

[5] National Museum of Korea, "Publication web service," 2021, https://www.museum.go.kr/site/main/archive/report/category/category_50?cp=1&pageSize=8&sortOrder=REPORT_YEAR%20DESC,%20ARC_CREATED&sortDirection=DESC&catId=50.

[6] K. K. Hirst, "What is an archaeological future? thoughtco," 2020, https://www.thoughtco.com/what-is-an-archaeological-feature-170909.

[7] M. Carver, "Excavation methods in archaeology," in *Encyclopedia of Global Archaeology*, C. Smith, Ed., Springer, New York, NY, USA, 2014.

[8] P. Historical, "Museum commission pennsylvania archeology," 2021, https://www.phmc.state.pa.us/portal/communities/archaeology/resources/value-archaeology.html.

[9] Jigsaw Cambridge shire, "Historic environment record," 2021, https://jigsawcambs.org/images/Introduction_to_Archaeological_Report_Writing.pdf.

[10] Pusan National University Museum, *Cultural Remains Survey Report — Gijang-Gun*, Pusan National University Museum, Busan, South Korea, 1998.

[11] Ulsan University Museum, *Ga-dong Site, Gi-jang I*, Ulsan University Museum, Ulsan, South Korea, 2008.

[12] Bu Kyung Cultural Relics Research Institute, *Gadong Historic Site and Gijang Bangok-Ri Historic Site I. II. III*, Bu Kyung Cultural Relics Research Institute, Busan, South Korea, 2014.

[13] J. M. Leimeister, "Collective intelligence," *Business & Information Systems Engineering*, vol. 2, no. 4, pp. 245–248, 2010.

[14] E. Yakel, "Digital curation," *OCLC Systems and Services: International digital library perspectives*, vol. 23, no. 4, pp. 335–340, 2007.

[15] Y. Feng and L. Richards, "A review of digital curation professional competencies: theory and current practices," *Records Management Journal*, vol. 28, no. 1, pp. 62–78, 2018.

[16] M. K. Buckland and L. R. Lancaster, "Advances in discovery: the electronic cultural atlas initiative experience," *First Monday*, vol. 11, no. 8, 2006.

[17] National Research Institute of Cultural Heritage, "The excavation timeline," 2021, https://portal.nrich.go.kr/kor/excavationChronologyUsrList.do?menuIdx=566.

[18] J.-H. Kang and S.-H. Moon, "Research on knowledge map using electronic cultural atlas," *Journal of the Korea Institute of Information and Communication Engineering*, vol. 18–6, no. 1383, pp. 2234–4772, 2014.

[19] M. Gigli and S. Koo, "Internet of things: services and applications categorization," *Advances in Internet of Things*, vol. 1, no. 2, pp. 27–31, 2011.

[20] J. W. Kim, S. H. Sul, and J. B. Choi, "Development of real-time internet of things motion detection platform applying non-contact sensor based on open source hardware," *International Journal of Distributed Sensor Networks*, vol. 16, no. 7, Article ID 155014772094402, 2020.

[21] A. F. Klaib, N. O. Alsrehin, W. Y. Melhem, H. O. Bashtawi, and A. A. Magableh, "Eye tracking algorithms, techniques, tools, and applications with an emphasis on machine learning and internet of things technologies," *Expert Systems with Applications*, vol. 166, pp. 114037–114174, 2021.

[22] M. Ge, H. Bangui, and B. Buhnova, "Big data for internet of things: a survey," *Future Generation Computer Systems*, vol. 87, no. 601–614, pp. 601–614, 2018.

[23] Z. Zheng, J. Zhu, and M. R. Lyu, "Service-generated big data and big data-as-a-service: an overview," in *Proceedings of the 2013 IEEE International Congress on Big Data, BigData*, vol. 403-410, July 2013.

[24] K. Ezéchiel, S. Kant, and R. Agarwal, "A systematic review on distributed databases systems and their techniques," *Journal of Theoretical and Applied Information Technology*, vol. 96, no. 1, pp. 233–266, 2019.