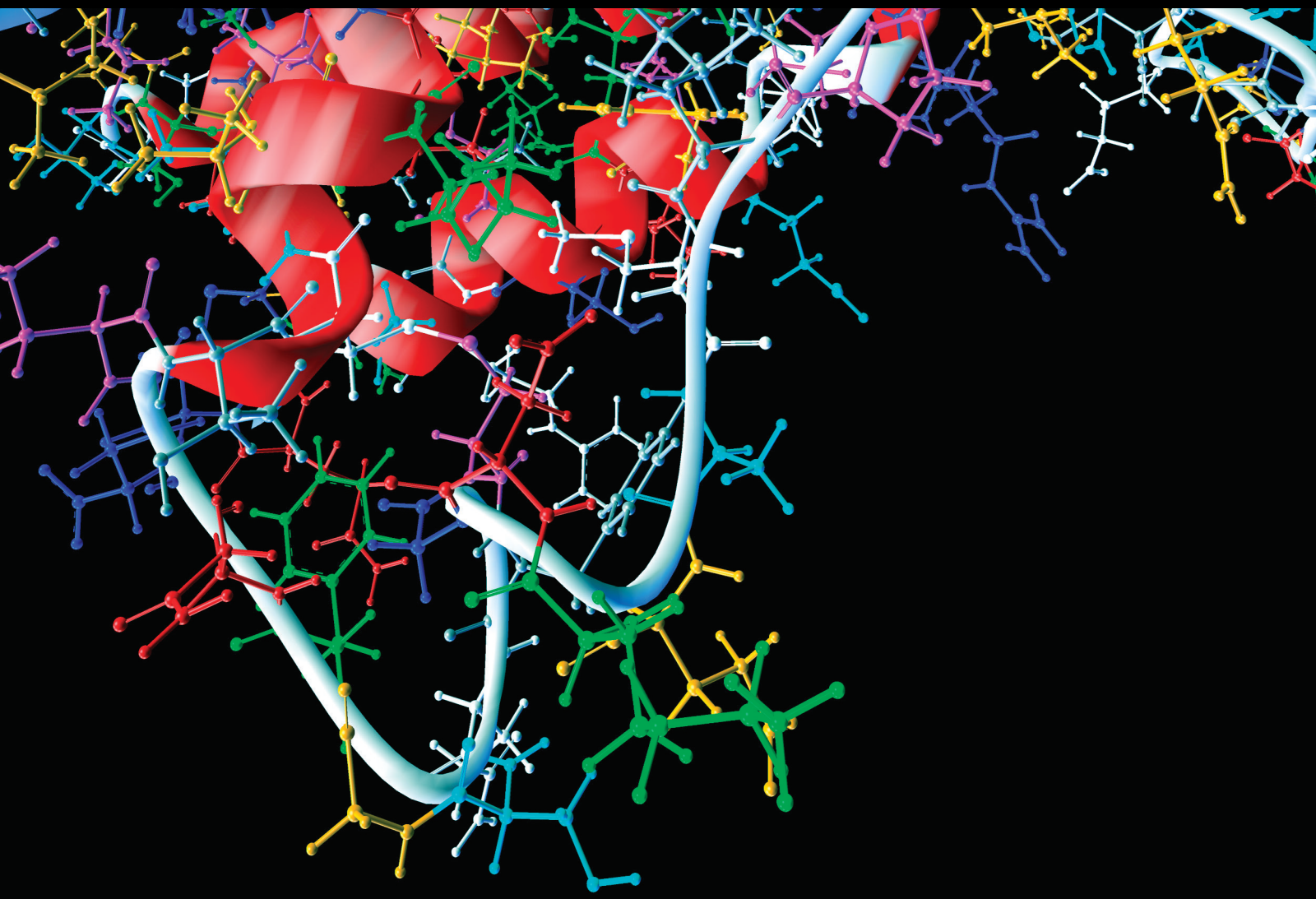
Computational and Mathematical Methods in Medicine

Advanced Computational Intelligence Methods and Ubiquitous Computing Model for Combating Infectious Disease

Lead Guest Editor: Kaijian Xia
Guest Editors: Alireza Jolfaei and Antonio Bucchiarone



Advanced Computational Intelligence Methods and Ubiquitous Computing Model for Combating Infectious Disease

Computational and Mathematical Methods in Medicine

Advanced Computational Intelligence Methods and Ubiquitous Computing Model for Combating Infectious Disease

Lead Guest Editor: Kaijian Xia
Guest Editors: Alireza Jolfaei and Antonio Bucchiarone

Contents

Hindawi
Computational and Mathematical Methods in Medicine
Volume 2023, Article ID 9859656, 1 page
https://doi.org/10.1155/2023/9859656

Retraction

Retracted: PSSPNN: PatchShuffle Stochastic Pooling Neural Network for an Explainable Diagnosis of COVID-19 with Multiple-Way Data Augmentation

Computational and Mathematical Methods in Medicine

Received 26 September 2023; Accepted 26 September 2023; Published 27 September 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] S. Wang, Y. Zhang, X. Cheng, X. Zhang, and Y. Zhang, "PSSPNN: PatchShuffle Stochastic Pooling Neural Network for an Explainable Diagnosis of COVID-19 with Multiple-Way Data Augmentation," *Computational and Mathematical Methods in Medicine*, vol. 2021, Article ID 6633755, 18 pages, 2021.

Hindawi
Computational and Mathematical Methods in Medicine
Volume 2022, Article ID 9934144, 8 pages
https://doi.org/10.1155/2022/9934144

Research Article

Computational Intelligence Method for Detection of White Blood Cells Using Hybrid of Convolutional Deep Learning and SIFT

Mohammad Manthouri,[1] **Zhila Aghajari**,[2] **and Sheida Safary**[3]

[1]*Electrical and Electronic Engineering Department, Shahed University, Tehran, Iran*
[2]*Khaje Nasir University of Technology, Electronic and Computer Engineering Faculty, Tehran, Iran*
[3]*Department of Computer Engineering, Islamic Azad University North Tehran Branch, Tehran, Iran*

Correspondence should be addressed to Mohammad Manthouri; mmanthouri@shahed.ac.ir

Received 26 March 2021; Accepted 25 October 2021; Published 11 January 2022

Academic Editor: Alireza Jolfaei

Infection diseases are among the top global issues with negative impacts on health, economy, and society as a whole. One of the most effective ways to detect these diseases is done by analysing the microscopic images of blood cells. Artificial intelligence (AI) techniques are now widely used to detect these blood cells and explore their structures. In recent years, deep learning architectures have been utilized as they are powerful tools for big data analysis. In this work, we are presenting a deep neural network for processing of microscopic images of blood cells. Processing these images is particularly important as white blood cells and their structures are being used to diagnose different diseases. In this research, we design and implement a reliable processing system for blood samples and classify five different types of white blood cells in microscopic images. We use the Gram-Schmidt algorithm for segmentation purposes. For the classification of different types of white blood cells, we combine Scale-Invariant Feature Transform (SIFT) feature detection technique with a deep convolutional neural network. To evaluate our work, we tested our method on LISC and WBCis databases. We achieved 95.84% and 97.33% accuracy of segmentation for these data sets, respectively. Our work illustrates that deep learning models can be promising in designing and developing a reliable system for microscopic image processing.

1. Introduction

Despite decades of efforts and research in controlling infection diseases, they are still among the most challenging issues in public health. According to the World Health Organization (WHO), infectious diseases are now the world's most deadly communicable disease and are ranked as the 4th leading cause of human death. They are among the top global problems with human, social, and economic impacts across the globe. Therefore, the development of robust systems for early diagnosis and investigating the source of the epidemic are critical to address this global, life-threatening issue.

One important part of the body's immune system is white blood cells (WBC). The white blood cells protect the body against infectious diseases. There are five different types of white blood cells, named as lymphocytes, monocytes, eosinophils, basophils, and neutrophil. The number of white blood cells, as well as their structure, is important in the diagnosis of different infection diseases, such as HIV, rubeola, poliovirus, and chickenpox [1, 2]. This test, named as hemogram test, is done by evaluating the blood cells under a microscope. However, due to the different types of white blood cells and their complex structures, the study of blood vessels manually is highly prone to error [3]. Therefore, a lot of researchers have explored different techniques to help with automatic detection of white blood cells with high accuracy accurately.

In recent years, researchers have investigated and proposed different computational intelligence techniques for infection diseases diagnosis. These techniques include but not limited to transfer learning and deep learning [4–6].

Many researchers have focused on using these computational techniques to detect the white blood cells due to their importance in diagnosing a variety of infectious diseases. Most of these studies have focused on classification and segmentation of the white blood cells. Given the importance of detecting white blood cells, in this paper, we will first review the prior literature on classification and segmentation of the white blood cells. We will present a deep learning method using convolutional neural networks to improve the prior studies. One of our motivations to use convolutional neural networks is because they do not require seeing the entire object. Therefore, it can be a good choice to deal with cells at the edge of the microscopic images as well. We use the Gram-Schmidt algorithm to segment the nuclei in the peripheral blood samples. Next, we use Scale-Invariant Feature Transform (SIFT) feature detection to extract the most predictable features. To keep spatial neighbourhood dependences, which are specifically important in processing image data, we will use convolutional neural networks to learn contextual dependencies. For the classification purpose, we use the weighted two-phase test sample sparse representation method (WTPTSSR) that is an improvement of the method two-phase test sample representation (TPTSR) method [7]. Our motivation to choose WTPTSSR over TPTSR is that this approach keeps the locality information. Therefore, it could be more appropriate for the image classification context.

The rest of this paper is organized as follows. In Section 2, we will review the different techniques that have been used for image segmentation and classification. In Section 3, we will describe our proposed method in two parts. First, we will explain the segmentation and classification steps in detail. Following that, the simulation experiments of segmentation and classification phases will be discussed in Section 4. Next, we will describe our experiments and will report the results of our proposed method. Eventually, we will provide a discussion on how our approach provides insights in detecting white blood cells and how our method can be further improved by future research.

2. Related Work

The diversities of the white blood cells make their detection very challenging. Many researchers have investigated different techniques in this domain. These studies mostly relied on image classification and segmentation to detect the white blood cells and investigate their structures. Otsu's thresholding method is used recurrently in the circular histogram to segment the white blood cells [8] by Wu et al. In this paper, Otsu's method is applied to components H and S of the HSI color model. Gautam and Bhadauria improved the contrast of the blood microscopic image and used Otsu's thresholding for the segmentation of the white cell nucleus [9]. Mohapatra et al. did the preprocessing step by applying the median filter on the images in order to eliminate possible noises and used K-means clustering in the Lab color model to divide pixels of the blood microscopic images [10]. K-means clustering and the Lab color model for segmentation of the white cells nuclei have been also explored [11, 12]. Theera-Umpon used c-means fuzzy clustering and morphological operators to segment white cell nuclei [13]. Pan et al. used ELM classifications to extract white blood cells via utilizing visual simulations [14]. They demonstrated that ELM has equivalent performance compared to the SVM and can find efficient samples actively and train the classification model in real time, without the need to adjust the parameters.

Ko et al. [15] provided a step-by-step integration method for nucleus segmentation based on the mean-shift clustering. They also used GVF (extreme learning machine) active curve to segment the cells' cytoplasm. Hamghalam et al. used a combination of Otsu's thresholding method and a snake method-based active curve to segment the nucleus and the cytoplasm in the white blood cells [16]. Rezatofighi et al. proposed a new method for the segmentation of white blood cell nuclei based on Gram-Schmidt orthogonality [17]. They further improved their work by proposing an active curve for cytoplasmic segmentation [18].

For microscopic image edge detection, Nakib et al. [19] used a microcanonical annealing approach to optimize their criterion function through benchmarking two-dimensional exponential entropy. In [20], genetic algorithms and wavelet were used to automatically estimate the number of thresholds for multilevel thresholding of the histogram. They examined their approach of different images, including microscopic blood images. The detection process is designed to detect the ovals in blood images and extract the best of the ovals with DE algorithm. They used the Gram-Schmidt orthogonality algorithm to segment the white blood cells. In order to characterize and extract the types of white blood cells, which could have five different types, the SIFT algorithm and deep convolutional neural network were used. The deep convolutional neural network they used consists of three layers of convolution and two full layers of pooling. To address the small data size, they used WTPTSSR algorithm [21].

3. The Proposed Method

In this section, we will first outline the segmentation process, which is primarily based on Gram-Schmidt orthogonality. We will then do the classification process, using Scale-Invariant Feature Transform (SIFT) feature detection and convolutional deep neural network.

3.1. Segmentation. We used Gram-Schmidt orthogonality to segment the nuclei in the peripheral blood samples. To do that, we first extracted a three-dimensional vector for each pixel based on their RGB components. Subsequently, the weight vector has been calculated, to tune the network for the input data set. To extract the area of interest, we used the idea presented in [22]. That is, we calculate the inner product of the weight vector w and the pixel feature vectors (Figure 1). This way, the purple area of the original image will have the highest brightness intensity, whereas the rest of the image will darken.

The Gram-Schmidt process takes a finite linearly independent set $S = \{v_1, \cdots, v_k\}$ for $k <= n$ and will generate an orthogonal set $S' = \{u_1, \cdots, u_k\}$ to span the same

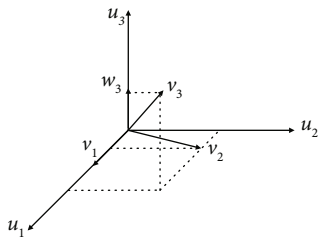


FIGURE 1: The relationship between the vector w_3 and the vectors v_1, v_2, and v_3 in the three-dimensional space.

k-dimensional subspace of R^n as S. To do that, a projection operation is defined as follows:

$$\text{proj}_u^v = \frac{<u, v>}{<u, u>} u = <u, v> \frac{u}{<u, u>}, \quad (1)$$

where $<u, v>$ represents the inner product of operator v on vector u.

Given this definition, the Gram-Schmidt orthogonality method will be as follows [19]:

$$\begin{aligned} u_1 &= v_1, \\ e_1 &= \frac{u_1}{\|u_1\|}, \\ u_2 &= v_2 - \text{proj}_{u_1}^{v_2}, \\ e_2 &= \frac{u_2}{\|u_2\|}, \\ u_3 &= v_3 - \text{proj}_{u_1}^{v_2} - \text{proj}_{u_2}^{v_3}, \\ e_3 &= \frac{u_3}{\|u_3\|}, \\ u_k &= v_k - \sum_{j=1}^{k-1} \text{proj}_{u_j}^{v_k}, \\ e_k &= \frac{u_k}{\|u_k\|}. \end{aligned} \quad (2)$$

Using this method, the w_k vector will be used for the set $S = \{v_1, \cdots, v_n\}$. Subsequently, the maximum projection on v_k and orthogonal to other vectors in the set is calculated as below:

$$w_k = v_k - \sum_{j=1}^{k-1} \text{proj}_{v_j}^{v_k}. \quad (3)$$

Eventually, we can do the segmentation based on appropriate thresholds that are chosen with respect to the histogram of the result. Given that the platelet areas are smaller than the nucleus, we can remove the small pieces and the remaining part will only include the nucleus. To eliminate the effect of the color difference and the nucleus illumination intensity between image samples, three different weighting vectors are calculated for each image. Eventually, we will apply the "AND" reasoned action on the three resulting images to get the segmentation phase results. This process is illustrated in Figure 2.

3.2. Classification

3.2.1. Scale-Invariant Feature Transform. Scale-Invariant Feature Transform (SIFT) feature detection has been used for feature extraction [23]. SIFT is based on the image gradients and is invariant to scaling and rotation [24]. It is rotation-invariant, which means even if the image is rotated, we can achieve the same result. It is scale-invariant which means changing the image scale will not affect the results. In addition, this method shows a high degree of resistance to other complex forms of transformation and illumination changes. SIFT extracts key points and feature vectors in three steps, presented in the following section [23].

Step 1. At this step, the incoming image is alternately convolved by Gaussian functions to obtain the smoothed samples of the original images. Then, the smoothed images are subtracted from each other to get the images of Difference of Gaussians (DOG).

Step 2. Next, the resulting DOG images are examined, and the maximum and minimum local points are selected as the key point. The maximum and minimum local points are the points that have the maximum or minimum values in both dimensions and scales compared to their neighbours. This feature ensures that the key points and the extracted feature vector remain invariant to the scale changes.

Step 3. Once the key points and the scales of each point are calculated in Step 1 and Step 2, the feature vectors for each key point will be calculated. First, the gradient image is calculated, which will be used to extract the key points. Subsequently, the direction of the region around the central pixel will be set on the gradient rotation of the central pixel. At this point, the gradient image is sampled for the $16 * 16$ regions around the central pixel of the gradient rotation. This step ensures that the extracted feature vector is invariant to rotation.

Next, the samples in a region are quantized in 8 main directions. The 16×16 region around the central pixel is divided into 16 regions of 4×4, and the histogram of gradient direction is calculated in each of these regions. Eventually, these sixteen 8-dimensional histograms form the final 128-dimensional feature vector [24].

3.3. Convolutional Neural Networks. In natural images, the values of pixels in a spatial neighbourhood have a high spatial dependency on each other and this dependence is independent from the neighbourhood location in the image [25]. To keep these dependencies and also to make the model invariant to spatial transformation, a convolutional neural network convolves a set of filters ($F = \{f_1, f_2, \cdots, f_{Nk}\}$) on

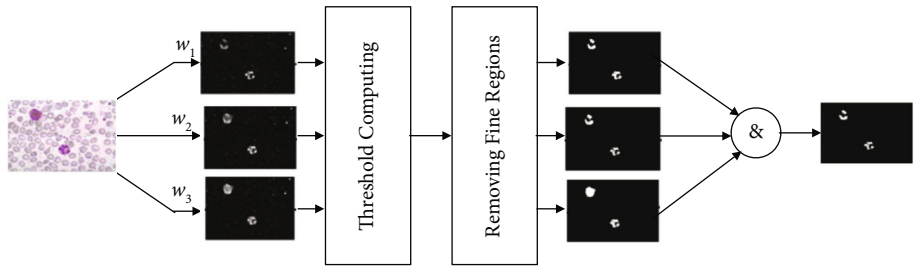


Figure 2: Segmentation of white cells using the Gram-Schmidt algorithm.

the input image and will result in the two-dimensional named as z in the following equation:

$$z = f_k * x. \tag{4}$$

These filters are learned from the input data and their gradients using a back propagation algorithm. To calculate the feature map units, the convolutional filters are transmitted through a nonlinear active function such as sigmoid function or Rectified Linear Unit function. Subsequently, a pooling layer is applied on the output of the feature map units, to make it invariant to the transmissions. Pooling action P could be done using maximizing or averaging of feature map unities of the neighbourhood G:

$$P_G = \max_{i \in G} h_i. \tag{5}$$

For the pooling phase, we used the maximum pooling method. This technique is used more often for the pooling phase as it takes care of negative values and does not blur the output units [26]. The result of the pooling layer will be sent to a regular fully connected network. In the last layer (the output layer), softmax activation is often used; however, in our work, we used the WTPSSR method instead of softmax function. The WTPTSSR method is a sparse method that will be described in detail in the next section. Subsequently, the whole network is trained using back propagation of the network error, which is calculated based on crossentropy of the last layer output.

$$c = -\sum_{c}^{N_c} y_c \text{loglog}\left(\widehat{y}_c\right). \tag{6}$$

The convolution network considered in this paper has the convolution layer and two max-pooling layers. Weight filters in the convolution layers are $3 * 3$, and zero padding is not considered in the layers.

3.4. Weighted Two-Phase Test Sample Sparse Representation. The weighted two-phase test sample sparse representation method (WTPTSSR) is an improvement of the method two-phase test sample representation (TPTSR) method [7]. The TPTSR method represents the test samples as a linear combination of the training samples. It then calculates the M nearest neighbours for each test sample based on the training samples that are most appropriate for the corresponding test sample. However, this method loses the local information, while in a lot of cases, locality is very important and holds a high recognition ratio [4]. The WTPTSSR method was presented to address this problem [27, 28]. WTPTSSR is identical to TPTSR, except that it adds locality on the l_2 regularization. The steps of the WTPTSSR methods are as follows:

(1) Input: $A \in R^{m \times N}$ training sample matrix, where N is the number of training samples and m is the number of features of each sample, and $y \in R^{m \times 1}$ is the pilot sample

(2) Columns A and Y are normalized to have normalized l_2 norms

(3) The M nearest neighbours for test samples are determined based on the equation below:

$$X = \left(A^T * A + t * W^T * W\right)^{-1} * A^T * y, \tag{7}$$

where W is a diagonal matrix and a local adaptor that penalties the distance between y and each pilot sample and is calculated as follows [29]:

$$\begin{aligned} W &= \text{diag}\left([\text{dist}(y, a_1), \text{dist}(y, a_2), \cdots \text{dist}(y, a_n)]\right) \\ \text{dist}(y, a_1) &= ||y - a_i||^k, \end{aligned} \tag{8}$$

where k is the local adaptor parameter. Note that if $k = 0$, the method will be transformed to TPTSSR. Then, the following equation will be calculated for all the training samples:

$$\text{con}_i = ||y - a_i x_i||_2. \tag{9}$$

Subsequently, M pilot samples with the lowest con_i value will be selected and construct matrix $\underline{A}$.

(4) In the next step, we will solve the linear equation (10), to calculate linear combination of the M training samples:

Table 1: Numerical results of core segmentation.

Evaluation metrics	Hamghalam et al. [16]		Salem [34]		The proposed method	
	LISC data set	Sharks WBCis data set	LISC data set	Sharks WBCis data set	LISC data set	Sharks WBCis data set
T_s	93.20	89.64	87.9	91.3	96.49	97.33
RDE	2.73	4.54	5.3	3.87	1.98	1.49
OR	0.076	0.084	0.103	0.066	0.062	0.052
UR	0.081	0.083	0.089	0.071	0.071	0.065
ER	0.179	0.194	0.24	0.156	0.128	0.134

Table 2: The confusion matrix when HoG descriptor is applied.

Predicted class	True class					
	Basophil	Eosinophil	Lymphocyte	Monocytes	Neutrophil	Accuracy
Basophil	45	4	0	5	1	81%
Eosinophil	5	24	1	7	2	61%
Lymphocyte	1	2	56	2	0	91%
Monocytes	13	9	2	22	2	45%
Neutrophil	0	2	0	1	54	94%

$$\underline{X} = \left(\underline{A}^T \underline{A} + \gamma I\right)^{-1} \underline{A}^T y, \tag{10}$$

where γ is a positive constant and I is the identity matrix

Since each of the M selected samples belongs to the same class, the degree of collaboration between each class needs to be determined [30]. Assume that t_i is a sample of the i^{th} class represented by $a_1^i, a_2^i, \cdots, a_{ti}^i$. Using the following equation, we will examine the degree of collaboration between training samples of the i_{th} class in representing y pilot sample:

$$\text{con}_i = \left\| y - \sum_{j=1}^{t_i} a_j^i x_j^i \right\|_2. \tag{11}$$

(5) Note that smaller con_i represents a greater contribution to the test sample. Therefore, the class of pilot sample y is determined as the class that gives the lowest value of collaboration

In the next section, we will present how our model worked on LISC and WBCis databases. We will also compare our model with four other baseline methods.

4. Results and Experiments

4.1. Segmentation Results. To assess the segmentation, we compare the similarity between manual and automatic segmentation. Higher similarity metric indicates more accurate segmentation. Similarity is calculated using the below equation:

$$T_S = \frac{A_{\text{Automatic}} U A_{\text{Manual}}}{\max\left(A_{\text{Automatic}}, A_{\text{Manual}}\right)} \times 100, \tag{12}$$

where $A_{\text{Automatic}}$ is the area of the automated segmented core and A_{Manual} is the area of the manually segmented core.

4.2. RDE Criterion. The relative distance error criterion (RDE) is used to assess the extracted segments [26, 31]. Assume that $e_1, e_2, \cdots . e_{n_T}$ are the E pixels and $t_1, t_2, \cdots . t_{n_T}$ are the T pixels, where E is the boundary of the image obtained from automated segmentation and T is the boundary of the image from manual segmentation. n_E and n_T are the number of segmented pixels in E and T boundaries, respectively. With these assumptions, RDE is defined according to the following equation:

$$\text{RDE} = \frac{1}{2}\left(\sqrt{\frac{1}{n_E}\sum_{i=1}^{n_E} d_{e_i}^2} + \sqrt{\frac{1}{n_t}\sum_{j=1}^{n_t} d_{t_j}^2}\right), \tag{13}$$

where $\mathbf{d_{t_j}}$ and $\mathbf{d_{e_j}}$ parameters are defined based on equation (14), and $\mathbf{distance}(\mathbf{e_i}, \mathbf{t_j})$ indicates the Euclidean distance between $\mathbf{e_i}$ and $\mathbf{t_j}$.

$$d_{e_i} = \min\left\{\text{distance}\left(e_i, t_j\right) \mid j = 1, 2, \cdots n_t\right\}. \tag{14}$$

4.3. OR, UR, and ER Criteria. $\mathbf{Q_p}$ indicates the number of pixels result from the manual segmentation that are not found in the automatic segmentation. $\mathbf{U_P}$ represents the number of pixels that result from automated segmentation and are not found in the manual segmentation. $\mathbf{D_P}$ represents the number of pixels in the manually segmented object.

The OR, UR, and ER criteria, which, respectively, indicate oversegmentation, subsegmentation, and error

Table 3: The confusion matrix when SIFT and CNN descriptors are applied.

Predicted class	True class					
	Basophil	Eosinophil	Lymphocyte	Monocytes	Neutrophil	Accuracy
Basophil	53	0	2	0	0	81%
Eosinophil	0	35	2	2	0	61%
Lymphocyte	0	1	58	1	1	91%
Monocytes	0	2	3	42	2	45%
Neutrophil	0	2	0	1	54	94%

Table 4: Comparing the accuracy of the proposed method in detecting the white blood cells with four baseline methods.

Reference	Segmentation method	Classification method	Sample size	Accuracy
The proposed method	Gram-Schmidt orthogonalization	WTPSSR	260	97.14%
The proposed model by Rezatofighi et al. [17]	Gram-Schmidt orthogonalization and snake	SVM	400	86.10%
The proposed model by Zhang et al. [26]	Histogram threshold	Distance classifier	199	92.46%
The proposed model by Balki et al. [6]	Entropy threshold and iterative threshold	Distance classifier	71	90.14%
The proposed model by Horne et al. [2]	Gram-Schmidt orthogonalization and snake	LVQ	400	94.10%

Table 5: Comparing the accuracy of the proposed approach when using different classification methods (the segmentation is done using SIFT and convolutional deep neural network across these models).

Feature extraction method	Classification method	Sample size	Accuracy
CNN + SIFT	WTPSSR	260	97.14%
CNN + SIFT	SVM	260	78.5%
CNN + SIFT	Distance classifier	260	81.2%

ratio, are calculated according to equations (15), (16), and (17) [3, 13, 32, 33].

$$\mathrm{OR} = \frac{Q_p}{U_P + D_P}, \tag{15}$$

$$\mathrm{UR} = \frac{U_p}{U_P + D_P}, \tag{16}$$

$$\mathrm{ER} = \frac{Q_p + U_P}{D_P}. \tag{17}$$

Table 1 illustrates the numerical results of the proposed method for nucleus segmentation in comparison with the methods proposed in [16, 34]. As mentioned earlier, in LISC (Leukocyte Images for Segmentation and Classification) and WBCis (Wight Blood Cell Images for Segmentation) databases, the evaluation parameters are only calculated for white blood cells.

4.4. Classification Results. We used 260 samples of images containing 720×576 pixels, all of which are colored images, to detect blood cells that contain 5 different classes. In Table 1, the results of the proposed procedure are applied to 260 white cell images such as neutrophil, basophil, monocyte, eosinophil, and lymphocyte. Tables 2 and 3 are the confusion matrix where HoG descriptor and SIFT are used along with CNN to extract features, respectively. In Table 4, we compare the accuracy of the proposed method against four baseline models. In Table 5, we compare how different classification techniques, namely, SVM, WTPSSR, and distance classification, perform. Note that the same feature extraction method (the combination of CNN and SIFT) is used for this comparison.

5. Conclusion

Infection diseases remain a major public health issue globally. One of the effective ways to detect several life-threatening infectious diseases is using white blood cells. In this paper, we present an approach to detect different types of white blood cells in microscopic images. We used the Gram-Schmidt process for the segmentation step, and for the classification, we used the Scale-Invariant Feature Transform (SIFT) technique along with a convolutional deep neural network. In the classification phase, instead of using a softmax classification method, we utilized a sparse method which improved accuracy of our model to 97.14%. While our work provides promising results, there are some areas for further improvement that future research should explore. The first limitation of our work, like many other researches in this domain, is lack of a benchmark to evaluate and compare our results. Future research should create a benchmark for this domain and analyse how different methods would work in a single data set comparatively. Second, we did not have access to a large enough data set. Increasing the data sample size as well as the variety of the

sample images could also greatly increase the accuracy and generalizability of the model. To increase the data set size and variety with the aim to increase the data independency and the classification accuracy, one potential solution would be to collect databases available in different health centres. Creating such a data set as the benchmark in this domain could be a very big step towards developing methods with higher accuracy and, more importantly, will improve the generalizability of the findings. Lastly, to apply our proposed model on a more complex data set, we can enhance the deep convolutional neural network by increasing the number of layers and the dimension of each layer to meet the complexity of a more complicated system.

Data Availability

The image data used to support the findings of this study have been deposited in the WBCis repository (https://github.com/zxaoyou/segmentation_WBC).

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] M. Toğaçar, B. Ergen, and Z. Cömert, "Classification of white blood cells using deep features obtained from convolutional neural network models based on the combination of feature selection methods," *Applied Soft Computing*, vol. 97, article 106810, 2020.

[2] B. D. Horne, J. L. Anderson, J. M. John et al., "Which white blood cell subtypes predict increased cardiovascular risk?," *Journal of the American College of Cardiology*, vol. 45, no. 10, pp. 1638–1643, 2005.

[3] H. T. Madhloom, S. A. Kareem, H. Ariffin, A. A. Zaidan, H. O. Alanazi, and B. B. Zaidan, "An automated white blood cell nucleus localization and segmentation using image arithmetic and automatic threshold," *Journal of Applied Sciences*, vol. 10, no. 11, pp. 959–966, 2010.

[4] L. B. Holder, M. M. Haque, and M. K. Skinner, "Machine learning for epigenetics and future medical applications," *Epigenetics*, vol. 12, no. 7, pp. 505–514, 2017.

[5] G. Currie, K. E. Hawk, E. Rohren, A. Vial, and R. Klein, "Machine learning and deep learning in medical imaging: intelligent imaging," *Journal of Medical Imaging and Radiation Sciences*, vol. 50, no. 4, pp. 477–487, 2019.

[6] I. Balki, A. Amirabadi, J. Levman et al., "Sample-size determination methodologies for machine learning in medical imaging research: a systematic review," *Canadian Association of Radiologists Journal*, vol. 70, no. 4, pp. 344–353, 2019.

[7] Y. Xu, D. Zhang, J. Yang, and J. Y. Yang, "A two-phase test sample sparse representation method for use with face recognition," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 21, no. 9, pp. 1255–1262, 2011.

[8] J. Wu, P. Zeng, Y. Zhou, and C. Olivier, "A novel color image segmentation method and its application to white blood cell image analysis," in *2006 8th International Conference on Signal Processing*, vol. 2, Guilin, China, 2006, November.

[9] A. Gautam and H. Bhadauria, "Classification of white blood cells based on morphological features," in *2014 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pp. 2363–2368, Delhi, India, 2014.

[10] S. Mohapatra, D. Patra, and S. Satpathy, "Automated leukemia detection in blood microscopic images using statistical texture analysis," in *Proceedings of the 2011 International Conference on Communication, Computing & Security*, pp. 184–187, Rourkela, Odisha, India, 2011.

[11] N. Sinha and A. G. Ramakrishnan, "Automation of differential blood count," in *TENCON 2003. Conference on convergent Technologies for Asia-Pacific Region*, vol. 2, pp. 547–551, Bangalore, India, 2003, October.

[12] J. Laosai and K. Chamnongthai, "Acute leukemia classification by using SVM and K-means clustering," in *2014 International Electrical Engineering Congress (iEECON)*, pp. 1–4, Chonburi, Thailand, 2014, March.

[13] N. Theera-Umpon, "Patch-based white blood cell nucleus segmentation using fuzzy clustering," *ECTI-EEC*, vol. 3, no. 1, pp. 15–19, 2005.

[14] C. Pan, D. S. Park, Y. Yang, and H. M. Yoo, "Leukocyte image segmentation by visual attention and extreme learning machine," *Neural Computing and Applications*, vol. 21, no. 6, pp. 1217–1227, 2012.

[15] B. C. Ko, J. W. Gim, and J. Y. Nam, "Automatic white blood cell segmentation using stepwise merging rules and gradient vector flow snake," *Micron*, vol. 42, no. 7, pp. 695–705, 2011.

[16] M. Hamghalam, M. Motameni, and A. E. Kelishomi, "Leukocyte segmentation in Giemsa-stained image of peripheral blood smears based on active contour," in *2009 International Conference on Signal Processing Systems*, pp. 103–106, Singapore, 2009.

[17] S. H. Rezatofighi, H. Soltanian-Zadeh, R. Sharifian, and R. A. Zoroofi, "A new approach to white blood cell nucleus segmentation based on Gram-Schmidt orthogonalization," in *2009 International Conference on Digital Image Processing*, pp. 107–111, Bangkok, Thailand, 2009.

[18] G. H. Golub and C. F. Van Loan, *Matrix Computations*, Johns Hopkins, 4th ed edition, 2013.

[19] A. Nakib, H. Oulhadj, and P. Siarry, "Microscopic image segmentation with two-dimensional exponential entropy based on hybrid microcanonical annealing," in *MVA2007 IAPR Conference on Machine Vision Applications*, pp. 420–423, Tokyo, JAPAN, 2007, May.

[20] K. Hammouche, M. Diaf, and P. Siarry, "A multilevel automatic thresholding method based on a genetic algorithm for a fast image segmentation," *Computer Vision and Image Understanding*, vol. 109, no. 2, pp. 163–175, 2008.

[21] X. Ma and N. Wu, "Two-phase test sample representation with efficient m-nearest neighbor selection in face recognition," in *International Symposium on Neural Networks*, pp. 221–228, Springer, Berlin, Heidelberg, 2012.

[22] D. Lopez-Puigdollers, V. J. Traver, and F. Pla, "Recognizing white blood cells with local image descriptors," *Expert Systems with Applications*, vol. 115, pp. 695–708, 2019.

[23] A. Vedaldi, *An Open Implementation of the SIFT Detector and Descriptor*, UCLA CSD, 2007.

[24] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[25] R. B. Hegde, K. Prasad, H. Hebbar, and B. M. K. Singh, "Feature extraction using traditional image processing and convolutional neural network methods to classify white blood cells: a

study," *Australasian Physical & Engineering Sciences in Medicine*, vol. 42, no. 2, pp. 627–638, 2019.

[26] C. Zhang, X. Xiao, X. Li et al., "White blood cell segmentation by color-space-based k-means clustering," *Sensors*, vol. 14, no. 9, pp. 16128–16147, 2014.

[27] Z. Liu, J. Pu, M. Xu, and Y. Qiu, "Face recognition via weighted two phase test sample sparse representation," *Neural Processing Letters*, vol. 41, no. 1, pp. 43–53, 2015.

[28] Y. Xu and Q. Zhu, "A simple and fast representation-based face recognition method," *Neural Computing and Applications*, vol. 22, no. 7-8, pp. 1543–1549, 2013.

[29] C. Y. Lu, H. Min, J. Gui, L. Zhu, and Y. K. Lei, "Face recognition via weighted sparse representation," *Journal of Visual Communication and Image Representation*, vol. 24, no. 2, pp. 111–116, 2013.

[30] X. Dong, H. Zhang, L. Zhu et al., "Weighted locality collaborative representation based on sparse subspace," *Journal of Visual Communication and Image Representation*, vol. 58, pp. 187–194, 2019.

[31] D. C. Huang, K. D. Hung, and Y. K. Chan, "A computer assisted method for leukocyte nucleus segmentation and recognition in blood smear images," *Journal of Systems and Software*, vol. 85, no. 9, pp. 2104–2118, 2012.

[32] A. K. Abbas and J. C. Aster, *Robbins and Cotran Pathologic Basis of Disease*, Elsevier/Saunders, 2015.

[33] S. F. Yang-Mao, Y. K. Chan, and Y. P. Chu, "Edge enhancement nucleus and cytoplast contour detector of cervical smear images," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 38, no. 2, pp. 353–366, 2008.

[34] N. M. Salem, "Segmentation of white blood cells from microscopic images using K-means clustering," in *2014 31st National Radio Science Conference (NRSC)*, pp. 371–376, Cairo, Egypt, 2014.

Hindawi
Computational and Mathematical Methods in Medicine
Volume 2021, Article ID 9976440, 7 pages
https://doi.org/10.1155/2021/9976440

Research Article

Texture Analysis in the Assessment of Rectal Cancer: Comparison of T2WI and Diffusion-Weighted Imaging

Ming Li,[1] Xiaodan Xu,[2] Pengjiang Qian,[3] Heng Jiang,[4] Jianlong Jiang,[1] Jinbing Sun,[1] and Zhihua Lu[4]

[1]*Department of General Surgery, Changshu No. 1 People's Hospital, Changshu, 215500 Jiangsu Province, China*
[2]*Department of Gastroenterology, Changshu No. 1 People's Hospital, Changshu, 215500 Jiangsu Province, China*
[3]*School of Artificial Intelligence and Computer Science, Jiangnan University, Wuxi, 214122 Jiangsu Province, China*
[4]*Department of Radiology, Changshu No. 1 People's Hospital, Changshu, 215500 Jiangsu Province, China*

Correspondence should be addressed to Jinbing Sun; sunjb515@163.com and Zhihua Lu; luzhihua4100@suda.edu.cn

Received 22 March 2021; Revised 5 July 2021; Accepted 27 August 2021; Published 15 September 2021

Academic Editor: Alireza Jolfaei

Texture analysis (TA) techniques derived from T2-weighted imaging (T2WI) and apparent diffusion coefficient (ADC) maps of rectal cancer can both achieve good diagnosis performance. This study was to compare TA from T2WI and ADC maps between different pathological T and N stages to confirm which TA analysis is better in diagnosis performance. 146 patients were enrolled in this study. Tumor TA was performed on every patient's T2WI and ADC maps, respectively; then, skewness, kurtosis, uniformity, entropy, energy, inertia, and correlation were calculated. Our results demonstrated that those significant different parameters derived from T2WI had better diagnostic performance than those from ADC maps in differentiating pT3b-4 and pN1-2 stage tumors. In particular, the energy derived from T2WI was an optimal parameter for diagnostic efficiency. High-resolution T2WI plays a key point in the local stage of rectal cancer; thus, TA derived from T2WI may be a more useful tool to aid radiologists and surgeons in selecting treatment.

1. Introduction

The incidence and mortality of colorectal cancer in China have maintained an upward trend, which ranks third and fifth among all malignant tumors. Among them, rectal cancer accounts for half of colorectal cancer in China. [1].

Pathological T and N stages of rectal cancer were important prognostic factors affecting patients [2]. The patients of rectal cancer with lymph node metastasis require neoadjuvant chemotherapy. pT1-2 stage patients do not require neoadjuvant chemotherapy due to a low recurrence rate; conversely, pT3-4 stage patients need neoadjuvant chemotherapy [3, 4]. The prognosis of pT3 stage rectal cancer varies greatly based on extramural depth (EMD) of tumor invasion [5]. The 5 mm cutoff of EMD has been determined to be the most discriminating and simple to use in clinical practice regardless of differences in overall survival and local recurrence [5–7]. Therefore, an accurate preoperative evaluation is essential because individualized treatment measures are required for different local staging patients.

Diffusion-weighted imaging (DWI) could improve contrast between the lesion and normal tissue, which could improve the accuracy of rectal cancer detection. Compared with T2WI alone, T2WI combined with DWI has the best accuracy and specificity for the diagnosis of rectal cancer; moreover, that can improve accuracy of rectal cancer staging [8, 9].

Texture analysis (TA), as a new image postprocessing technology, can quantitatively describe tissue heterogeneity in medical images. About rectal cancer, there are many TA studies involved on T2WI and ADC maps. The study of Yang et al. [10] reported that T2WI histogram parameters could differentiate the positive and negative lymph node. TA derived from T2WI can be used for prediction of the

Table 1: MRI sequences.

MRI sequence	TR (ms)	TE (ms)	TSE factor	Slice thickness (mm)	Gap (mm)	FOV (cm)	Matrix	NSA	*b*-value
Sagittal T2WI	3577	70	20	3	0	$24 \times 24 \times 7.2$	300×266	3	—
Coronal T2WI	3000	75	18	2	0	$18 \times 18 \times 0.4$	300×218	3	—
Axial T2WI	3000	75	18	3	0	$18 \times 18 \times 7.2$	368×273	3	—
DWI	2750	76	—	3	0	$22 \times 22 \times 7.2$	112×108	2	0, 1000

TR: repetition time; TE: echo time; TSE: turbo spin echo; FOV: field of view; NSA: number of signal averaged.

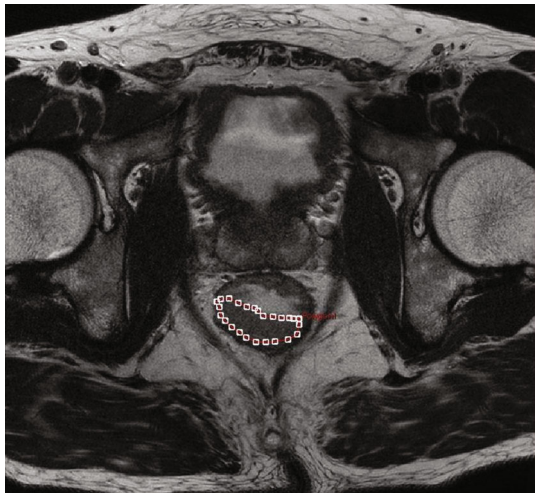


Figure 1: Example image for delineating a rectal lesion on T2WI.

Figure 2: Example image for delineating a rectal lesion tumor edge on the ADC map.

rectal cancer T stage. [11]. In addition, our previous research results showed that TA from ADC maps could predict the local stage of rectal cancer [12].

In the present study, we investigated to compare TA from T2WI and ADC maps between different pathological T and N stages and to evaluate diagnostic performance to further confirm which TA analysis is better in diagnosis performance.

Table 2: Clinical and pathological data of patients.

Characteristics	No. of patients ($n = 146$)
Gender	
Male	86
Female	60
Tumor location	
Upper rectum	55
Middle rectum	66
Low rectum	25
Differentiated grade	
Well differentiated	28
Moderately differentiated	67
Poorly differentiated	51
Pathological T stage	
T1	8
T2	37
T3a	25
T3b	54
T3c	17
T4	5
Pathological N stage	
N0	91
N1	40
N2	15

2. Materials and Methods

2.1. Patients. This retrospective study was approved by the Institutional Review Board of Changshu Hospital of Soochow University. As this study has retrospective nature, requirements for written informed consent were waived. Inclusion criteria were (1) biopsy-proven nonmucinous adenocarcinoma and (2) the complete pathological T and N stage report. Exclusion criteria were (1) motion artifacts and magnetic sensitive artifacts affecting image quality and (2) treatment with preoperative chemotherapy or radiotherapy.

Finally, 146 patients were enrolled in the final study between October 2016 and October 2020.

2.2. MRI. All MRI examinations were performed on 3.0 T scanner (Intera Achieva TX, Philips Medical System) using a phased-array surface coil.

MR sequences included sagittal, oblique axial (perpendicular to the tumor axis), and oblique coronal (parallel to

Table 3: Texture parameters between the pT1-3a and pT3b-4 stages.

Texture parameters	pT1-3a ($n = 70$)	pT3b-4 ($n = 76$)	p value	ICC
T2WI				
Skewness	0.785 ± 0.397	0.871 ± 0.384	0.185	0.863
Kurtosis	1.109 ± 0.425	1.378 ± 0.593	0.002	0.885
Energy ($\times 10^{-3}$)	0.685 (0.580, 0.871)	0.515 (0.452, 0.605)	<0.001	0.847
Entropy	9.226 (8.480, 9.488)	9.590 (9.342, 9.869)	<0.001	0.913
Uniformity	0.812 ± 0.068	0.798 ± 0.073	0.228	0.905
Inertia	159.572 (126.822, 195.283)	163.749 (125.463, 203.394)	0.738	0.865
Correlation ($\times 10^{-3}$)	1.358 ± 0.470	1.304 ± 0.502	0.508	0.886
ADC				
Skewness	0.382 ± 0.311	0.516 ± 0.360	0.018	0.838
Kurtosis	0.661 ± 0.375	0.809 ± 0.390	0.020	0.875
Energy ($\times 10^{-3}$)	0.949 (0.625, 0.693)	0.690 (0.583, 1.445)	0.102	0.912
Entropy	9.672 ± 1.203	10.093 ± 1.235	0.039	0.903
Uniformity	0.761 ± 0.057	0.769 ± 0.070	0.502	0.887
Inertia	922.235 (621.034, 1189.230)	737.177 (520.237, 1060.946)	0.086	0.866
Correlation ($\times 10^{-3}$)	0.448 (0.349, 0.638)	0.482 (0.384, 0.641)	0.267	0.903

Table 4: Texture parameters between the pN0 and pN1-2 stages.

Texture parameters	pN0 ($n = 91$)	pN1-2 ($n = 55$)	p value	ICC
T2WI				
Skewness	0.763 ± 0.380	0.940 ± 0.389	0.007	0.852
Kurtosis	1.253 ± 0.552	1.243 ± 0.510	0.911	0.856
Energy ($\times 10^{-3}$)	0.639 (0.530, 0.786)	0.516 (0.437, 0.613)	<0.001	0.867
Entropy	9.342 (8.788, 9.662)	9.537 (9.135, 9.773)	0.039	0.883
Uniformity	0.810 ± 0.064	0.797 ± 0.081	0.333	0.915
Inertia	163.749 (125.463, 203.394)	157.387 (119.347, 184.395)	0.186	0.856
Correlation ($\times 10^{-3}$)	1.341 ± 0.473	1.312 ± 0.510	0.729	0.863
ADC				
Skewness	0.403 ± 0.283	0.532 ± 0.414	0.045	0.878
Kurtosis	0.682 ± 0.346	0.832 ± 0.437	0.033	0.863
Energy ($\times 10^{-3}$)	0.832 (0.589, 1.493)	0.710 (0.592, 1.503)	0.747	0.902
Entropy	9.863 ± 1.212	9.938 ± 1.278	0.722	0.893
Uniformity	0.762 ± 0.067	0.771 ± 0.060	0.392	0.907
Inertia	830.845 (529.492, 1116.510)	918.167 (593.935, 1180.040)	0.202	0.916
Correlation ($\times 10^{-3}$)	0.484 (0.377, 0.734)	0.458 (0.373, 0.637)	0.330	0.905

the tumor axis) T2WI and oblique axial DWI (Table 1). The position and scanning parameters of oblique axial DWI were consistent with oblique axial T2WI. The ADC map is automatically generated by DWI with a b value of 0 and 1000 s/mm^2.

2.3. Textural Feature Calculation. TA was performed on axial T2WI and ADC maps, respectively, using in-house software (Omni-Kinetics, GE Healthcare) by two authors (Zhihua Lu and Ming Li). The regions of interest (ROIs) involve as much tumor tissue as possible on the largest tumor slice, excluding necrosis, cysts, and gas (Figures 1 and 2). Then, texture features based on T2WI and ADC maps were calculated automatically. The seven texture features that we chose included skewness, kurtosis, and uniformity (first-order statistics) and entropy, energy, inertia, and correlation (second-order statistics).

2.4. Histopathologic Analysis. A pathological T (pT) and N (pN) staging report was performed based on the eighth edition of the American Joint Committee on Cancer (AJCC 8th) [13]. The histopathological report of local T and N staging was referred to as the gold standard in this study.

Table 5: Performance of texture parameters to discriminate high-stage tumors.

Parameters	AUC	p value	95% CI	Sensitivity (%)	Specificity (%)	Cutoff value
Discriminate pT3b-4 stage						
T2-kurtosis	0.638	0.0028	0.555, 0.716	53.9	75.7	>1.332
T2-energy	0.793	<0.0001	0.718, 0.856	60.5	87.1	≤0.528
T2-entropy	0.743	<0.0001	0.664, 0.811	76.3	64.3	>9.329
A-skewness	0.605	0.0248	0.521, 0.685	30.3	91.4	>0.730
A-kurtosis	0.618	0.0121	0.534, 0.697	50.0	77.1	>0.818
A-entropy	0.602	0.0292	0.518, 0.682	75.0	42.9	>9.395
Discriminate pN1-2 stage						
T2-skewness	0.640	0.0029	0.556, 0.718	70.9	54.9	>0.729
T2-energy	0.713	<0.0001	0.632, 0.784	67.3	73.6	≤0.547
T2-entropy	0.602	0.0313	0.518, 0.682	72.7	48.4	>9.288
A-skewness	0.618	0.0243	0.535, 0.698	40.0	90.1	>0.729
A-kurtosis	0.601	0.0489	0.516, 0.681	29.1	93.4	>1.219

Notes: T2-: texture parameters from the T2WI map; A-: texture parameters from the ADC map; CI: confidence interval; AUC: area under the curve.

2.5. Statistical Analysis. First, the Kolmogorov-Smirnov test was used; data were presented as mean ± standard deviation or median ± interquartile range. To assess the difference of texture parameters between pT1-3a vs. pT3b-4 stage and pN0 vs. pN1-2 stage from the T2WI and ADC maps, the t-test or Mann-Whitney U test was performed. Subsequent receiver operating characteristic (ROC) curve analysis was performed with those significant different parameters from the T2WI and ADC maps. The correlations between those significant different parameters from the T2WI and ADC maps and pathological T and N stages were used by the Spearman analysis. Intraclass correlation coefficients (ICCs) were calculated for interobserver agreements. SPSS and MedCalc software was used for statistical analysis. The significant level was set as p value ≤ 0.05.

3. Results

3.1. Clinical and Pathological Findings. 146 patients comprised 86 males (41-89 years; median, 68 years) and 60 females (41-84 years; median, 67 years). Detailed clinical and pathological data are listed in Table 2. According to the design of this study, 70 patients were grouped as pT1-3a and 76 patients were grouped as pT3b-4. 91 patients were grouped as pN0, and 55 patients were grouped as pN1-2.

3.2. Texture Parameters in Differentiating pT1-3a and pT3b-4 Stages. On T2WI, the pT3b-4 stage tumor demonstrated significantly higher kurtosis and entropy and lower energy than the pT1-3a stage tumor (all $p < 0.05$, Table 3). On the ADC map, the pT3b-4 stage tumor demonstrated significantly higher skewness, kurtosis, and entropy than the pT1-3a stage tumor (all $p < 0.05$). No significant differences were found in other TA parameters between different T stages. The results of ICCs are shown in Table 3.

3.3. Texture Parameters in Differentiating pN0 and pN1-2 Stages. On T2WI, the pN1-2 stage tumor demonstrated significantly higher skewness and entropy and lower energy than the pN0 stage tumor (all $p < 0.05$, Table 4). On the ADC map, the pN1-2 stage tumor demonstrated significantly higher skewness and kurtosis than the pN0 stage tumor (all $p < 0.05$). No significant differences were found in other TA parameters between different N stages. The results of ICCs are shown in Table 4.

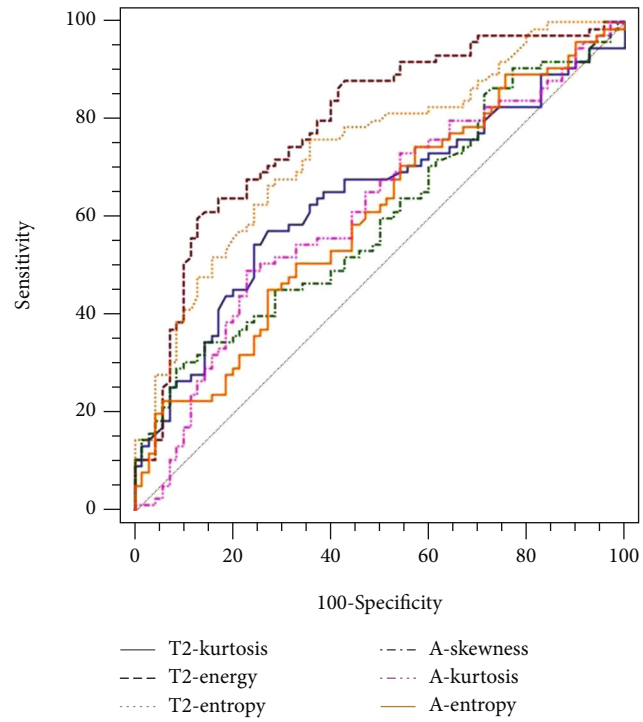


Figure 3: The ROC map of texture parameters in differentiating pT3b-4 from the pT1-3a stage.

3.4. Performance of Texture Parameters to Distinguish the pT3b-4 and pN1-2 Stages. For differentiating pT3b-4 from pT1-3a stage tumors, energy derived from T2WI had the largest AUC (79.3%). Kurtosis derived from the ADC map

Table 6: Correlations between significant different parameters from the T2WI and ADC maps and pathological T and N stages.

Parameters	Pathological T stage		Parameters	Pathological N stage	
	R_s	p value		R_s	p value
T2-kurtosis	0.266	0.001	T2-skewness	0.229	0.005
T2-energy	-0.494	<0.001	T2-energy	-0.359	<0.001
T2-entropy	0.445	<0.001	T2-entropy	0.175	0.035
A-skewness	0.199	0.016	A-skewness	0.196	0.018
A-kurtosis	0.235	0.004	A-kurtosis	0.168	0.043
A-entropy	0.221	0.007			

Notes: T2-: texture parameters from the T2WI map; A-: texture parameters from the ADC map.

had the largest AUC (61.8%) (Table 5, Figure 3). Furthermore, energy derived from T2WI showed a better correlation with the pathological T stage ($R_s = -0.494$) than kurtosis derived from ADC maps (Table 6).

For differentiating pN1-2 from pN0 stage tumors, energy derived from T2WI had the largest AUC (71.3%). Skewness derived from ADC maps had the largest AUC (61.8%) (Figure 4). Furthermore, energy derived from T2WI showed a better correlation with the pathological N stage ($R_s = -0.359$) than skewness derived from ADC maps (Table 6).

4. Discussion

The present study showed that kurtosis, energy, and entropy from T2WI and skewness, kurtosis, and entropy from ADC maps could significantly differentiate pT1-3a and pT3b-4 stages. In addition, we found that skewness, energy, and entropy from T2WI and skewness and kurtosis from ADC maps could significantly differentiate pN0 and pN1-2 stages. Among the aforementioned significant different parameters, those parameters derived from T2WI had higher AUC differentiating pT3b-4 and pN1-2 stage tumors than those derived from ADC maps. Furthermore, those parameters derived from T2WI showed a better correlation with pathological T and N stages than those derived from ADC maps. The results indicate that the diagnostic performance of TA parameters derived from T2WI was better than that derived from ADC maps.

High-resolution T2WI plays a key point in the local stage of rectal cancer. In recent years, some studies applied TA derived from T2WI in order to improve the preoperative stage and predict prognostic factors of rectal cancer [10, 11, 14]. In our study, the pT3b-4 stage tumor demonstrated higher kurtosis and entropy and lower energy than the pT1-3a stage tumor. In addition, the pN1-2 stage tumor demonstrated higher skewness and entropy and lower energy than the pN0 stage tumor. Theoretically, higher entropy and lower energy reflect higher heterogeneity of the lesion [15]. This may interpret our results. Yang et al. [10] have reported that the lymph node-positive group had a significant lower energy and higher entropy than the lymph node-negative group.

DWI could improve contrast between the lesion and normal tissue. It has been widely used in preoperative imaging of rectal cancer because T2WI combined with DWI can improve the accuracy of rectal cancer staging [8]. Our previous research and other studies have shown that the ADC value estimated from DWI was related to tumor aggressiveness [16–18]. TA parameters derived from ADC maps were more sensitive and reliable markers for the accurate operative stage and evaluate the efficacy to treatment in rectal cancer [12, 19]. In our study, the pT3b-4 stage tumor demonstrated higher skewness, kurtosis, and entropy than the pT1-3a stage tumor. In addition, the pN1-2 stage tumor demonstrated higher skewness and kurtosis than the pN0 stage tumor. Higher value of skewness and kurtosis reflects more complexity and heterogeneity in tumors. Our result concluded that skewness and kurtosis derived from ADC maps as imaging biomarkers might be differentiating the positive lymph node of rectal cancer.

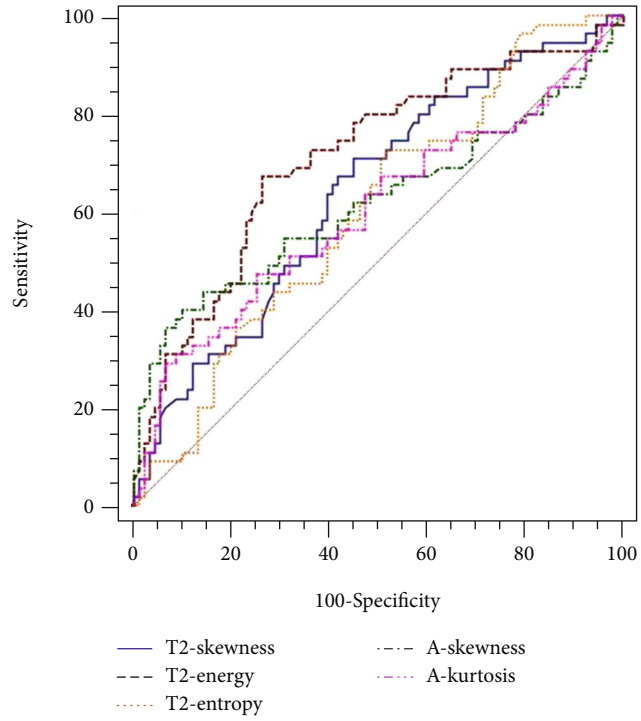


Figure 4: The ROC map of texture parameters in differentiating pN1-2 from the pN0 stage.

ROC curves showed that significant different parameters derived from T2WI had higher AUC differentiating pT3b-4

and pN1-2 stage tumors than those derived from ADC maps. Energy was shown to be the best performance in predicting pT3b-4 and pN1-2 stage tumors among TA parameters with yielding AUC of 79.3% and 71.3%, respectively. Furthermore, the Spearman correlation analysis showed that energy derived from T2WI had the better correlation with pathological T and N stages than parameters derived from the ADC map. Therefore, our results indicated that the diagnostic performance of TA parameters derived from T2WI was better than that derived from ADC maps. The reason may be that high-resolution T2WI can reflect much more details in the image of the lesion. In addition, T2WI is the preferred sequence for the preoperative stage of rectal cancer. Thus, TA derived from T2WI may be a more useful tool to aid radiologists and surgeons in selecting treatment.

There are some limitations. First, rectal cancer often grows infiltrating along the bowel wall. Lesions and normal bowel walls are often unclear; then, the measurement cannot accurately delineate the entire tumor. So, the measurement in our study was performed on a single slice of the largest tumor slice. Second, we evaluated seven TA parameters in our study because these parameters have been proven by previous studies to be related to the heterogeneity of the disease. No other more parameters were included for the group study. Third, we only included pathological T and N stage as the grouping standard. More prognostic factors should be included in the later period for further research.

In conclusion, our study showed that the diagnostic performance of TA parameters derived from T2WI was better than that derived from ADC maps. In particular, the energy derived from T2WI was the optimal parameter for diagnostic efficiency. Thus, TA derived from T2WI may be a more useful tool to aid radiologists and surgeons in selecting treatment.

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 no conflicts of interest.

Authors' Contributions

Ming Li and Xiaodan Xu contributed equally to this work.

Acknowledgments

This work was supported in part by the Jiangsu Provincial Medical Youth Talent under Grant QNRC2016212, Jiangsu Committee of Health under Grant H2018071, Suzhou Clinical Special Disease Diagnosis and Treatment Program under Grant LCZX20182, Suzhou GuSu Medical Talent Project under Grants GSWS2019077 and GSWS2020108, and Suzhou Science and Technology Development Program under Grant SYS2020058.

References

[1] National Health Commission of the People's Republic of China, "Chinese protocol of diagnosis and treatment of colorectal cancer (2020 edition)," *Chinese Journal of Digestive Surgery*, vol. 19, no. 6, pp. 563–588, 2020.

[2] H. J. Schmoll, E. van Cutsem, A. Stein et al., "ESMO consensus guidelines for management of patients with colon and rectal cancer. A personalized approach to clinical decision making," *Annals of Oncology*, vol. 23, no. 10, pp. 2479–2516, 2012.

[3] Oncology Branch of the Chinese Medical Association, Medical and Hospital Administration Bureau of the Health and Family Planning Commission of the People's Republic of China, "Chinese standards for diagnosis and treatment of colorectal cancer (2017 edition) (in Chinese)," *Chinese Journal of Gastrointestinal Surgery*, vol. 21, pp. 92–106, 2018.

[4] R. Glynne-Jones, L. Wyrwicz, E. Tiret et al., "Corrections to "Rectal cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up"," *Annals of Oncology*, vol. 29, article iv263, Supplement 4, 2018.

[5] R. Zinicola, G. Pedrazzi, N. Haboubi, and R. J. Nicholls, "The degree of extramural spread of T3 rectal cancer: an appeal to the American Joint Committee on Cancer," *Colorectal Disease*, vol. 19, no. 1, pp. 8–15, 2017.

[6] S. Merkel, K. Weber, V. Schellerer et al., "Prognostic subdivision of ypT3 rectal tumours according to extension beyond the muscularis propria," *The British Journal of Surgery*, vol. 101, no. 5, pp. 566–572, 2014.

[7] H. Kaur, H. Choi, Y. N. You et al., "MR imaging for preoperative evaluation of primary rectal cancer: practical considerations," *Radiographics*, vol. 32, no. 2, pp. 389–409, 2012.

[8] G. N. Cong, M. W. Qin, H. You et al., "Diffusion weighted imaging combined with magnetic resonance conventional sequences for the diagnosis of rectal cancer," *Zhongguo Yi Xue Ke Xue Yuan Xue Bao. Acta Academiae Medicinae Sinicae*, vol. 31, no. 2, pp. 200–205, 2009.

[9] A. Delli Pizzi, D. Caposiena, D. Mastrodicasa et al., "Tumor detectability and conspicuity comparison of standard b1000 and ultrahigh b2000 diffusion-weighted imaging in rectal cancer," *Abdominal Radiology*, vol. 44, no. 11, pp. 3595–3605, 2019.

[10] L. Yang, D. Liu, X. Fang et al., "Rectal cancer: can T2WI histogram of the primary tumor help predict the existence of lymph node metastasis?," *European Radiology*, vol. 29, no. 12, pp. 6469–6476, 2019.

[11] H. C. Lu, F. Wang, and J. D. Yin, "Texture analysis based on sagittal fat-suppression and transverse T2-weighted magnetic resonance imaging for determining local invasion of rectal cancer," *Frontiers in Oncology*, vol. 10, p. 1476, 2020.

[12] Z. Lu, L. Wang, K. Xia et al., "Prediction of clinical pathologic prognostic factors for rectal adenocarcinoma: volumetric texture analysis based on apparent diffusion coefficient maps," *Journal of Medical Systems*, vol. 43, no. 12, p. 331, 2019.

[13] M. B. Amin, S. Edge, F. Greene et al., *AJCC Cancer Staging Manual*, Springer, New York, 8th edition, 2016.

[14] A. Delli Pizzi, A. M. Chiarelli, P. Chiacchiaretta et al., "MRI-based clinical-radiomics model predicts tumor response before treatment in locally advanced rectal cancer," *Scientific Reports*, vol. 11, no. 1, p. 5379, 2021.

[15] R. Duvauferrier, J. Bezy, V. Bertaud, G. Toussaint, J. Morelli, and J. Lasbleiz, "Texture analysis software: integration with a

radiological workstation," *Studies in Health Technology and Informatics*, vol. 180, pp. 1030–1034, 2012.

[16] L. Curvo-Semedo, D. M. Lambregts, M. Maas, G. L. Beets, F. Caseiro-Alves, and R. G. Beets-Tan, "Diffusion-weighted MRI in rectal cancer: apparent diffusion coefficient as a potential noninvasive marker of tumor aggressiveness," *Journal of Magnetic Resonance Imaging*, vol. 35, no. 6, pp. 1365–1371, 2012.

[17] E. M. Hecht, M. Z. Liu, M. R. Prince et al., "Can diffusion-weighted imaging serve as a biomarker of fibrosis in pancreatic adenocarcinoma?," *Journal of Magnetic Resonance Imaging*, vol. 46, no. 2, pp. 393–402, 2017.

[18] E. Nerad, A. Delli Pizzi, D. M. J. Lambregts et al., "The apparent diffusion coefficient (ADC) is a useful biomarker in predicting metastatic colon cancer using the ADC-value of the primary tumor," *PLoS One*, vol. 14, no. 2, article e0211830, 2019.

[19] L. Yang, M. Qiu, C. Xia et al., "Value of high-resolution DWI in combination with texture analysis for the evaluation of tumor response after preoperative chemoradiotherapy for locally advanced rectal cancer," *AJR. American Journal of Roentgenology*, vol. 12, pp. 1–8, 2019.

Hindawi
Computational and Mathematical Methods in Medicine
Volume 2021, Article ID 6648959, 18 pages
https://doi.org/10.1155/2021/6648959

Research Article

Stability and Hopf Bifurcation Analysis of a Vector-Borne Disease Model with Two Delays and Reinfection

Yanxia Zhang,[1] **Long Li**,[1] **Junjian Huang**,[2] **and Yanjun Liu**[3]

[1]*School of Mathematics and Information Engineering, Chongqing University of Education, Chongqing 400065, China*
[2]*Chongqing Key Laboratory of Nonlinear Circuits and Intelligent Information Processing, College of Electronic and Information Engineering, Southwest University, Chongqing 400715, China*
[3]*School of Mathematical Sciences, Chongqing Normal University, Chongqing 401131, China*

Correspondence should be addressed to Yanxia Zhang; zhangyx@cque.edu.cn

Received 30 November 2020; Revised 9 May 2021; Accepted 20 May 2021; Published 8 June 2021

Academic Editor: Alireza Jolfaei

In this paper, a vector-borne disease model with two delays and reinfection is established and considered. First of all, the existence of the equilibrium of the system, under different cases of two delays, is discussed through analyzing the corresponding characteristic equation of the linear system. Some conditions that the system undergoes Hopf bifurcation at the endemic equilibrium are obtained. Furthermore, by employing the normal form method and the center manifold theorem for delay differential equations, some explicit formulas used to describe the properties of bifurcating periodic solutions are derived. Finally, the numerical examples and simulations are presented to verify our theoretical conclusions. Meanwhile, the influences of the degree of partial protection for recovered people acquired by a primary infection on the endemic equilibrium and the critical values of the two delays are analyzed.

1. Introduction

Vector-borne diseases (VBDs) are one of the complex infectious diseases that endanger human beings. Vectors are living biological agents, such as ticks, mosquitoes, and fleas, with the ability to transmit parasites, bacteria, or viruses between people or from animals to people. It is reported that VBDs cause more than 1 billion infections and 1 million deaths worldwide every year [1]. Malaria is the most prevalent parasitic vector-borne disease caused by plasmodium parasitizing the human body [2–5]. Plasmodium enters stem cells through the blood to parasitize and reproduce. It then invades red blood cells to reproduce after maturity, causing red blood cells to burst in batches and attack. The source of malaria infection is malaria patients and those with Plasmodium. The natural transmission medium of malaria is female mosquitoes of the genus Anopheles, which is transmitted by biting the human body, and a few by blood transfusion and vertical transmission, with Anopheles gambiae being a major carrier of the disease [2, 5, 6]. Different types of malaria have different incubation periods, some of which are about 7-12 days, and some of which are more than 6 months. The morbidity and mortality of malaria are high [7]. According to the latest WHO reports, in 2018, there were estimated 228 million malaria cases worldwide, of which the death toll was about 405,000. It has led to great global economic and social losses, especially in the tropical and subtropical regions on five continents [3].

The population is generally susceptible to malaria. Although there is a certain degree of immunity after several infections, its acquisition process is so slow that it may take years or decades to develop, and it gradually weakens over time [5, 7, 8]. The remaining plasmodium may escape from the immune function due to the antigenic variation and reproduce again (recrudescence); even if the parasite has been eliminated by human immunity or drugs after the initial onset of malaria, the possibility of relapse is not excluded with time. That is, reinfection likely occurs [6, 7]. It is shown that time delay is of great significance in many biological modelling, and its change may affect the dynamic behavior of the system [4–6, 9–19], such as stability, uniqueness, and

oscillation of solution. So, it is instructive to consider a mathematical model with time delays to research the influence of immunity on disease control of malaria transmission.

Considering the incubation period of virus transmission in the vector population, some vector-borne epidemic models with a delay were considered [4–6]. In [5], Xu and Zhou proposed a delayed vector-borne epidemic model and reinfection, investigated its existence and stability of equilibrium, and analyzed its dynamical behavior. It suggested that there are two effective preventive measures to reduce infections: one is to minimize vector to human contacts and the other is to use insecticides to control vector. According to [4, 5], a vector-borne disease model with delay-saturated infection rate and cure rate was given by [6]. The existence and local stability of the epidemic equilibrium were discussed, and the length of the delay of the system preserving stability was estimated. This paper develops an improved vector-borne disease model with two delays and reinfection to consider the time required for the malaria virus to spread to the host population and vector population.

The remainder of this paper is organized as follows. In Section 2, an improved vector-borne disease model is formulated. In Section 3, the stability of the equilibrium and the existence of local Hopf bifurcation are discussed. In Section 4, some explicit formulas determining properties bifurcating periodic solutions are obtained by employing the normal form method and the center manifold theorem for the delay differential equations developed by Hassard et al. [20]. Some numerical examples and simulations are performed in Section 5 to demonstrate the main theoretical results, and the conclusions of this paper are summarized in Section 6.

2. Model Description

Generally, in the mathematical model of a vector-transmitted disease, the host population size at time t, denoted by $N_1(t)$, is divided into three subclasses: susceptible, infected, and recovered, with numbers denoted by $S(t)$, $I(t)$, and $R(t)$, respectively. The vector population size at time t, given by $N_2(t)$, is partitioned into two subclasses: susceptible vectors $M(t)$ and infectious vectors $V(t)$. Based on the models [4–6], we consider an improved vector-borne disease model with two delays and reinfection as follows:

$$\begin{cases} \dfrac{dS}{dt} = \Lambda_1 - b\beta_1 S(t-\tau_1)V(t-\tau_1) - \mu_1 S(t), \\ \dfrac{dI}{dt} = b\beta_1 S(t-\tau_1)V(t-\tau_1) + \sigma b\beta_1 R(t)V(t) - (\mu_1+\gamma)I(t), \\ \dfrac{dR}{dt} = \gamma I(t) - \sigma b\beta_1 R(t)V(t) - \mu_1 R(t), \\ \dfrac{dM}{dt} = \Lambda_2 - b\beta_2 M(t-\tau_2)I(t-\tau_2) - \mu_2 M(t), \\ \dfrac{dV}{dt} = b\beta_2 M(t-\tau_2)I(t-\tau_2) - \mu_2 V(t), \end{cases} \tag{1}$$

where Λ_1 and Λ_2 are the recruitment rates of the hosts and vectors, respectively. b is the average number of bites per mosquito per day. β_1 and β_2 represent the infection rates from vector to human and human to vector, respectively. μ_1 and μ_2 are the natural death rates of the hosts and vectors, respectively. $\sigma(0 \le \sigma \le 1)$ is the degree of partial protection of individuals that recovered from primary infection. γ is the recovery rate per capita of the infective host population. τ_1 and τ_2 are two delays, in which τ_1 represents the incubation period of the host population, and τ_2 denotes the incubation period of the vector population. The term $b\beta_1 S(t-\tau_1)V(t-\tau_1)$ represents the incidence number of the susceptible host infections caused by the infective vector at time $t-\tau_1$. $b\beta_2 M(t-\tau_2)I(t-\tau_2)$ denotes the number of infections of the susceptible vectors that bite the infected host at time $t-\tau_2$ and become infective at time t.

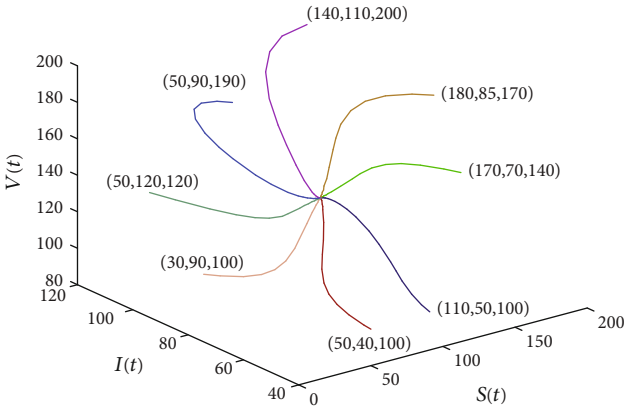


FIGURE 1: The phase graph of system (2) with $\tau_1 = \tau_2 = 0$.

According to [21], system (1) is equivalent to the following model:

$$\begin{cases} \dfrac{dS}{dt} = \Lambda_1 - b\beta_1 S(t-\tau_1)V(t-\tau_1) - \mu_1 S(t), \\ \dfrac{dI}{dt} = b\beta_1 S(t-\tau_1)V(t-\tau_1) + \sigma b\beta_1 \left(\dfrac{\Lambda_1}{\mu_1} - S(t) - I(t)\right)V(t) - (\mu_1+\gamma)I(t), \\ \dfrac{dV}{dt} = b\beta_2 \left(\dfrac{\Lambda_2}{\mu_2} - V(t-\tau_2)\right)I(t-\tau_2) - \mu_2 V(t). \end{cases} \tag{2}$$

Notice that $N_i'(t) = \Lambda_i - \mu_i N_i(t)$, $i = 1, 2$, and its solution is $N_i(t) = N_i(0)e^{-\mu_i t} + (\Lambda_i/\mu_i)(1 - e^{-\mu_i t})$.

Obviously, $\lim_{t\to\infty} N_i(t) = \Lambda_i/\mu_i$, $(i = 1, 2)$. So, for $t \to \infty$ and any $\varepsilon > 0$, $0 \le S + I \le N_1(t) < (\Lambda_1/\mu_1) + \varepsilon$ holds by limit theorem, which follows that $0 \le S + I \le \Lambda_1/\mu_1$. Similarly, $0 \le V \le \Lambda_2/\mu_2$ also holds. Thus, all the solutions of system (2) enter into the region:

$$\Omega = \left\{ (S, I, V) \in R_+^3 \middle| 0 \le S + I \le \frac{\Lambda_1}{\mu_1},\ 0 \le V \le \frac{\Lambda_2}{\mu_2},\ S \ge 0, I \ge 0 \right\}. \tag{3}$$

For the existence of the equilibrium of system (2), according to [5], one can have the following result.

Lemma 1. *For system (2), $E^0 = (\Lambda_1/\mu_1, 0, 0)$ is a disease-free equilibrium if $R_0 \le 1$, and $E^* = (S^*, I^*, V^*)$ is an endemic equilibrium if $R_0 > 1$, where $R_0 = (b^2\Lambda_1\Lambda_2\beta_1\beta_2)/(\mu_1\mu_2{}^2(\mu_1 + \gamma))$ is the basic reproductive number, $S^* = (\Lambda_1\mu_2(b\beta_2 I^* + \mu_2))/(b^2\Lambda_2\beta_1\beta_2 I^* + \mu_1\mu_2(b\beta_2 I^* + \mu_2))$, $V^* = (b\Lambda_2\beta_2 I^*)/(\mu_2(b\beta_2 I^* + \mu_2))$, and I^* satisfy the following equation:*

$$P_2(I^*)^2 + P_1 I^* + P_0 = 0, \tag{4}$$

where

$$\begin{aligned} P_2 &= \mu_1 b^2\beta_2^2(b\sigma\beta_1\Lambda_2 + \mu_1\mu_2)(b\Lambda_2\beta_1 + \mu_1\mu_2) \\ &\quad + \gamma b^2\beta_2^2\mu_1\mu_2(b\Lambda_2\beta_1 + \mu_1\mu_2), \end{aligned}$$

$$\begin{aligned} P_1 &= b\beta_2\mu_1\mu_2^2(\mu_1 + \gamma)(b\Lambda_2\beta_1 + \mu_1\mu_2) \\ &\quad + b\beta_2\mu_1{}^2\mu_2{}^2(b\sigma\beta_1\Lambda_2 + \mu_1\mu_2) + b\gamma\beta_2\mu_1{}^2\mu_2{}^3 \\ &\quad - b^3\Lambda_1\Lambda_2\beta_1\beta_2{}^2(b\sigma\beta_1\Lambda_2 + \mu_1\mu_2), \end{aligned}$$

$$P_0 = \mu_1{}^2\mu_2{}^4(\mu_1 + \gamma)(1 - R_0). \tag{5}$$

The disease-free equilibrium E^0 denotes no infection, and the endemic equilibrium E^* represents that the disease will exist and persist. The basic reproductive number R_0 describes the expected number of secondary infections, which is mainly sensitive to parameters b and β_1 but not affected by parameter σ [5]. The process of deriving R_0 using the next-generation method [22–24] is presented in Appendix A.

3. Stability of Endemic Equilibrium and Hopf Bifurcation

The Jacobian matrix of the linear system (2) at $E^* = (S^*, I^*, V^*)$ is

$$J = \begin{pmatrix} -\mu_1 - b\beta_1 V^* e^{-\lambda\tau_1} & 0 & -b\beta_1 S^* e^{-\lambda\tau_1} \\ -\left(\sigma - e^{-\lambda\tau_1}\right) b\beta_1 V^* & -\mu_1 - \gamma - b\sigma\beta_1 V^* & b\beta_1 S^* e^{-\lambda\tau_1} + b\sigma\beta_1\left(\dfrac{\Lambda_1}{\mu_1} - I^* - S^*\right) \\ 0 & b\beta_2\left(\dfrac{\Lambda_2}{\mu_2} - V^*\right)e^{-\lambda\tau_2} & -\mu_2 - b\beta_2 I^* e^{-\lambda\tau_2} \end{pmatrix}, \tag{6}$$

and the characteristic equation $\det(\lambda I - J) = 0$ is

$$\begin{aligned} &\lambda^3 + m_2\lambda^2 + m_1\lambda + m_0 + \left(n_2\lambda^2 + n_1\lambda + n_0\right)e^{-\lambda\tau_1} \\ &\quad + \left(q_2\lambda^2 + q_1\lambda + q_0\right)e^{-\lambda\tau_2} + (r_1\lambda + r_0)e^{-\lambda(\tau_1+\tau_2)} = 0, \end{aligned} \tag{7}$$

where

$$\begin{aligned} m_2 &= \mu_2 + 2\mu_1 + \gamma + b\sigma\beta_1 V^*, \\ m_1 &= (\mu_1 + \gamma + b\sigma\beta_1 V^*)(\mu_1 + \mu_2) + \mu_1\mu_2, \\ m_0 &= \mu_1\mu_2(\mu_1 + \gamma + b\sigma\beta_1 V^*), \\ n_2 &= b\beta_1 V^*, \\ n_1 &= b\beta_1 V^*(\mu_2 + \mu_1 + \gamma + b\sigma\beta_1 V^*), \\ n_0 &= b\beta_1 V^*(\mu_1 + \gamma + b\sigma\beta_1 V^*)\mu_2, \\ q_2 &= b\beta_2 I^*, \\ q_1 &= b\beta_2 I^*(2\mu_1 + \gamma + b\sigma\beta_1 V^*) - b\sigma\beta_1\left(\frac{\Lambda_1}{\mu_1} - I^* - S^*\right)b\beta_2 \\ &\quad \cdot \left(\frac{\Lambda_2}{\mu_2} - V^*\right), \\ q_0 &= b\beta_2 I^*\mu_1(\mu_1 + \gamma + b\sigma\beta_1 V^*) - b\sigma\beta_1\left(\frac{\Lambda_1}{\mu_1} - I^* - S^*\right)b\beta_2 \\ &\quad \cdot \left(\frac{\Lambda_2}{\mu_2} - V^*\right)\mu_1, \\ r_1 &= b^2\beta_1\beta_2 V^* I^* - b\beta_2\left(\frac{\Lambda_2}{\mu_2} - V^*\right)b\beta_1 S^*, \\ r_0 &= b^2\beta_1\beta_2 V^* I^*(\mu_1 + \gamma + b\sigma\beta_1 V^*) - b^3\beta_1^2\beta_2 S^*\sigma V^*\left(\frac{\Lambda_2}{\mu_2} - V^*\right) \\ &\quad - \left[b^3\beta_1^2\beta_2\sigma V^*\left(\frac{\Lambda_1}{\mu_1} - I^* - S^*\right)\left(\frac{\Lambda_2}{\mu_2} - V^*\right)\right. \\ &\quad \left. + \mu_1 b^2\beta_1\beta_2 S^*\left(\frac{\Lambda_2}{\mu_2} - V^*\right)\right]. \end{aligned} \tag{8}$$

Case (1). $\tau_1 = \tau_2 = 0$.

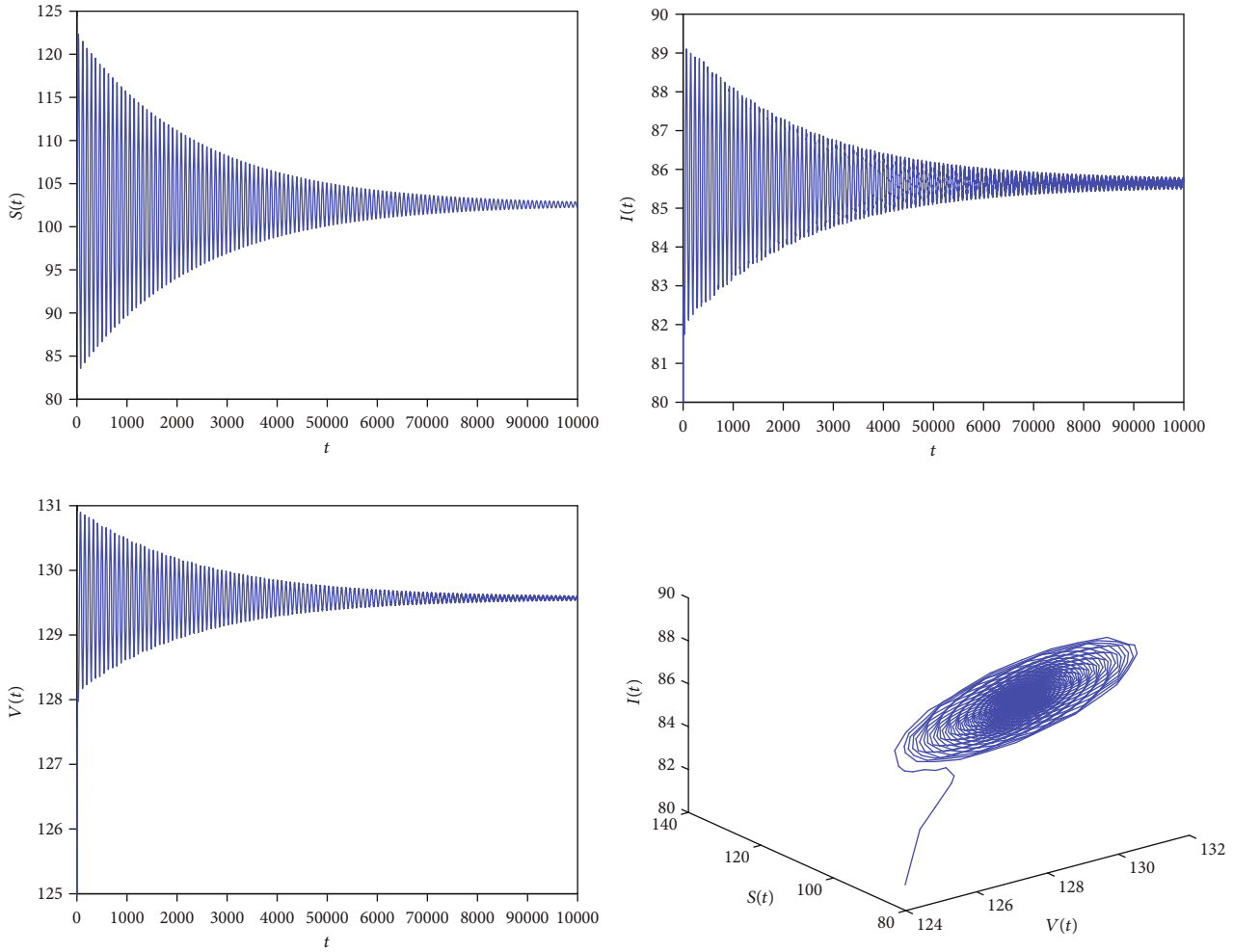


FIGURE 2: The endemic equilibrium E^* is asymptotically stable with $\tau_1 = 23 < \tau_{10}$.

Characteristic equation (7) becomes

$$\lambda^3 + A_{12}\lambda^2 + A_{11}\lambda + A_{10} = 0, \tag{9}$$

where $A_{12} = m_2 + n_2 + q_2$, $A_{11} = m_1 + n_1 + q_1 + r_1$, and $A_{10} = m_0 + n_0 + q_0 + r_0$.

When the following condition (H1) holds, all roots of equation (9) have negative real parts.

$$\text{(H1)}: A_{12} > 0, A_{11} > 0, A_{10} > 0, A_{12}A_{11} - A_{10} > 0. \tag{10}$$

Hence, according to the Routh-Hurwitz criterion, the following conclusion can be drawn.

Theorem 2. *If (H1) holds, the endemic equilibrium E^* is locally asymptotically stable when $\tau_1 = \tau_2 = 0$.*

Case (2). $\tau_1 > 0, \tau_2 = 0$.

The characteristic equation (7) becomes

$$\lambda^3 + A_{22}\lambda^2 + A_{21}\lambda + A_{20} + \left(B_{22}\lambda^2 + B_{21}\lambda + B_{20}\right)e^{-\lambda\tau_1} = 0, \tag{11}$$

where $A_{22} = m_2 + q_2$, $A_{21} = m_1 + q_1$, $A_{20} = m_0 + q_0$, $B_{22} = n_2$, $B_{21} = n_1 + r_1$, and $B_{20} = n_0 + r_0$.

Suppose $\lambda = i\omega_1(\omega_1 > 0)$ is a root of equation (11), replacing it into equation (11) and separating the real and imaginary parts, then, we can obtain

$$\begin{cases} A_{22}{\omega_1}^2 - A_{20} = \left(B_{20} - B_{22}{\omega_1}^2\right)\cos\omega_1\tau_1 + B_{21}\omega_1 \sin\omega_1\tau_1, \\ {\omega_1}^3 - A_{21}\omega_1 = -\left(B_{20} - B_{22}{\omega_1}^2\right)\sin\omega_1\tau_1 + B_{21}\omega_1\cos\omega_1\tau_1. \end{cases} \tag{12}$$

It follows that

$${\omega_1}^6 + \left({A_{22}}^2 - 2A_{21} - {B_{22}}^2\right){\omega_1}^4 + \left({A_{21}}^2 - {B_{21}}^2 - 2A_{20}A_{22} + 2B_{20}B_{22}\right){\omega_1}^2 + {A_{20}}^2 - {B_{20}}^2 = 0. \tag{13}$$

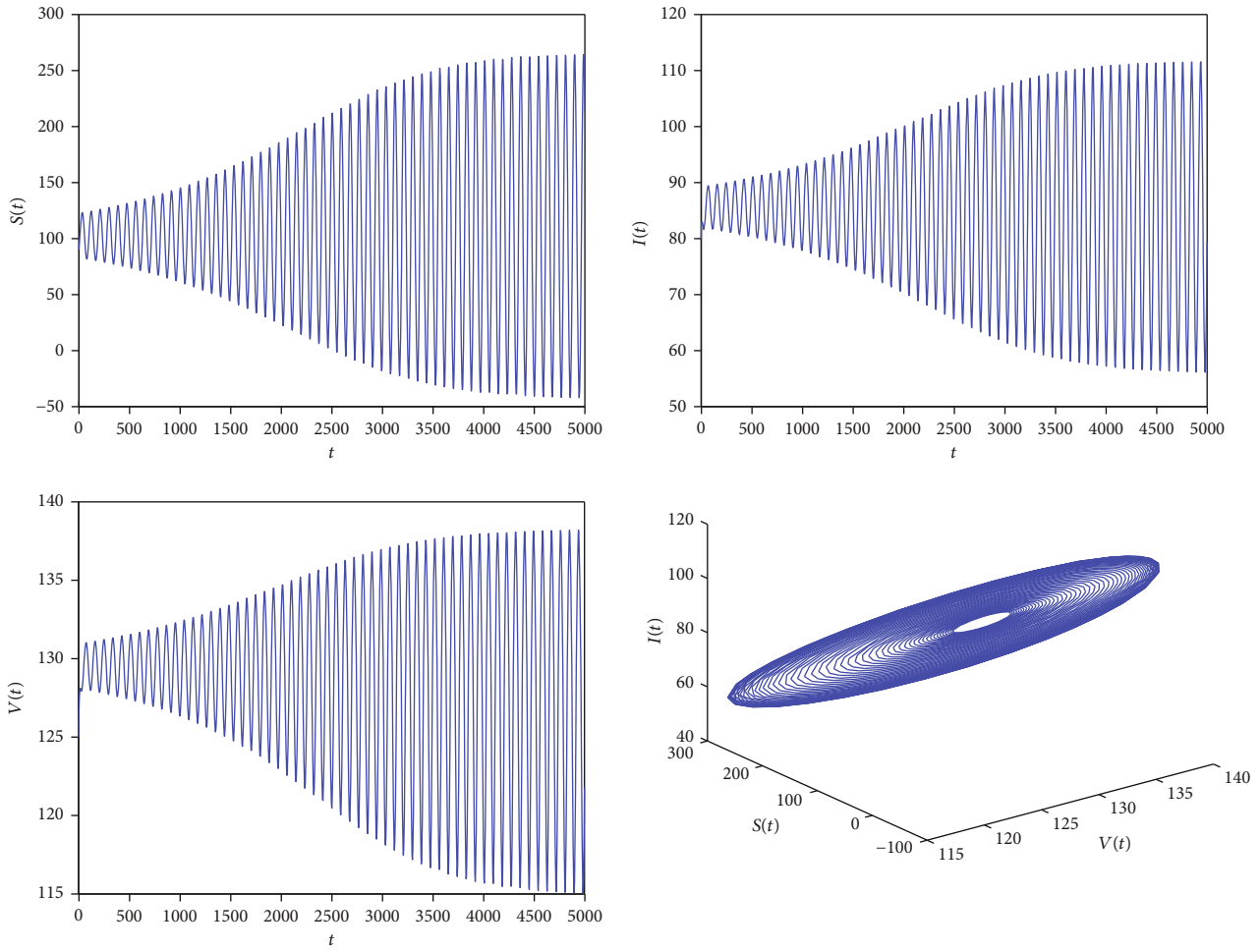


Figure 3: The endemic equilibrium E^* is unstable with $\tau_1 = 24 > \tau_{10}$.

Let $z = \omega_1^2$, $p_2 = A_{22}^2 - 2A_{21} - B_{22}^2$, $p_1 = A_{21}^2 - B_{21}^2 - 2A_{20}A_{22} + 2B_{20}B_{22}$, and $p_0 = A_{20}^2 - B_{20}^2$, then, equation (13) is equivalent to

$$f(z) = z^3 + p_2 z^2 + p_1 z + p_0 = 0. \tag{14}$$

It is easy to get that the two roots of $f'(z) = 3z^2 + 2p_2 z + p_1 = 0$ are $z_1 = (1/3)(-p_2 + \sqrt{p_2^2 - 3p_1})$ and $z_2 = (1/3)(-p_2 - \sqrt{p_2^2 - 3p_1})$.

According to [25, 26], we can get the conditions that equation (14) exists positive roots.

Lemma 3. *For equation (14),*

(i) If $p_0 < 0$, then equation (14) has at least one positive root

(ii) If $p_0 \geq 0$ and $\Delta_1 = p_2^2 - 3p_1 \leq 0$, then equation (14) has no positive roots

(iii) If $p_0 \geq 0$ and $\Delta_1 = p_2^2 - 3p_1 > 0$, then equation (14) has positive roots if and only if $z_1 = (1/3)(-p_2 + \sqrt{\Delta_1}) > 0$ and $f(z_1) \leq 0$

By Lemma 3, it is easy to see that the stability of E^* will not change when τ changes if (ii) is set up. However, the stability may change when τ crosses through some critical values if equation (14) has a positive root.

Assume that z_1, z_2, and z_3 are three positive roots of equation (14), then, there are three positive numbers $\omega_{11} = \sqrt{z_1}$, $\omega_{12} = \sqrt{z_2}$, and $\omega_{13} = \sqrt{z_3}$. It follows from equation (12) that

$$\cos \omega_1 \tau_1 = \frac{\left(A_{22}\omega_1^2 - A_{20}\right)\left(B_{20} - B_{22}\omega_1^2\right) + B_{21}\omega_1\left(\omega_1^3 - A_{21}\omega_1\right)}{\left(B_{20} - B_{22}\omega_1^2\right)^2 + B_{21}^2\omega_1^2}. \tag{15}$$

Denote

$$\tau_{1k}^{(j)} = \frac{1}{\omega_{1k}} \arccos\left\{\frac{\left(A_{22}\omega_1^2 - A_{20}\right)\left(B_{20} - B_{22}\omega_1^2\right) + B_{21}\omega_1\left(\omega_1^3 - A_{21}\omega_1\right)}{\left(B_{20} - B_{22}\omega_1^2\right)^2 + B_{21}^2\omega_1^2}\right\} + \frac{2j\pi}{\omega_{1k}}, \tag{16}$$

where $j = 0, 1, 2, \cdots$, $k = 1, 2, 3$, then, $\pm i\omega_{1k}$ is a pair of purely imaginary roots of equation (11) for $\tau_1 = \tau_{1k}^{(j)}$. Let $\tau_{10} =$

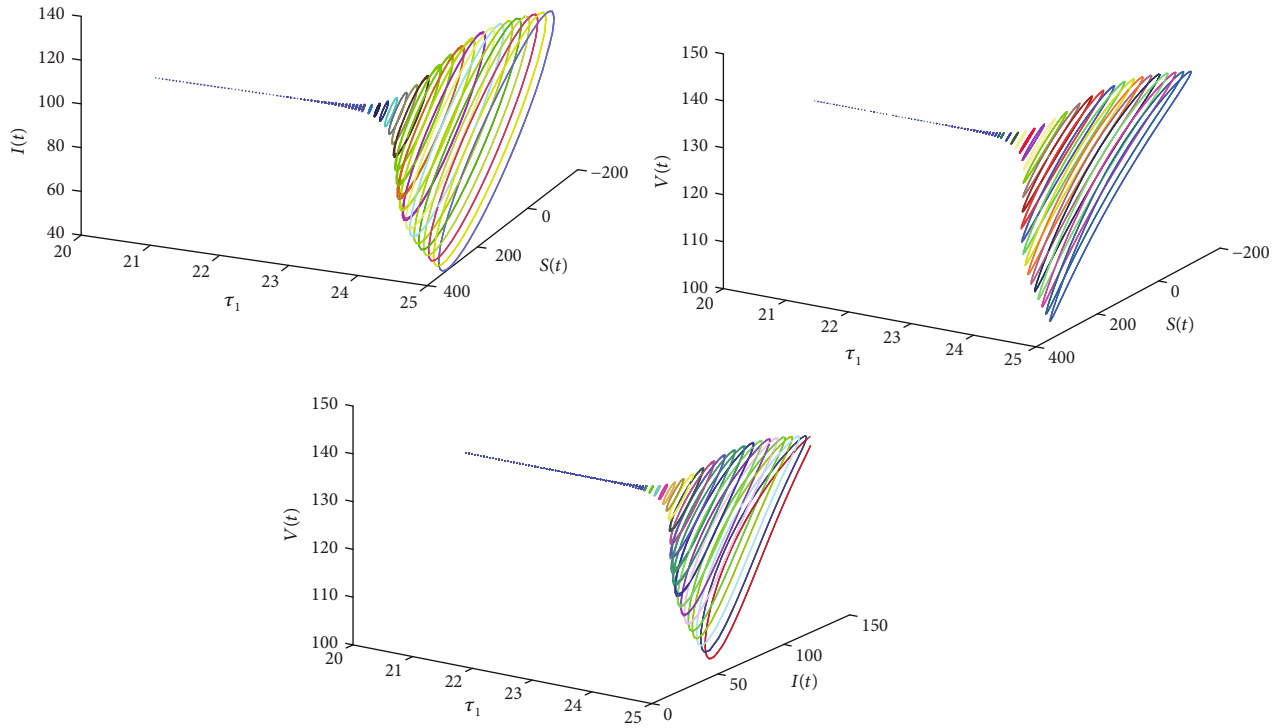


FIGURE 4: The bifurcation diagrams with τ_1 as parameter and $\tau_2 = 0$. A Hopf bifurcation exists at $\tau_1 = \tau_{10}$.

$\min_{k\in\{1,2,3\}}\{\tau_{1k}^{(0)}\}$, ω_{10} be the ω_{1k} that corresponds to the minimum $\tau_{1k}^{(0)}$, and $\lambda(\tau_1) = \alpha(\tau_1) + i\omega(\tau_1)$ be the root of equation (11) near $\tau_1 = \tau_{1k}^{(j)}$ satisfying $\alpha(\tau_{1k}^{(j)}) = 0$ and $\omega(\tau_{1k}^{(j)}) = \omega_{1k}$.

Furthermore, we consider the transversality condition. By differentiating both sides of equation (11) with respect to τ_1, we get

$$\left[\frac{d\lambda}{d\tau_1}\right]^{-1} = \frac{3\lambda^2 + 2A_{22}\lambda + A_{21}}{-\lambda\left(\lambda^3 + A_{22}\lambda^2 + A_{21}\lambda + A_{20}\right)} + \frac{2B_{22}\lambda + B_{21}}{\lambda\left(B_{22}\lambda^2 + B_{21}\lambda + B_{20}\right)} - \frac{\tau_1}{\lambda}. \tag{17}$$

It follows that

$$\begin{aligned}\operatorname{Re}\left\{\left[\frac{d\lambda}{d\tau_1}\right]^{-1}_{\tau_1=\tau_{10}}\right\} &= \operatorname{Re}\left\{\frac{\left(3\lambda^2 + 2A_{22}\lambda + A_{21}\right)e^{\lambda\tau_1} + 2B_{22}\lambda + B_{21}}{\lambda\left(B_{22}\lambda^2 + B_{21}\lambda + B_{20}\right)} - \frac{\tau_1}{\lambda}\right\}_{\tau_1=\tau_{10}} \\ &= \frac{\left[3\omega_1{}^4 + \left(2A_{22}^2 - 4A_{21} - 2B_{22}^2\right)\omega_1{}^2 + A_{21}^2 - 2A_{20}A_{22} - B_{21}^2 + 2B_{20}B_{22}\right]\omega_1{}^2}{B_{21}^2\omega_1{}^4 + \left(B_{20} - B_{22}\omega_1{}^2\right)^2\omega_1{}^2} \\ &= \frac{\omega_1{}^2 f'\left(\omega_1{}^2\right)}{B_{21}^2\omega_1{}^4 + \left(B_{20} - B_{22}\omega_1{}^2\right)^2\omega_1{}^2}.\end{aligned} \tag{18}$$

Thus,

$$\operatorname{sign}\left\{\frac{d(\operatorname{Re}\lambda)}{d\tau_1}\bigg|_{\tau_1=\tau_{10}}\right\} = \operatorname{sign}\left\{\operatorname{Re}\left\{\left[\frac{d\lambda}{d\tau_1}\right]^{-1}_{\tau_1=\tau_{10}}\right\}\right\} = \operatorname{sign}\left\{f'\left(\omega_1^2\right)\right\}. \tag{19}$$

The transversality condition is satisfied when $f'(\omega_1^2) \neq 0$.

According to [27], and applying Lemma 3 and the above transversal condition to system (2), the following theorem is obtained.

Theorem 4. *For system (2), when $\tau_2 = 0$,*

(i) if $p_0 \geq 0$ and $\Delta_1 = p_2^2 - 3p_1 \leq 0$ hold, then, the endemic equilibrium E^* is locally asymptotically stable for all $\tau_1 > 0$

(ii) if $p_0 < 0$ or $p_0 \geq 0$, $\Delta_1 = p_2^2 - 3p_1 > 0$, $z_1 > 0$, and $f(z_1) \leq 0$, then, the endemic equilibrium E^* is locally asymptotically stable for $\tau_1 \in (0, \tau_{10})$

(iii) if the conditions of (ii) and $f'(z_1) \neq 0$ hold, then, it undergoes a Hopf bifurcation at E^* when $\tau_1 = \tau_{10}$

Case (3). $\tau_1 > 0, \tau_2 > 0$.

Fixing τ_1 in its stable interval and regarding τ_2 as a parameter, we analyze the roots of characteristic equation (7). Let $\lambda = i\omega(\omega > 0)$ be a characteristic root of equation (7). By performing some calculations as those in case (2), we can obtain

$$\begin{aligned}&\left(\lambda^3 + m_2\lambda^2 + m_1\lambda + m_0\right)e^{\lambda\tau_2} + \left(n_2\lambda^2 + n_1\lambda + n_0\right)e^{-\lambda\tau_1}e^{\lambda\tau_2} \\ &\quad + \left(q_2\lambda^2 + q_1\lambda + q_0\right) + \left(r_1\lambda + r_0\right)e^{-\lambda\tau_1} = 0,\end{aligned} \tag{20}$$

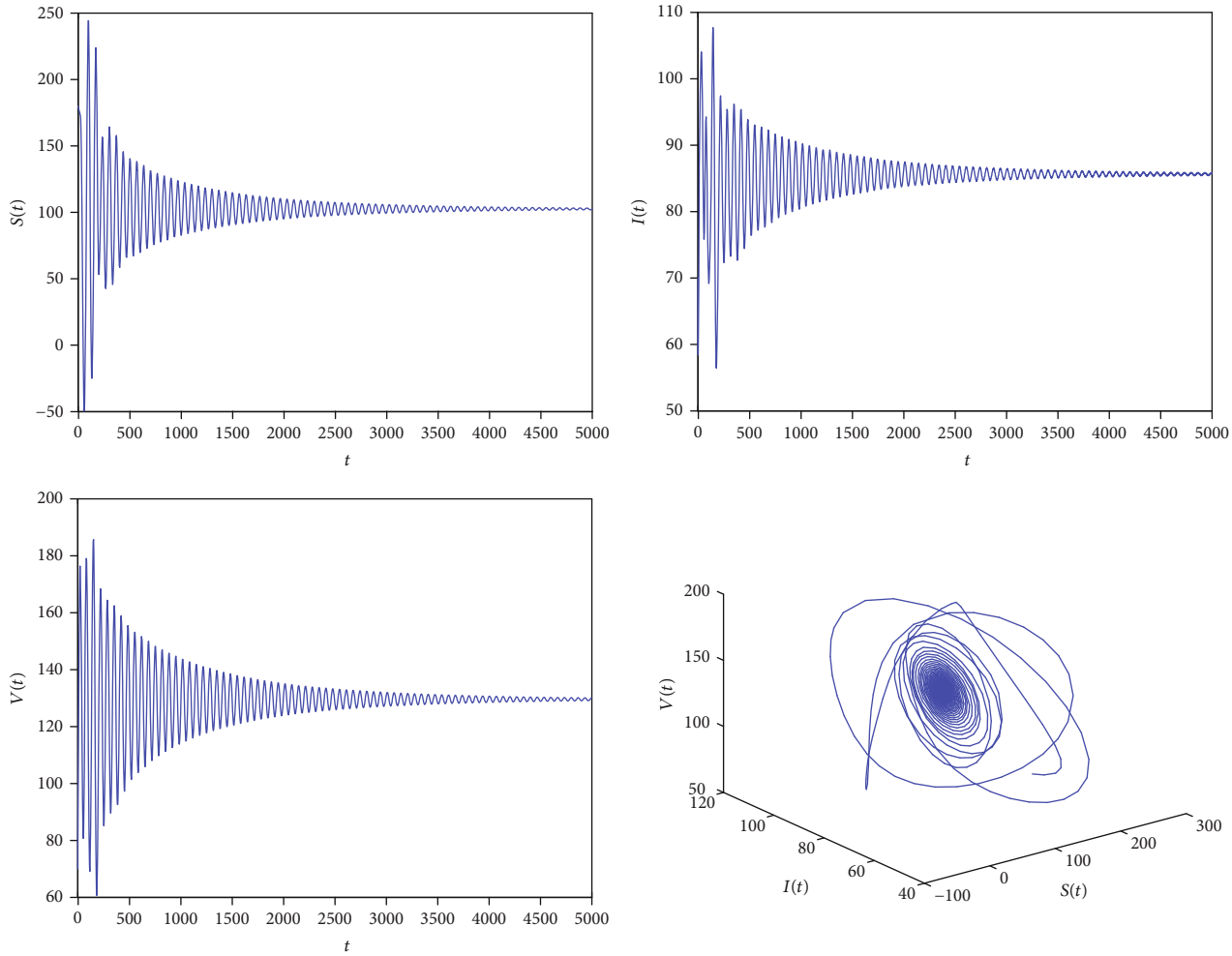


FIGURE 5: The endemic equilibrium E^* is asymptotically stable with $\tau_1 = 23$, $\tau_2 = 21 < \tau_2^*$.

$$\begin{cases} A_{31}\cos\omega\tau_2 + A_{32}\sin\omega\tau_2 = A_{34}, \\ A_{31}\sin\omega\tau_2 - A_{32}\cos\omega\tau_2 = A_{35}, \end{cases} \tag{21}$$

$$\omega^6 + k_5\omega^5 + k_4\omega^4 + k_3\omega^3 + k_2\omega^2 + k_1\omega + k_0 = 0, \tag{22}$$

where

$$A_{31} = \left(n_0 - n_2\omega^2\right)\cos\omega\tau_1 + n_1\omega\sin\omega\tau_1 + m_0 - m_2\omega^2,$$

$$A_{32} = \left(n_0 - n_2\omega^2\right)\sin\omega\tau_1 - n_1\omega\cos\omega\tau_1 + \omega^3 - m_1\omega,$$

$$A_{34} = -r_1\omega\sin\omega\tau_1 - r_0\cos\omega\tau_1 + q_2\omega^2 - q_0,$$

$$A_{35} = -r_1\omega\cos\omega\tau_1 + r_0\sin\omega\tau_1 - q_1\omega,$$

$$k_0 = n_0^2 + m_0^2 - r_0^2 - q_0^2 + 2n_0m_0\cos\omega\tau_1 - 2r_0q_0\cos\omega\tau_1,$$

$$k_1 = 2\left(n_1m_0 - n_0m_1 - r_1q_0 + r_0q_1\right)\sin\omega\tau_1,$$

$$k_2 = n_1^2 + m_1^2 - 2n_0n_2 - 2n_0m_2\cos\omega\tau_1 - 2n_2m_0\cos\omega\tau_1 - 2m_0m_2 + 2n_1m_1\cos\omega\tau_1 - r_1^2 + 2r_0q_2\cos\omega\tau_1 + 2q_0q_2 - 2r_1q_1\cos\omega\tau_1 - q_1^2,$$

$$k_3 = 2\left(n_0 + m_1n_2 - n_1m_2 + r_1q_2\right)\sin\omega\tau_1,$$

$$k_4 = n_2^2 + m_2^2 + 2n_2m_2\cos\omega\tau_1 - 2n_1\cos\omega\tau_1 - 2m_1 - q_2^2,$$

$$k_5 = -2n_2\sin\omega\tau_1. \tag{23}$$

Furthermore, we suppose that

(H2): equation (22) has finite positive roots $\omega_2^{(i)}$, $(i = 1, 2, \cdots, k)$.

Then, for every fixed $\omega_2^{(i)}$, there exists a sequence of critical values $\tau_{2i}^{(j)}$ such that equation (22) holds, where

$$\tau_{2i}^{(j)} = \frac{1}{\omega_2^{(i)}}\arccos\left\{\frac{A_{31}A_{34} - A_{32}A_{35}}{A_{31}^2 + A_{32}^2}\right\} + \frac{2j\pi}{\omega_2^{(i)}}, \quad (i = 1, 2, \cdots, k; j \in \mathbb{N}_0). \tag{24}$$

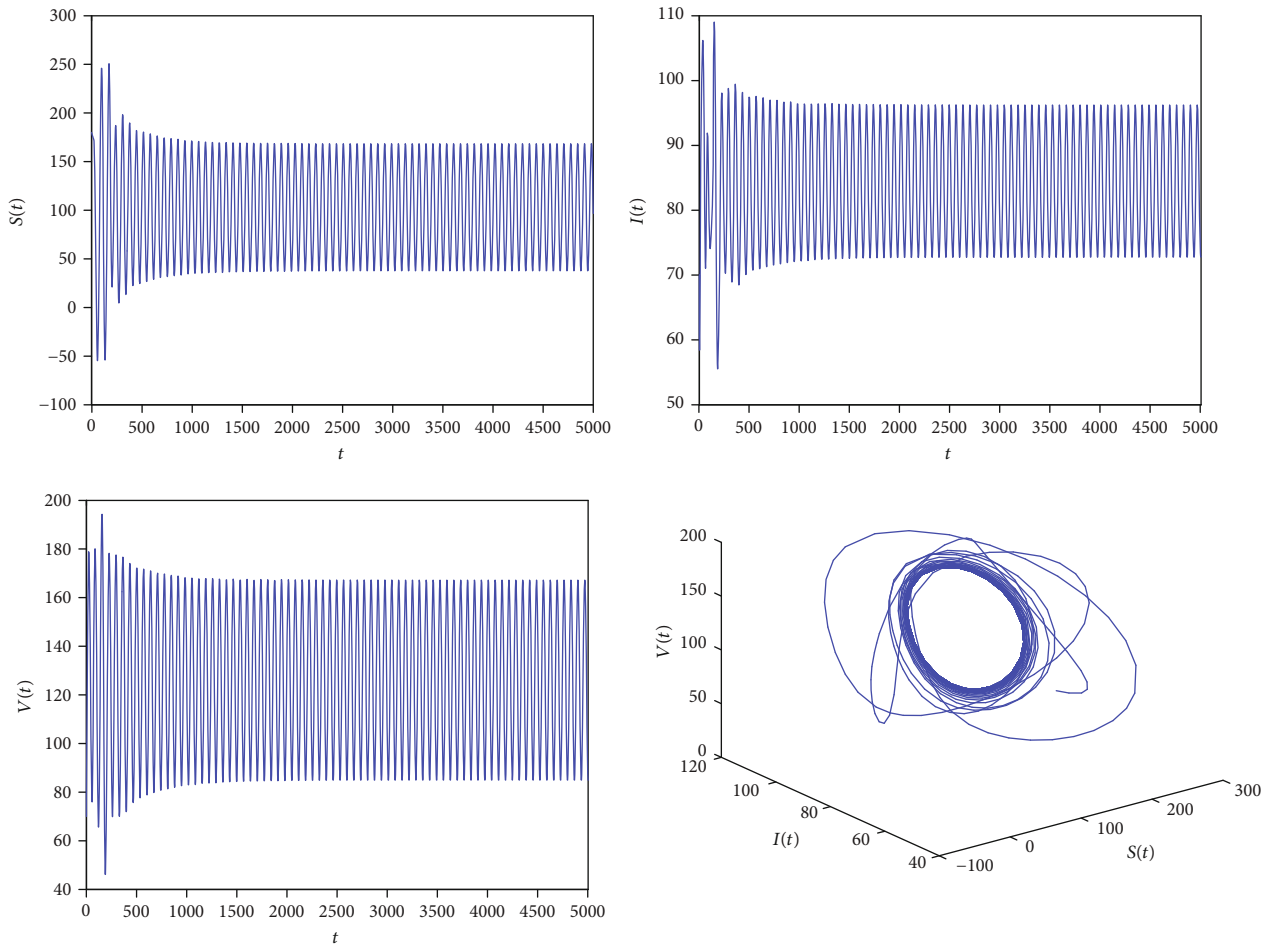


FIGURE 6: The endemic equilibrium E^* is unstable with $\tau_1 = 23$ and $\tau_2 = 22 > \tau_2^*$.

Let $\tau_2^* = \min\{\tau_{2i}^0 | i = 1, 2, \cdots, k\}$, when $\tau_2 = \tau_2^*$, then equation (7) has a pair of purely imaginary roots $\pm i\omega_2^*$.

Next, we also check the transversality condition. By taking the differentiation of (20) with respect to τ_2 and further calculating, we can obtain

$$\left[\frac{d\lambda}{d\tau_2}\right]^{-1} = -\frac{f_{31}(\lambda)}{f_{32}(\lambda)} + \frac{\tau_2}{\lambda}, \tag{25}$$

where

$$\begin{aligned} f_{31}(\lambda) = {} & \left(3\lambda^2 + 2m_2\lambda + m_1\right)e^{\lambda\tau_2} + \left(2n_2\lambda + n_1\right)e^{\lambda(\tau_2-\tau_1)} \\ & - \left(n_2\lambda^2 + n_1\lambda + n_0\right)\tau_1 e^{\lambda(\tau_2-\tau_1)} + \left(2q_2\lambda + q_1\right) \\ & + r_1 e^{-\lambda\tau_1} - \left(r_1\lambda + r_0\right)\tau_1 e^{-\lambda\tau_1}, \end{aligned}$$

$$f_{32}(\lambda) = \left(\lambda^4 + m_2\lambda^3 + m_1\lambda^2 + m_0\lambda\right)e^{\lambda\tau_2} + \left(n_2\lambda^3 + n_1\lambda^2 + n_0\lambda\right)e^{\lambda(\tau_2-\tau_1)}. \tag{26}$$

Thus, we can get

$$\mathrm{Re}\left\{\left[\frac{d\lambda}{d\tau_2}\right]^{-1}_{\tau_2=\tau_2^*}\right\} = -\frac{P_{31}Q_{31} + P_{32}Q_{32}}{Q_{31}^2 + Q_{32}^2}, \tag{27}$$

where

$$\begin{aligned} Q_{31} = {} & \left[\left(n_0\omega_2^* - n_2(\omega_2^*)^3\right)\sin\omega_2^*\tau_1 - n_1(\omega_2^*)^2\cos\omega_2^*\tau_1 + (\omega_2^*)^4\right. \\ & \left. - m_1(\omega_2^*)^2\right]\cos\omega_2^*\tau_2^* - \left[\left(n_0\omega_2^* - n_2(\omega_2^*)^3\right)\cos\omega_2^*\tau_1\right. \\ & \left. + n_1(\omega_2^*)^2\sin\omega_2^*\tau_1 + m_0\omega_2^* - m_2(\omega_2^*)^3\right]\sin\omega_2^*\tau_2^*, \end{aligned}$$

$$\begin{aligned} Q_{32} = {} & \left[\left(n_0\omega_2^* - n_2(\omega_2^*)^3\right)\sin\omega_2^*\tau_1 - n_1(\omega_2^*)^2\cos\omega_2^*\tau_1 + (\omega_2^*)^4\right. \\ & \left. - m_1(\omega_2^*)^2\right]\sin\omega_2^*\tau_2^* + \left[\left(n_0\omega_2^* - n_2(\omega_2^*)^3\right)\cos\omega_2^*\tau_1\right. \\ & \left. + n_1(\omega_2^*)^2\sin\omega_2^*\tau_1 - m_2(\omega_2^*)^3 + m_0\omega_2^*\right]\cos\omega_2^*\tau_2^*, \end{aligned}$$

$$\begin{aligned} P_{31} = {} & \left[\left(\tau_1 n_2(\omega_2^*)^2 - \tau_1 n_0 + n_1\right)\cos\omega_2^*\tau_1\right. \\ & \left. + (2n_2 - \tau_1 n_1)\omega_2^*\sin\omega_2^*\tau_1 - 3(\omega_2^*)^2 + m_1\right]\cos\omega_2^*\tau_2^* \\ & + \left[\left(\tau_1 n_2(\omega_2^*)^2 - \tau_1 n_0 + n_1\right)\sin\omega_2^*\tau_1\right. \\ & \left. - (2n_2 - \tau_1 n_1)\omega_2^*\cos\omega_2^*\tau_1 - 2m_2\omega_2^*\right]\sin\omega_2^*\tau_2^* \\ & + (r_1 - r_0\tau_1)\cos\omega_2^*\tau_1 - \tau_1 r_1\omega_2^*\sin\omega_2^*\tau_1 + q_1, \end{aligned}$$

$$P_{32} = \left[\left(\tau_1 n_2(\omega_2^*)^2 - \tau_1 n_0 + n_1\right)\cos\omega_2^*\tau_1 + (2n_2 - \tau_1 n_1)\omega_2^* \sin\omega_2^*\tau_1 - 3(\omega_2^*)^2 + m_1\right]\sin\omega_2^*\tau_2^* - \left[\left(\tau_1 n_2(\omega_2^*)^2 - \tau_1 n_0 + n_1\right)\sin\omega_2^*\tau_1 - (2n_2 - \tau_1 n_1)\omega_2^*\cos\omega_2^*\tau_1 - 2m_2\omega_2^*\right]\cos\omega_2^*\tau_2^* - (r_1 - r_0\tau_1)\sin\omega_2^*\tau_1 - \tau_1 r_1\omega_2^*\cos\omega_2^*\tau_1 + 2q_2\omega_2^*. \tag{28}$$

Therefore, when

(H3): $P_{31}Q_{31} + P_{32}Q_{32} \neq 0$ holds, then $\mathrm{Re}\{[d\lambda/d\tau_2]^{-1}_{\tau_2=\tau_2^*}\} \neq 0$, i.e., the transversality condition is satisfied.

According to the above discussions and based on [27], we can get the following result.

Theorem 5. *Let $\tau_1 \in (0, \tau_{10})$, if (H2) and (H3) hold, then, the endemic equilibrium E^* is locally asymptotically stable for $\tau_2 \in [0, \tau_2^*)$ and is unstable for $\tau_2 \in (\tau_2^*, +\infty)$. System (2) undergoes Hopf bifurcation at the endemic equilibrium E^* for $\tau_2 = \tau_2^*$.*

4. Properties of Hopf Bifurcation

From the analysis in the last section, we can see that for some critical values of delays, system (2) can occur in a series of periodic solutions at the equilibrium. On the basis of Theorem 5, we will employ the normal form method and the center manifold theorem introduced by Hassard et al. [20] to provide the properties of bifurcating periodic solutions in this section. Without loss of generality, we assume that $\tau_1^* < \tau_2^*$, $\tau_1^* \in (0, \tau_{10})$, $\tau_2 = \tau_2^* + \mu$, $u_1(t) = S(\tau t) - S^*$, $u_2(t) = I(\tau t) - I^*$, and $u_3(t) = V(\tau t) - V^*$. Then, equation (2) can be rewritten into the following functional differential equation in the Banach Space $C = C([-1, 0], \mathbb{R}^3)$.

$$\dot{u}(t) = L_\mu(u_t) + F(\mu, u_t), \tag{29}$$

where $u(t) = (u_1(t), u_2(t), u_3(t))^T \in \mathbb{R}^3$, $u_t(\theta) = u(t + \theta)$, $L_\mu : C \to \mathbb{R}^3$, and $F(\mu, \cdot): \mathbb{R} \times C \to \mathbb{R}^3$ are defined as follows:

$$L_\mu(\phi) = (\tau_2^* + \mu)\left[A'\phi(0) + B'\phi\left(-\frac{\tau_1^*}{\tau_2^*}\right) + C'\phi(-1)\right],$$
$$F(\mu, \phi) = (\tau_2^* + \mu)(F_1, F_2, F_3)^T, \tag{30}$$

where

$$\phi(\theta) = (\phi_1(\theta), \phi_2(\theta), \phi_3(\theta))^T \in C,$$
$$A' = \begin{pmatrix} -\mu_1 & 0 & 0 \\ -\sigma b\beta_1 V^* & -(\mu_1 + \gamma + b\sigma\beta_1 V^*) & b\sigma\beta_1\left(\frac{\Lambda_1}{\mu_1} - I^* - S^*\right) \\ 0 & 0 & -\mu_2 \end{pmatrix},$$
$$B' = \begin{pmatrix} -b\beta_1 V^* & 0 & -b\beta_1 S^* \\ b\beta_1 V^* & 0 & b\beta_1 S^* \\ 0 & 0 & 0 \end{pmatrix},$$
$$C' = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & b\beta_2\left(\frac{\Lambda_2}{\mu_2} - V^*\right) & -b\beta_2 I^* \end{pmatrix},$$
$$F_1 = -b\beta_1\phi_1\left(-\frac{\tau_1^*}{\tau_2^*}\right)\phi_3\left(-\frac{\tau_1^*}{\tau_2^*}\right),$$
$$F_2 = b\beta_1\phi_1\left(-\frac{\tau_1^*}{\tau_2^*}\right)\phi_3\left(-\frac{\tau_1^*}{\tau_2^*}\right) - b\sigma\beta_1\phi_1(0)\phi_3(0) - b\sigma\beta_1\phi_2(0)\phi_3(0),$$
$$F_3 = -b\beta_2\phi_2(-1)\phi_3(-1). \tag{31}$$

Based on the Riesz representation theorem, we know that there exists a function of bounded variation components $\eta(\theta, \mu)$, $\theta \in [-1, 0]$, such that

$$L_\mu(\phi) = \int_{-1}^{0} \mathrm{d}\eta(\theta, \mu)\phi(\theta), \phi(\theta) \in C\left([-1, 0], \mathbb{R}^3\right). \tag{32}$$

In fact, we can take

$$\eta(\theta, \mu) = \begin{cases} (\tau_2^* + \mu)\left(A' + B' + C'\right), & \theta = 0, \\ (\tau_2^* + \mu)\left(B' + C'\right), & \theta \in \left[-\frac{\tau_1^*}{\tau_2^*}, 0\right], \\ (\tau_2^* + \mu)C', & \theta \in \left(-1, -\frac{\tau_1^*}{\tau_2^*}\right), \\ 0, & \theta = -1. \end{cases} \tag{33}$$

Define $A(\mu)$ and $R(\mu)$ by

$$A(\mu)\phi = \begin{cases} \frac{\mathrm{d}\phi(\theta)}{d\theta}, & \theta \in [-1, 0), \\ \int_{-1}^{0} \mathrm{d}\eta(s, \mu)\phi(s), & \theta = 0, \end{cases} \tag{34}$$

$$R(\mu)\phi = \begin{cases} 0, & \theta \in [-1, 0), \\ F(\mu, \phi), & \theta = 0. \end{cases} \tag{35}$$

Then, system (29) can be further represented as

$$\dot{u}(t) = A(\mu)u_t + R(\mu)u_t, \tag{36}$$

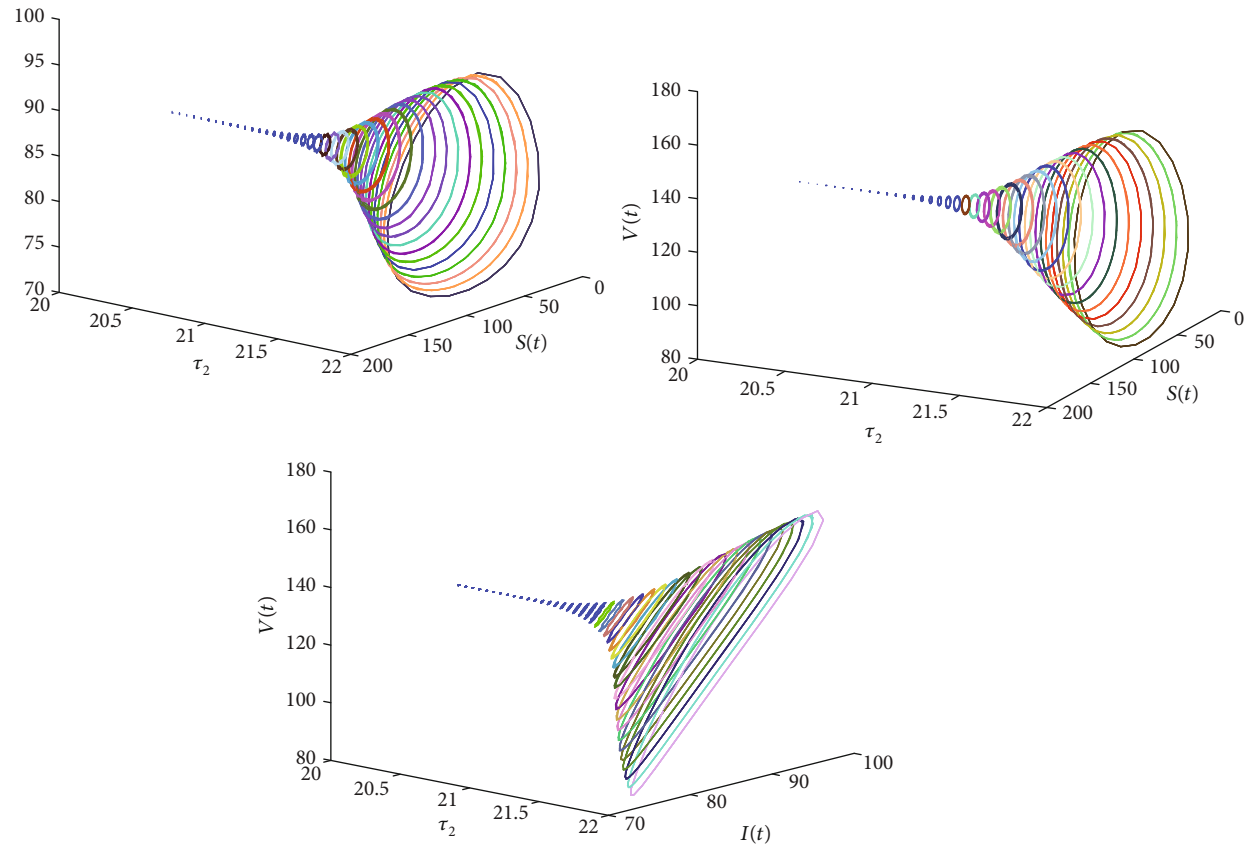


FIGURE 7: The bifurcation diagrams with τ_2 as parameter and $\tau_1 = 23$. A Hopf bifurcation exists at $\tau_2 = \tau_2^*$.

where $u_t(\theta) = u(t+\theta) = (u_1(t+\theta),\ u_2(t+\theta),\ u_3(t+\theta)), \theta \in [-1, 0]$.

For $\varphi \in C([0,1], (\mathbb{R}^3)^*)$, we define

$$A^*\varphi(s) = \begin{cases} -\dfrac{d\psi(s)}{ds}, & s \in (0,1], \\ \displaystyle\int_{-1}^{0} d\eta^T(t,0)\varphi(-t), & s = 0, \end{cases} \tag{37}$$

and the bilinear inner product

$$\langle \varphi(s), \phi(\theta) \rangle = \bar{\varphi}(0)\phi(0) - \int_{-1}^{0}\int_{\xi=0}^{\theta} \bar{\varphi}(\xi-\theta)d\eta(\theta)\phi(\xi)d\xi, \tag{38}$$

where $\eta(\theta) = \eta(\theta, 0)$. Let $A = A(0)$, then A and A^* are a pair of adjoint operators. From the discussions in Section 3, we know that $\pm i\omega_2^*\tau_2^*$ are a pair of eigenvalues of $A(0)$, it follows that they are also a pair of eigenvalues of A^*. Next, we calculate the eigenvectors of $A(0)$ and A^* with respect to $i\omega_2^*\tau_2^*$ and $-i\omega_2^*\tau_2^*$, respectively.

Suppose that $q(\theta)$ is an eigenfunction of $A(0)$ corresponding to $i\omega_2^*\tau_2^*$. Then, by the definition of A, we can take

$$q(\theta) = (1, q_2, q_3)^T e^{i\omega_2^*\tau_2^*\theta}, \tag{39}$$

where

$$q_2 = \frac{\left(i\omega_2^* + \mu_2 + b\beta_2 I^* e^{-i\omega_2^*\tau_2^*}\right)\left(i\omega_2^* + \mu_1 + b\beta_1 V^* e^{-i\omega_2^*\tau_1^*}\right)}{-b\beta_2((\Lambda_2/\mu_2) - V^*)e^{-i\omega_2^*\tau_2^*}},$$
$$q_3 = \frac{i\omega_2^* + \mu_1 + b\beta_1 V^* e^{-i\omega_2^*\tau_1^*}}{-b\beta_1 S^* e^{-i\omega_2^*\tau_1^*}}. \tag{40}$$

Similarly, $q^*(s) = D(1, q_2^*, q_3^*)^T e^{i\omega_2^*\tau_2^* s}, s \in [0,1]$ is the eigenfunction of $A^*(0)$ corresponding to $-i\omega_2^*\tau_2^*$, where

$$q_2^* = -\frac{-i\omega_2^* + \mu_1 + b\beta_1 V^* e^{i\omega_2^*\tau_1^*}}{\left(\sigma - e^{i\omega_2^*\tau_1^*}\right)b\beta_1 V^*},$$
$$q_3^* = \frac{\left(-i\omega_2^* + \mu_1 + \gamma + b\sigma\beta_1 V^*\right)\left(-i\omega_2^* + \mu_1 + b\beta_1 V^* e^{i\omega_2^*\tau_1^*}\right)}{b\beta_2((\Lambda_2/\mu_2) - V^*)e^{i\omega_2^*\tau_2^*}\left(\sigma - e^{i\omega_2^*\tau_1^*}\right)b\beta_1 V^*}. \tag{41}$$

From equation (38), we can get

$$\begin{aligned}\langle q^*(s), q(\theta)\rangle &= \bar{q}^*(0)q(0) - \int_{-1}^{0}\int_{\xi=0}^{\theta}\bar{q}^*(\xi-\theta)d\eta(\theta)q(\xi)d\xi \\ &= \bar{D}(1,\bar{q}_2^*,\bar{q}_3^*)(1,q_2,q_3)^T - \bar{D}\int_{-1}^{0}\int_{\xi=0}^{\theta} \\ &\quad\cdot(1,\bar{q}_2^*,\bar{q}_3^*)e^{i\omega_2^*\tau_2^*(\theta-\xi)}d\eta(\theta)(1,q_2,q_3)^T e^{i\omega_2^*\tau_2^*\xi}d\xi \\ &= \bar{D}\left[1+q_2\bar{q}_2^*+q_3\bar{q}_3^* - (1,\bar{q}_2^*,\bar{q}_3^*)\int_{-1}^{0}\theta e^{i\omega_2^*\tau_2^*\theta}d\eta(\theta)(1,q_2,q_3)^T\right] \\ &= \bar{D}\Big[1+q_2\bar{q}_2^*+q_3\bar{q}_3^*+\tau_1^*b\beta_1e^{-i\omega_2^*\tau_1^*}(-V^*+\bar{q}_2^*-S^*q_3+\bar{q}_2^*q_3) \\ &\quad+\tau_2^*b\beta_2e^{-i\omega_2^*\tau_2^*}\left(\left(\frac{\Lambda_2}{\mu_2}-V^*\right)q_2-I^*q_3\right)\bar{q}_3^*\Big].\end{aligned} \tag{42}$$

Thus, one can choose D as

$$\begin{aligned}D = \Big[&1+\bar{q}_2q_2^*+\bar{q}_3q_3^*+\tau_1^*b\beta_1e^{i\omega_2^*\tau_1^*}(-V^*+\bar{q}_2^*-S^*q_3+\bar{q}_2^*q_3) \\ &+\tau_2^*b\beta_2e^{i\omega_2^*\tau_2^*}\left(\left(\frac{\Lambda_2}{\mu_2}-V^*\right)q_2-I^*q_3\right)\bar{q}_3^*\Big]^{-1},\end{aligned} \tag{43}$$

which satisfies $\langle q^*(s), q(\theta)\rangle = 1$.

Next, using the methods given in [20], we can calculate some explicit expressions as follows that are used to determine the qualities of bifurcating periodic solution.

$$\begin{aligned}g_{20} = 2\bar{D}\tau_2^*\Big[&-b\beta_1q_3e^{-2i\omega_2^*\tau_1^*}+\bar{q}_2^*\left(b\beta_1q_3e^{-2i\omega_2^*\tau_1^*}-\sigma b\beta_1q_3-\sigma b\beta_1q_2q_3\right) \\ &-\bar{q}_3^*b\beta_2q_2q_3e^{-2i\omega_2^*\tau_2^*}\Big],\end{aligned} \tag{44}$$

$$\begin{aligned}g_{11} = \bar{D}\tau_2^*[&-b\beta_1(\bar{q}_3+q_3)+\bar{q}_2^*((1-\sigma)b\beta_1(\bar{q}_3+q_3) \\ &-b\sigma\beta_1(q_2\bar{q}_3+\bar{q}_2q_3))-\bar{q}_3^*b\beta_2(q_2\bar{q}_3+\bar{q}_2q_3)],\end{aligned} \tag{45}$$

$$\begin{aligned}g_{02} = 2\bar{D}\tau_2^*\Big[&-b\beta_1\bar{q}_3e^{2i\omega_2^*\tau_1^*}+\bar{q}_2^*\left(b\beta_1\bar{q}_3e^{2i\omega_2^*\tau_1^*}-\sigma b\beta_1\bar{q}_3-\sigma b\beta_1\bar{q}_2\bar{q}_3\right) \\ &-\bar{q}_3^*b\beta_2\bar{q}_2\bar{q}_3e^{2i\omega_2^*\tau_2^*}\Big],\end{aligned} \tag{46}$$

$$\begin{aligned}g_{21} = \bar{D}\tau_2^*\Big\{&-b\beta_1\left[2W_{11}^{(3)}\left(-\frac{\tau_1^*}{\tau_2^*}\right)e^{-i\omega_2^*\tau_1^*}+W_{20}^{(3)}\left(-\frac{\tau_1^*}{\tau_2^*}\right)e^{i\omega_2^*\tau_1^*}\right. \\ &\left.+W_{20}^{(1)}\left(-\frac{\tau_1^*}{\tau_2^*}\right)\bar{q}_3e^{i\omega_2^*\tau_1^*}+2W_{11}^{(1)}\left(-\frac{\tau_1^*}{\tau_2^*}\right)q_3e^{-i\omega_2^*\tau_1^*}\right] \\ &+\bar{q}_2^*\left[b\beta_1\left(2W_{11}^{(3)}\left(-\frac{\tau_1^*}{\tau_2^*}\right)e^{-i\omega_2^*\tau_1^*}+W_{20}^{(3)}\left(-\frac{\tau_1^*}{\tau_2^*}\right)e^{i\omega_2^*\tau_1^*}\right.\right. \\ &\left.+W_{20}^{(1)}\left(-\frac{\tau_1^*}{\tau_2^*}\right)\bar{q}_3e^{i\omega_2^*\tau_1^*}+2W_{11}^{(1)}\left(-\frac{\tau_1^*}{\tau_2^*}\right)q_3e^{-i\omega_2^*\tau_1^*}\right) \\ &-\sigma b\beta_1\left(2W_{11}^{(3)}(0)+W_{20}^{(3)}(0)+W_{20}^{(1)}(0)\bar{q}_3+2W_{11}^{(1)}(0)q_3\right) \\ &\left.-b\sigma\beta_1\left(2W_{11}^{(3)}(0)q_2+W_{20}^{(3)}(0)\bar{q}_2+W_{20}^{(2)}(0)\bar{q}_32W_{11}^{(2)}(0)q_3\right)\right] \\ &-\bar{q}_3^*b\beta_2\left[2W_{11}^{(3)}(-1)q_2e^{-i\omega_2^*\tau_2^*}+W_{20}^{(3)}(-1)\bar{q}_2e^{i\omega_2^*\tau_2^*}\right. \\ &\left.+W_{20}^{(2)}(-1)\bar{q}_3e^{i\omega_2^*\tau_2^*}+2W_{11}^{(2)}(-1)q_3e^{-i\omega_2^*\tau_2^*}\right]\Big\},\end{aligned} \tag{47}$$

where

$$\begin{aligned}W_{20}(\theta) &= \frac{ig_{20}}{\omega_2^*\tau_2^*}q(0)e^{i\omega_2^*\tau_2^*\theta}+\frac{i\bar{g}_{02}}{3\omega_2^*\tau_2^*}\bar{q}(0)e^{-i\omega_2^*\tau_2^*\theta}+E_1e^{2i\omega_2^*\tau_2^*\theta}, \\ W_{11}(\theta) &= -\frac{ig_{11}}{\omega_2^*\tau_2^*}q(0)e^{i\omega_2^*\tau_2^*\theta}+\frac{i\bar{g}_{11}}{\omega_2^*\tau_2^*}\bar{q}(0)e^{-i\omega_2^*\tau_2^*\theta}+E_2, \\ E_1 &= 2\left(\frac{\Delta_{11}}{\Delta_1},\frac{\Delta_{12}}{\Delta_1},\frac{\Delta_{13}}{\Delta_1}\right), \\ E_2 &= \left(\frac{\Delta_{21}}{\Delta_2},\frac{\Delta_{22}}{\Delta_2},\frac{\Delta_{23}}{\Delta_2}\right),\end{aligned}$$

$$\Delta_1 = \begin{vmatrix} 2i\omega_2^*+\mu_1+b\beta_1V^*e^{-2i\omega_2^*\tau_1^*} & 0 & b\beta_1S^*e^{-2i\omega_2^*\tau_1^*} \\ \left(\sigma-e^{-2i\omega_2^*\tau_1^*}\right)b\beta_1V^* & 2i\omega_2^*+\mu_1+\gamma+b\sigma\beta_1V^* & -b\beta_1S^*e^{-2i\omega_2^*\tau_1^*}-b\sigma\beta_1\left(\frac{\Lambda_1}{\mu_1}-I^*-S^*\right) \\ 0 & -b\beta_2\left(\frac{\Lambda_2}{\mu_2}-V^*\right)e^{-2i\omega_2^*\tau_2^*} & 2i\omega_2^*+\mu_2+b\beta_2I^*e^{-2i\omega_2^*\tau_2^*} \end{vmatrix}, \tag{48}$$

$$\Delta_2 = \begin{vmatrix} \mu_1+b\beta_1V^* & 0 & b\beta_1S^* \\ -(1-\sigma)b\beta_1V^* & \mu_1+\lambda+b\sigma\beta_1V^* & -\frac{(\mu_1+\gamma)I^*}{V^*} \\ 0 & -b\beta_2\left(\frac{\Lambda_2}{\mu_2}-V^*\right) & \mu_2+b\beta_2I^* \end{vmatrix}.$$

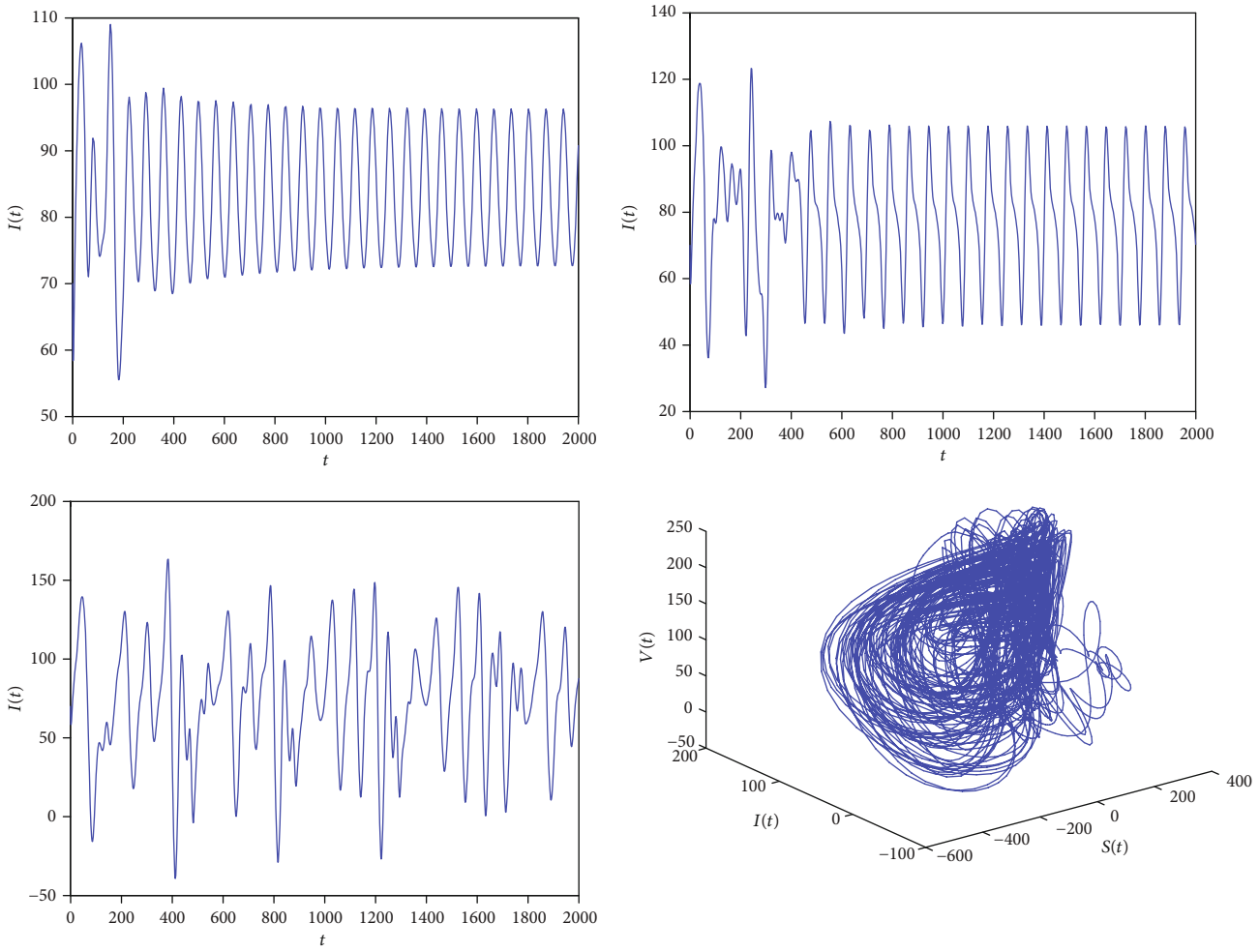


Figure 8: The bifurcation periodic solutions of $I(t)$ with $\tau_1 = 23$ and $\tau_2 = 22, 28, 38$ (from the top to the first of the bottom). The bifurcation periodic solution is unstable with $\tau_1 = 23$ and $\tau_2 = 38$.

Δ_{11}, Δ_{12}, and Δ_{13} are the third-order determinants obtained by substituting the first, second, and third columns of Δ_1 by vector $(P_1, P_2, P_3)^T$, respectively. Δ_{21}, Δ_{22}, and Δ_{23} are also third-order determinants obtained by replacing the first, second, and third columns of Δ_2 by vector $(Q_1, Q_2, Q_3)^T$, respectively, where

$$\begin{aligned}
P_1 &= -b\beta_1 q_3 e^{2i\omega_2^*\tau_1^*},\\
P_2 &= b\beta_1 q_3 e^{-2i\omega_2^*\tau_1^*} - \sigma b\beta_1 q_3 - \sigma b\beta_1 q_2 q_3,\\
P_3 &= -b\beta_2 q_2 q_3 e^{-2i\omega_2^*\tau_2^*},\\
Q_1 &= -b\beta_1 \operatorname{Re}\{q_3\},\\
Q_2 &= (1-\sigma)b\beta_1 \operatorname{Re}\{q_3\} - b\sigma\beta_1 \operatorname{Re}\{q_2\bar{q}_3\},\\
Q_3 &= -b\beta_2 \operatorname{Re}\{q_2\bar{q}_3\}.
\end{aligned} \tag{49}$$

Thus, the following quantities can be calculated:

$$\begin{cases}
c_1(0) = \dfrac{i}{2\omega_2^*\tau_2^*}\left(g_{20}g_{11} - 2|g_{11}|^2 - \dfrac{1}{3}|g_{02}|^2\right) + \dfrac{1}{2}g_{21},\\
\mu_2 = -\dfrac{\operatorname{Re}\{c_1(0)\}}{\operatorname{Re}\left\{\lambda'(\tau_2^*)\right\}},\\
\beta_2 = 2\operatorname{Re}\{c_1(0)\},\\
T_2 = -\dfrac{1}{\omega_2^*\tau_2^*}\left(\operatorname{Im}\{c_1(0)\} + \mu_2 \operatorname{Im}\left\{\lambda'(\tau_2^*)\right\}\right).
\end{cases} \tag{50}$$

Therefore, based on [20], we can obtain the main results in the present section as follows.

Theorem 6. *For system (2), the periodic solution is supercritical (resp., subcritical) if $\mu_2 > 0$ (resp., $\mu_2 < 0$). The bifurcating periodic solutions are orbitally asymptotically stable with an*

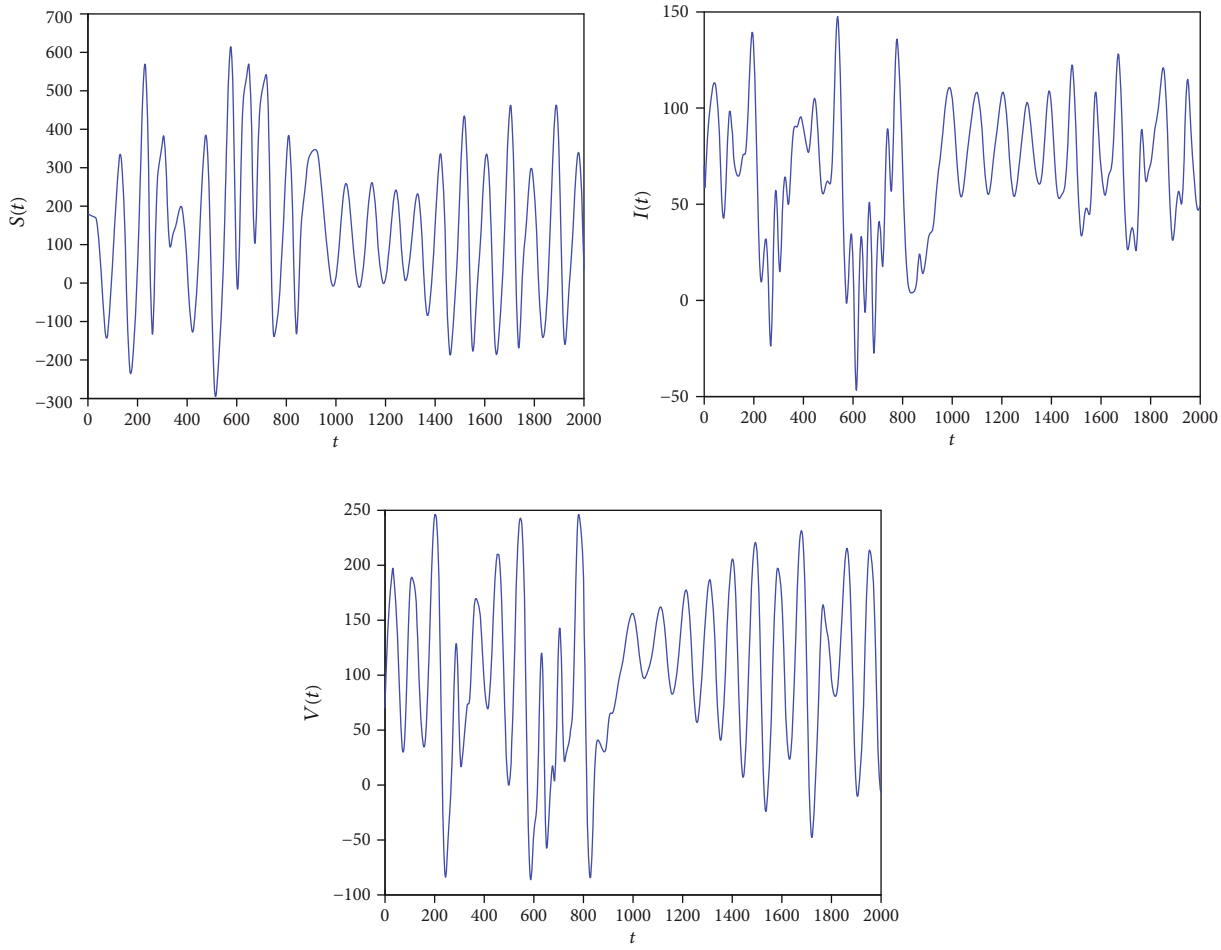


Figure 9: The system appears irregular periodic behaviors when $\tau_1 = 31$ and $\tau_2 = 31$.

asymptotical phase (resp., unstable) if $\beta_2 < 0$ (resp., $\beta_2 > 0$). The period of the bifurcating periodic solutions increases (resp., decreases) if $T_2 > 0$ (resp. $T_2 < 0$).

5. Numerical Examples and Simulations

We select the parameters to simulate the above theoretical results as follows:

$$\begin{aligned} \Lambda_1 &= 9, b = 0.3, \\ \beta_1 &= 0.002, \mu_1 = 0.01, \\ \sigma &= 0.8, \gamma = 0.6, \\ \Lambda_2 &= 9, \\ \beta_2 &= 0.005, \\ \mu_2 &= 0.05. \end{aligned} \tag{51}$$

By calculations, we can obtain that the endemic equilibrium of system (2) is $E^* = (S^*, I^*, V^*) = (102.5734, 85.6437, 129.5701)$. Under the above parameters, the condition (H1) is satisfied. According to Theorem 2, we know that

Table 1: The relationships among the partial immunity σ, the endemic equilibrium E^*, and the critical values τ_{10} and τ_2^*.

σ	1	0.8	0.6	0.4	0.2
S^*	97.4478	102.5734	111.3948	129.1222	170.4044
I^*	107.0566	85.6437	63.4246	41.2030	21.8956
V^*	137.2619	129.5701	117.9896	99.5023	71.3613
τ_{10}	21.5895	23.3291	26.6112	34.3647	56.7849
τ_2^*	17.8120	21.4739	25.5925	30.6267	44.0800

the endemic equilibrium E^* of system (2) is asymptotically stable when $\tau_1 = \tau_2 = 0$ (see Figure 1). When $\tau_1 > 0$, $\tau_2 = 0$, we get $p_0 = -4.7847 \times 10^{-5} < 0$, $\tau_{10} = 23.3291$, $z_1 = 0.0053$, and $f'(z_1) = 0.017 \neq 0$; from Theorem 4, the endemic equilibrium E^* is locally asymptotically stable for $\tau_1 \in (0, \tau_{10})$ and is unstable for $\tau_1 > \tau_{10}$, and a Hopf bifurcation occurs at $\tau_1 = \tau_{10}$ (see Figures 2–4). When $\tau_2 > 0$, fixing $\tau_1 = 23$ that belongs to its stable interval $(0, \tau_{10})$, then, by calculation, we obtain that a positive root of equation (22) is $\omega = 0.0933$ and $\tau_2^* = 21.4739$, which follow from Theorem 5 that the endemic equilibrium E^* is locally asymptotically stable for $\tau_2 \in (0, \tau_2^*)$

and is unstable for $\tau_2 \in (\tau_2^*, +\infty)$, a Hopf bifurcation exits at $\tau_2 = \tau_2^*$ (see Figures 5–7).

The numerical results show that $c_1(0) = 1.0266 \times 10^{-5} - 4.7626 \times 10^{-6} i$, $\mu_2 = -0.0024 < 0$, $\beta_2 = 2.0531 \times 10^{-5} > 0$, and $T_2 = 2.3115 \times 10^{-6} > 0$, which imply that, for system (2), its periodic solution is subcritical at τ_2^* and unstable, and the period of the bifurcating periodic solutions increases (see Figure 8). Also, when the two delays are greater than their critical values, the system will produce irregular periodic behaviors, which is depicted in Figure 9.

Besides, the influences of the parameter σ on the endemic equilibrium E^* and the critical values τ_{10} and τ_2^* of the two delays τ_1 and τ_2 are shown in Table 1. When σ decreases, I^* and V^* also decrease, but S^*, τ_{10}, and τ_2^* increase, which suggests that the degree of partial protection for recovered people acquired by a primary infection can change the densities of distributions of susceptible and infective people and can also affect on the stability of positive equilibrium. It is noteworthy that it can not change the direction of Hopf bifurcation and the stability of bifurcating periodic solutions.

6. Conclusions

In this paper, we considered a vector-borne disease model with two delays and reinfection based on the models proposed by [4–6]. Some theoretical analyses of this model were implemented. From the characteristic equations, firstly, we discussed the conditions of the endemic equilibrium E^*, which is locally asymptotically stable in the model without delays by the Routh-Hurwitz criterion. Then, by setting $\tau_2 = 0$, regarding τ_1 as a bifurcating parameter, and setting τ_1 in its stable interval, regarding τ_2 as a parameter, respectively, we employed the Hopf bifurcation theorem to analyze the conditions for the stability of equilibrium and existence of Hopf bifurcation. Furthermore, when τ_1 is in its stable interval, for parameter τ_2, the directions and stability of Hopf bifurcation were investigated, and some explicit formulas were derived by using the normal form theory and center manifold theorem. In the end, our numerical examples and simulations verified the correctness of theoretical conclusions. As the facts show, the disease will be persistent under the set of parameters, and the number of the host population and vector population will fluctuate. Also, when both delays exceed their critical values, the stability of the system and the properties and behaviors of the bifurcating periodic solutions may lead to change. The oscillatory behavior and phenomenon that an increased partial immunity may give rise to instability will prompt medical workers engaged in infectious diseases to remain vigilant against the virus even if they have noticed that the number of infected populations decreases. The qualitative analysis and numerical simulation used in the research of dynamic behavior of system can provide theoretical basis for revealing the development process, transmission pattern, and change trend of diseases and give important reference to the analysis, prediction, and control of human infectious diseases.

Appendix

A. The Derivation Process of R_0 Based on the Next-Generation Method

R_0 is the dominant eigenvalue of the next generation matrix $G = FV^{-1}$ whose entries are interpreted as the number of secondary infections produced by infected vectors and hosts in the process of their infection. Here, the notation we use follows [22–24]. For system (2), we calculate

$$F = \begin{pmatrix} 0 & \dfrac{b\beta_1\Lambda_1}{\mu_1} \\ \dfrac{b\beta_2\Lambda_2}{\mu_2} & 0 \end{pmatrix}, \quad V = \begin{pmatrix} \mu_1 + \gamma & 0 \\ 0 & \mu_2 \end{pmatrix}, \quad G = \begin{pmatrix} 0 & \dfrac{b\beta_1\Lambda_1}{\mu_1\mu_2} \\ \dfrac{b\beta_2\Lambda_2}{(\mu_1+\gamma)\mu_2} & 0 \end{pmatrix}. \tag{A.1}$$

Thus, $R_0 = (b^2\Lambda_1\Lambda_2\beta_1\beta_2)/(\mu_1\mu_2^2(\mu_1+\gamma))$. Note that we use R_0^2 as in [5], though it is called R_0. Either way, its threshold behavior $R_0 > 1$ remains unchanged.

B. The Derivation Process of Coefficients g_{20}, g_{11}, g_{02}, and g_{21} of Equation (47)

Using the algorithms and marks presented in [20], when $\mu = 0$, the projection system of system (2) on the center manifold C_0 can be calculated. For the solution u_t of equation (29), we define $z(t) = \langle q^*, u_t \rangle$, which follows from equations (35)–(39) that

$$\dot{z}(t) = \langle q^*, \dot{u}_t \rangle = \langle A^*(0)q^*, u_t \rangle + \bar{q}^*(0)F(0, u_t) \triangleq i\omega_2^*\tau_2^* z + g(z, \bar{z}), \tag{B.1}$$

where $g(z, \bar{z}) = \bar{q}^*(0)F(0, u_t) = g_{20}(\theta)(z^2/2) + g_{11}(\theta)z\bar{z} + g_{02}(\theta)(\bar{z}^2/2) + g_{21}(\theta)(z^2\bar{z}/2) \cdots$.

Let

$$W(t, \theta) = u_t(\theta) - 2\,\mathrm{Re}\,\{z(t)q(\theta)\}, \tag{B.2}$$

then, on the center manifold C_0, we can have

$$W(t, \theta) = W(z(t), \bar{z}(t), \theta), \tag{B.3}$$

where

$$W(z, \bar{z}, \theta) = W_{20}(\theta)\frac{z^2}{2} + W_{11}(\theta)z\bar{z} + W_{02}(\theta)\frac{\bar{z}^2}{2} + W_{30}(\theta)\frac{z^3}{6} + \cdots. \tag{B.4}$$

z and $\bar{z}$ are local coordinates on C_0 in the directions of q^*

and $\bar{q}^*$. From equations (39) and (B.2), we can see

$$\begin{aligned} u_t(\theta) &= W(t,\theta) + z(t)q(\theta) + \bar{z}(t)\bar{q}(\theta) = (1, q_2, q_3)^T e^{i\omega_2^*\tau_2^*\theta} z \\ &\quad + (1, \bar{q}_2, \bar{q}_3)^T e^{-i\omega_2^*\tau_2^*\theta}\bar{z} + W_{20}(\theta)\frac{z^2}{2} + W_{11}(\theta)z\bar{z} \\ &\quad + W_{02}(\theta)\frac{\bar{z}^2}{2} + W_{30}(\theta)\frac{z^3}{6} + \cdots. \end{aligned} \tag{B.5}$$

Therefore, we can obtain

$$u_{1t}(0) = z + \bar{z} + W_{20}^{(1)}(0)\frac{z^2}{2} + W_{11}^{(1)}(0)z\bar{z} + W_{02}^{(1)}(0)\frac{\bar{z}^2}{2} + \cdots,$$

$$u_{2t}(0) = zq_2 + \bar{z}\bar{q}_2 + W_{20}^{(2)}(0)\frac{z^2}{2} + W_{11}^{(2)}(0)z\bar{z} + W_{02}^{(2)}(0)\frac{\bar{z}^2}{2} + \cdots,$$

$$u_{3t}(0) = zq_3 + \bar{z}\bar{q}_3 + W_{20}^{(3)}(0)\frac{z^2}{2} + W_{11}^{(3)}(0)z\bar{z} + W_{02}^{(3)}(0)\frac{\bar{z}^2}{2} + \cdots,$$

$$\begin{aligned} u_{1t}(-1) &= ze^{-i\omega_2^*\tau_2^*} + \bar{z}e^{i\omega_2^*\tau_2^*} + W_{20}^{(1)}(-1)\frac{z^2}{2} + W_{11}^{(1)}(-1)z\bar{z} \\ &\quad + W_{02}^{(1)}(-1)\frac{\bar{z}^2}{2} + \cdots, \end{aligned}$$

$$\begin{aligned} u_{2t}(-1) &= zq_2e^{-i\omega_2^*\tau_2^*} + \bar{z}\bar{q}_2e^{i\omega_2^*\tau_2^*} + W_{20}^{(2)}(-1)\frac{z^2}{2} + W_{11}^{(2)}(-1)z\bar{z} \\ &\quad + W_{02}^{(2)}(-1)\frac{\bar{z}^2}{2} + \cdots, \end{aligned}$$

$$\begin{aligned} u_{3t}(-1) &= zq_3e^{-i\omega_2^*\tau_2^*} + \bar{z}\bar{q}_3e^{i\omega_2^*\tau_2^*} + W_{20}^{(3)}(-1)\frac{z^2}{2} + W_{11}^{(3)}(-1)z\bar{z} \\ &\quad + W_{02}^{(3)}(-1)\frac{\bar{z}^2}{2} + \cdots, \end{aligned}$$

$$\begin{aligned} u_{1t}\left(-\frac{\tau_1^*}{\tau_2^*}\right) &= ze^{-i\omega_2^*\tau_1^*} + \bar{z}e^{i\omega_2^*\tau_1^*} + W_{20}^{(1)}\left(-\frac{\tau_1^*}{\tau_2^*}\right)\frac{z^2}{2} \\ &\quad + W_{11}^{(1)}\left(-\frac{\tau_1^*}{\tau_2^*}\right)z\bar{z} + W_{02}^{(1)}\left(-\frac{\tau_1^*}{\tau_2^*}\right)\frac{\bar{z}^2}{2} + \cdots, \end{aligned}$$

$$\begin{aligned} u_{2t}\left(-\frac{\tau_1^*}{\tau_2^*}\right) &= zq_2e^{-i\omega_2^*\tau_1^*} + \bar{z}\bar{q}_2e^{i\omega_2^*\tau_1^*} + W_{20}^{(2)}\left(-\frac{\tau_1^*}{\tau_2^*}\right)\frac{z^2}{2} \\ &\quad + W_{11}^{(2)}\left(-\frac{\tau_1^*}{\tau_2^*}\right)z\bar{z} + W_{02}^{(2)}\left(-\frac{\tau_1^*}{\tau_2^*}\right)\frac{\bar{z}^2}{2} + \cdots, \end{aligned}$$

$$\begin{aligned} u_{3t}\left(-\frac{\tau_1^*}{\tau_2^*}\right) &= zq_3e^{-i\omega_2^*\tau_1^*} + \bar{z}\bar{q}_3e^{i\omega_2^*\tau_1^*} + W_{20}^{(3)}\left(-\frac{\tau_1^*}{\tau_2^*}\right)\frac{z^2}{2} \\ &\quad + W_{11}^{(3)}\left(-\frac{\tau_1^*}{\tau_2^*}\right)z\bar{z} + W_{02}^{(3)}\left(-\frac{\tau_1^*}{\tau_2^*}\right)\frac{\bar{z}^2}{2} + \cdots. \end{aligned} \tag{B.6}$$

Furthermore, we can get $g(z,\bar{z})$ as follows:

$$\begin{aligned} g(z,\bar{z}) &= \bar{q}^*(0)F(0,u_t) = \bar{D}(1,\bar{q}_2^*,\bar{q}_3^*)F(0,u_t) = \bar{D}(1,\bar{q}_2^*,\bar{q}_3^*)\tau_2^*\begin{pmatrix} -b\beta_1 u_{1t}\left(-\frac{\tau_1^*}{\tau_2^*}\right)u_{3t}\left(-\frac{\tau_1^*}{\tau_2^*}\right) \\ b\beta_1 u_{1t}\left(-\frac{\tau_1^*}{\tau_2^*}\right)u_{3t}\left(-\frac{\tau_1^*}{\tau_2^*}\right) - \sigma b\beta_1 u_{1t}(0)u_{3t}(0) - b\sigma\beta_1 u_{2t}(0)u_{3t}(0) \\ -b\beta_2 u_{2t}(-1)u_{3t}(-1) \end{pmatrix} \\ &= \bar{D}\tau_2^*\left\{-b\beta_1 u_{1t}\left(-\frac{\tau_1^*}{\tau_2^*}\right)u_{3t}\left(-\frac{\tau_1^*}{\tau_2^*}\right) + \bar{q}_2^*\left[b\beta_1 u_{1t}\left(-\frac{\tau_1^*}{\tau_2^*}\right)u_{3t}\left(-\frac{\tau_1^*}{\tau_2^*}\right) - \sigma b\beta_1 u_{1t}(0)u_{3t}(0) - b\sigma\beta_1 u_{2t}(0)u_{3t}(0)\right]\right. \\ &\quad \left.-\bar{q}_3^* b\beta_2 u_{2t}(-1)u_{3t}(-1)\right\}. \end{aligned} \tag{B.7}$$

Substituting $u_{1t}(0)$, $u_{2t}(0)$, $u_{3t}(0)$, $u_{2t}(-1)$, $u_{3t}(-1)$, $u_{1t}(-\tau_1^*/\tau_2^*)$, and $u_{3t}(-\tau_1^*/\tau_2^*)$ into the above expression of $g(z,\bar{z})$ and comparing the coefficients yield the coefficients g_{20}, g_{11}, g_{02}, and g_{21} of equation (47).

Obviously, g_{20}, g_{11}, and g_{02} can be determined directly by substituting system parameters, but $W_{20}(\theta)$ and $W_{11}(\theta)$ in g_{21} need to be calculated through some methods [25, 26].

By equations (36) and (B.2), we can get

$$\begin{aligned} \dot{W} = \dot{u}_t - \dot{z}q - \dot{\bar{z}}\bar{q} &= \begin{cases} AW - 2\,\mathrm{Re}\,\{\bar{q}^*(0)F(0,u_t)q(\theta)\}, \theta \in [-1,0), \\ AW - 2\,\mathrm{Re}\,\{\bar{q}^*(0)F(0,u_t)q(0)\} + F(0,u_t), \theta = 0, \end{cases} \\ &\triangleq AW + H(z,\bar{z},\theta), \end{aligned} \tag{B.8}$$

where

$$H(z,\bar{z},\theta)=H_{20}(\theta)\frac{z^2}{2}+H_{11}(\theta)z\bar{z}+H_{02}(\theta)\frac{\bar{z}^2}{2}+\cdots. \quad \text{(B.9)}$$

Thus, we have

$$AW(t,\theta)-\dot{W}=-H(z,\bar{z},\theta)=-H_{20}(\theta)\frac{z^2}{2}-H_{11}(\theta)z\bar{z}-H_{02}(\theta)\frac{\bar{z}^2}{2}-\cdots. \quad \text{(B.10)}$$

From equation (B.4), we can get

$$AW(t,\theta)=AW_{20}(\theta)\frac{z^2}{2}+AW_{11}(\theta)z\bar{z}+AW_{02}(\theta)\frac{\bar{z}^2}{2}+AW_{30}(\theta)\frac{z^3}{6}+\cdots,$$
$$\dot{W}=W_{20}(\theta)z\dot{z}+W_{11}(\theta)\left(\dot{z}\bar{z}+z\dot{\bar{z}}\right)+\cdots=2i\omega^*\tau_2^*W_{20}(\theta)\frac{z^2}{2}+\cdots. \quad \text{(B.11)}$$

Therefore, we can obtain

$$(A-2i\omega^*\tau_2^*)W_{20}(\theta)=-H_{20}(\theta),\ AW_{11}(\theta)=-H_{11}(\theta). \quad \text{(B.12)}$$

For $\theta\in[-1,0)$, we have

$$H(z,\bar{z},\theta)=-g(z,\bar{z})q(\theta)-\bar{g}(z,\bar{z})\bar{q}(\theta). \quad \text{(B.13)}$$

By comparing coefficients of (B.9) with (B.13), one can get

$$H_{20}(\theta)=-(g_{20}q(\theta)+\bar{g}_{02}\bar{q}(\theta)), \quad \text{(B.14)}$$

$$H_{11}(\theta)=-(g_{11}q(\theta)+\bar{g}_{11}\bar{q}(\theta)). \quad \text{(B.15)}$$

From equations (B.12) and (B.14) and the definition of A, we can obtain

$$\dot{W}_{20}(\theta)=2i\omega_2^*\tau_2^*W_{20}(\theta)+g_{20}q(\theta)+\bar{g}_{02}\bar{q}(\theta). \quad \text{(B.16)}$$

Since $q(\theta)=q(0)e^{i\omega_2^*\tau_2^*\theta}=(1,q_2,q_3)^Te^{i\omega_2^*\tau_2^*\theta}$, then according to (B.16), we can get

$$W_{20}(\theta)=\frac{ig_{20}}{\omega_2^*\tau_2^*}q(0)e^{i\omega_2^*\tau_2^*\theta}+\frac{i\bar{g}_{02}}{3\omega_2^*\tau_2^*}\bar{q}(0)e^{-i\omega_2^*\tau_2^*\theta}+E_1e^{2i\omega_2^*\tau_2^*\theta}, \quad \text{(B.17)}$$

where $E_1=(E_1^{(1)},E_1^{(2)},E_1^{(3)})\in\mathbb{R}^3$.

Similarly,

$$\dot{W}_{11}(\theta)=g_{11}q(\theta)+\bar{g}_{11}\bar{q}(\theta), \quad \text{(B.18)}$$

$$W_{11}(\theta)=-\frac{ig_{11}}{\omega_2^*\tau_2^*}q(0)e^{i\omega_2^*\tau_2^*\theta}+\frac{i\bar{g}_{11}}{\omega_2^*\tau_2^*}\bar{q}(0)e^{-i\omega_2^*\tau_2^*\theta}+E_2, \quad \text{(B.19)}$$

where $E_2=(E_2^{(1)},E_2^{(2)},E_2^{(3)})\in\mathbb{R}^3$.

Next, we seek the appropriate constant vectors E_1 and E_2 in expressions (B.17) and (B.19). From equation (B.12) and the definition of A, we can have

$$\int_{-1}^{0}d\eta(\theta)W_{20}(\theta)=2i\omega_2^*\tau_2^*W_{20}(0)-H_{20}(0), \quad \text{(B.20)}$$

$$\int_{-1}^{0}d\eta(\theta)W_{11}(\theta)=-H_{11}(0), \quad \text{(B.21)}$$

where $\eta(\theta)=\eta(0,\theta)$. Thus,

$$H_{20}(0)=-g_{20}q(0)-\bar{g}_{02}\bar{q}(0)+2\tau_2^*(P_1,P_2,P_3)^T, \quad \text{(B.22)}$$

$$H_{11}(0)=-g_{11}q(0)-\bar{g}_{11}\bar{q}(0)+\tau_2^*(Q_1,Q_2,Q_3)^T. \quad \text{(B.23)}$$

where $P_1=-b\beta_1q_3e^{2i\omega_2^*\tau_1^*}$, $P_2=b\beta_1q_3e^{-2i\omega_2^*\tau_1^*}-\sigma b\beta_1q_3-\sigma b\beta_1q_2q_3$, and $P_3=-b\beta_2q_2q_3e^{-2i\omega_2^*\tau_2^*}$ and $Q_1=-b\beta_1\,\mathrm{Re}\{q_3\}$, $Q_2=(1-\sigma)b\beta_1\,\mathrm{Re}\{q_3\}-b\sigma\beta_1\,\mathrm{Re}\{q_2\bar{q}_3\}$, and $Q_3=-b\beta_2\,\mathrm{Re}\{q_2\bar{q}_3\}$.

Note that

$(i\omega_2^*\tau_2^*I-\int_{-1}^{0}e^{i\omega_2^*\tau_2^*\theta}d\eta(\theta))q(0)=0$,

$(-i\omega_2^*\tau_2^*I-\int_{-1}^{0}e^{-i\omega_2^*\tau_2^*\theta}d\eta(\theta))\bar{q}(0)=0$.

Substituting expressions (39) and (B.22) into equation (B.20), we can have

$$\left(2i\omega_2^*\tau_2^*I-\int_{-1}^{0}e^{2i\omega_2^*\tau_2^*\theta}d\eta(\theta)\right)E_1=2\tau_2^*(P_1,P_2,P_3)^T, \quad \text{(B.24)}$$

which leads to $K_1E_1=2(P_1,P_2,P_3)^T$, where

$$K_1=\begin{pmatrix}2i\omega_2^*+\mu_1+b\beta_1V^*e^{-2i\omega_2^*\tau_1^*} & 0 & b\beta_1S^*e^{-2i\omega_2^*\tau_1^*}\\ \left(\sigma-e^{-2i\omega_2^*\tau_1^*}\right)b\beta_1V^* & 2i\omega_2^*+\mu_1+\gamma+b\sigma\beta_1V^* & -b\beta_1S^*e^{-2i\omega_2^*\tau_1^*}-b\sigma\beta_1\left(\dfrac{\Lambda_1}{\mu_1}-I^*-S^*\right)\\ 0 & -b\beta_2\left(\dfrac{\Lambda_2}{\mu_2}-V^*\right)e^{-2i\omega_2^*\tau_2^*} & 2i\omega_2^*+\mu_2+b\beta_2I^*e^{-2i\omega_2^*\tau_2^*}\end{pmatrix}. \quad \text{(B.25)}$$

It follows that

$$E_1^{(1)} = \frac{2\Delta_{11}}{\Delta_1}, \quad E_1^{(2)} = \frac{2\Delta_{12}}{\Delta_1}, \quad E_1^{(3)} = \frac{2\Delta_{13}}{\Delta_1}, \tag{B.26}$$

where $\Delta_1 = |K_1|$,

$$\Delta_{11} = \begin{vmatrix} P_1 & 0 & b\beta_1 S^* e^{-2i\omega_2^*\tau_1^*} \\ P_2 & 2i\omega_2^* + \mu_1 + \gamma + b\sigma\beta_1 V^* & -b\beta_1 S^* e^{-2i\omega_2^*\tau_1^*} - b\sigma\beta_1\left(\frac{\Lambda_1}{\mu_1} - I^* - S^*\right) \\ P_3 & -b\beta_2\left(\frac{\Lambda_2}{\mu_2} - V^*\right)e^{-2i\omega_2^*\tau_2^*} & 2i\omega_2^* + \mu_2 + b\beta_2 I^* e^{-2i\omega_2^*\tau_2^*} \end{vmatrix},$$

$$\Delta_{12} = \begin{vmatrix} 2i\omega_2^* + \mu_1 + b\beta_1 V^* e^{-2i\omega_2^*\tau_1^*} & P_1 & b\beta_1 S^* e^{-2i\omega_2^*\tau_1^*} \\ \left(\sigma - e^{-2i\omega_2^*\tau_1^*}\right)b\beta_1 V^* & P_2 & -b\beta_1 S^* e^{-2i\omega_2^*\tau_1^*} - b\sigma\beta_1\left(\frac{\Lambda_1}{\mu_1} - I^* - S^*\right) \\ 0 & P_3 & 2i\omega_2^* + \mu_2 + b\beta_2 I^* e^{-2i\omega_2^*\tau_2^*} \end{vmatrix},$$

$$\Delta_{13} = \begin{vmatrix} 2i\omega_2^* + \mu_1 + b\beta_1 V^* e^{-2i\omega_2^*\tau_1^*} & 0 & P_1 \\ \left(\sigma - e^{-2i\omega_2^*\tau_1^*}\right)b\beta_1 V^* & 2i\omega_2^* + \mu_1 + \gamma + b\sigma\beta_1 V^* & P_2 \\ 0 & -b\beta_2\left(\frac{\Lambda_2}{\mu_2} - V^*\right)e^{-2i\omega_2^*\tau_2^*} & P_3 \end{vmatrix}. \tag{B.27}$$

Similarly, substituting expressions (B.19) and (B.23) into equation (B.21), we can have

$$\int_{-1}^{0} d\eta(\theta)E_2 = -\tau_2^*(Q_1, Q_2, Q_3)^T, \tag{B.28}$$

which leads to $K_2E_2 = (Q_1, Q_2, Q_3)^T$, where

$$K_2 = \begin{pmatrix} \mu_1 + b\beta_1 V^* & 0 & b\beta_1 S^* \\ -(1-\sigma)b\beta_1 V^* & \mu_1 + \gamma + b\sigma\beta_1 V^* & -\frac{(\mu_1+\gamma)I^*}{V^*} \\ 0 & -b\beta_2\left(\frac{\Lambda_2}{\mu_2} - V^*\right) & \mu_2 + b\beta_2 I^* \end{pmatrix}. \tag{B.29}$$

It follows that

$$E_2^{(1)} = \frac{\Delta_{21}}{\Delta_2}, \quad E_2^{(2)} = \frac{\Delta_{22}}{\Delta_2}, \quad E_2^{(3)} = \frac{\Delta_{23}}{\Delta_2}, \tag{B.30}$$

where $\Delta_2 = |K_2|$,

$$\Delta_{21} = \begin{vmatrix} Q_1 & 0 & b\beta_1 S^* \\ Q_2 & \mu_1 + \gamma + b\sigma\beta_1 V^* & -\frac{(\mu_1+\gamma)I^*}{V^*} \\ Q_3 & -b\beta_2\left(\frac{\Lambda_2}{\mu_2} - V^*\right) & \mu_2 + b\beta_2 I^* \end{vmatrix},$$

$$\Delta_{22} = \begin{vmatrix} \mu_1 + b\beta_1 V^* & Q_1 & b\beta_1 S^* \\ -(1-\sigma)b\beta_1 V^* & Q_2 & -\frac{(\mu_1+\gamma)I^*}{V^*} \\ 0 & Q_3 & \mu_2 + b\beta_2 I^* \end{vmatrix},$$

$$\Delta_{23} = \begin{vmatrix} \mu_1 + b\beta_1 V^* & 0 & Q_1 \\ -(1-\sigma)b\beta_1 V^* & \mu_1 + \gamma + b\sigma\beta_1 V^* & Q_2 \\ 0 & -b\beta_2\left(\frac{\Lambda_2}{\mu_2} - V^*\right) & Q_3 \end{vmatrix}. \tag{B.31}$$

Data Availability

The data included in this paper is only to demonstrate the validity of the results.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

This work was supported in part by the Fundamental Research Funds for the Central Universities, China (SWU020005), in part by the Natural Science Foundation Project of Chongqing CSTC (cstc2018jcyjAX0810), in part by the Science and Technology Research Program of Chongqing Municipal Education Commission (KJQN201801605), and in part by CQUE (KY201708B and KY20200146).

References

[1] K. S. Moore, "Vector-borne diseases: an ongoing threat," *The journal Nurse Practitioners*, vol. 15, no. 6, pp. 449–457, 2019.

[2] F. Catteruccia, "Malaria-carrying mosquitoes get a leg up on insecticides," *Nature*, vol. 577, no. 7790, pp. 319-320, 2020.

[3] J. Sachs and P. Malaney, "The economic and social burden of malaria," *Nature*, vol. 415, no. 6872, pp. 680–685, 2002.

[4] H.-M. Wei, X.-Z. Li, and M. Martcheva, "An epidemic model of a vector-borne disease with direct transmission and time delay," *Journal of Mathematical Analysis and Application*, vol. 342, no. 2, pp. 895–908, 2008.

[5] J. Xu and Y. Zhou, "Hopf bifurcation and its stability for a vector-borne disease model with delay and reinfection," *Applied Mathematical Modeling*, vol. 40, no. 3, pp. 1685–1702, 2016.

[6] Z. Hu, S. Yin, and H. Wang, "Stability and Hopf bifurcation of a vector-borne disease model with saturated infection rate and

reinfection," *Computational and Mathematical Methods in Medicine*, vol. 2019, Article ID 1352698, 17 pages, 2019.

[7] A. M. Niger and A. B. Gumel, "Mathematical analysis of the role of repeated exposure on malaria transmission dynamics," *Differential Equations and Dynamical Systems*, vol. 16, no. 3, pp. 251–287, 2008.

[8] L. Hviid, "Naturally acquired immunity to Plasmodium falciparum malaria in Africa," *Acta Tropica*, vol. 95, no. 3, pp. 270–275, 2005.

[9] Y. Xue and T. Li, "Stability and Hopf bifurcation for a delayed SIR epidemic model with logistic growth," *Abstract and Applied Analysis*, vol. 2013, Article ID 916130, 11 pages, 2013.

[10] Z. Zhang, S. Kundu, J. P. Tripathi, and S. Bugalia, "Stability and Hopf bifurcation analysis of an SVEIR epidemic model with vaccination and multiple time delays," *Chaos Solitons & Fractals*, vol. 131, article 109483, 2020.

[11] H. Wan and J. A. Cui, "A malaria model with two delays," *Discrete Dynamics in Nature and Society*, vol. 2013, Article ID 601265, 8 pages, 2013.

[12] J. L. Liu and T. L. Zhang, "Stability and Hopf bifurcation analysis of a plant virus propagation model with two delays," *Discrete Dynamics in Nature and Society*, vol. 2018, Article ID 7126135, 12 pages, 2018.

[13] L. Li, C. H. Zhang, and X. P. Yan, "Stability and Hopf bifurcation analysis for a two-enterprise interaction model with delays," *Communications in Nonlinear Science and Numerical Simulation*, vol. 30, no. 1–3, pp. 70–83, 2016.

[14] C. Xu and M. Liao, "Bifurcation analysis of an autonomous epidemic predator-prey model with delay," *Annali di Matematica Pura ed Applicata*, vol. 193, no. 1, pp. 23–38, 2014.

[15] C. Xu and P. Li, "Oscillations for a delayed predator-prey model with Hassell–Varley-type functional response," *Comptes Rendus Biologies*, vol. 338, no. 4, pp. 227–240, 2015.

[16] S. Khajanchi, M. Perc, and D. Ghosh, "The influence of time delay in a chaotic cancer model," *Chaos*, vol. 28, article 103101, 2018.

[17] C. Xu, M. Liao, and P. Li, "Bifurcation control of a fractional-order delayed competition and cooperation model of two enterprises," *Science China Technological Sciences*, vol. 62, no. 12, pp. 2130–2143, 2019.

[18] R. Kumar, A. K. Sharma, and K. Agnihotri, "Hopf bifurcation analysis in a multiple delayed innovation diffusion model with Holling II functional response," *Mathematical Methods in the Applied Sciences*, vol. 43, no. 4, pp. 2056–2075, 2020.

[19] Z. Zhang and H. Yang, "Stability and bifurcation in a stage-structured predator–prey system with Holling-II functional response and multiple delays," *International Journal of Computer Mathematics*, vol. 92, no. 3, pp. 542–561, 2015.

[20] B. D. Hassard, N. D. Kazarinoff, and Y. H. Wan, "Theory and applications of Hopf bifurcation," *London Mathematical Society Lecture Note Series*, vol. 41, 1981.

[21] C. Castillochavez and H. R. Thieme, "Asymptotically autonomous epidemic models," in *Mathematical Population Dynamics: Analysis of Heterogeneity, Theory of Epidemics*, vol. 1, Wuerz, Winnipeg, Canada, 1994.

[22] M. Jackson and B. M. Chen-Charpentier, "Modeling plant virus propagation with delays," *Journal of Computational and Applied Mathematics*, vol. 309, no. 47, pp. 611–621, 2017.

[23] P. Driessche and J. Watmough, "Further notes on the basic reproduction number," in *Mathematical Epidemiology Lecture Notes in Mathematics*, vol. 1945, Springer, Berlin, 2008.

[24] J. M. Heffernan, R. J. Smith, and L. M. Wahl, "Perspectives on the basic reproductive ratio," *Journal of the Royal Society Interface*, vol. 2, no. 4, pp. 281–293, 2005.

[25] Y. Song and S. Yuan, "Bifurcation analysis in a predator-prey system with time delay," *Nonlinear Analysis: Real World Applications*, vol. 7, no. 2, pp. 265–284, 2006.

[26] Y. Song and J. Wei, "Bifurcation analysis for Chen's system with delayed feedback and its application to control of chaos," *Chaos, Solitons & Fractals*, vol. 22, no. 1, pp. 75–91, 2004.

[27] J. Hale, "Theory and application of Hopf bifurcation," in *London Math Society Lecture Note Series*, Cambridge University Press, Cambridge, UK, 1977.

Hindawi
Computational and Mathematical Methods in Medicine
Volume 2021, Article ID 5528144, 17 pages
https://doi.org/10.1155/2021/5528144

Review Article

An Overview of Deep Learning Techniques on Chest X-Ray and CT Scan Identification of COVID-19

Woan Ching Serena Low,[1] **Joon Huang Chuah**,[1] **Clarence Augustine T. H. Tee**,[1] **Shazia Anis**,[2] **Muhammad Ali Shoaib**,[1] **Amir Faisal**,[3] **Azira Khalil**,[4] **and Khin Wee Lai**[2]

[1]*Department of Electrical Engineering, Faculty of Engineering, University of Malaya, 40603 Kuala Lumpur, Malaysia*
[2]*Department of Biomedical Engineering, Faculty of Engineering, University of Malaya, 40603 Kuala Lumpur, Malaysia*
[3]*Department of Biomedical Engineering, Faculty of Production and Industrial Technology, Institut Teknologi Sumatera, Lampung 35365, Indonesia*
[4]*Faculty of Science and Technology, Universiti Sains Islam Malaysia, 71800 Nilai, Negeri Sembilan, Malaysia*

Correspondence should be addressed to Khin Wee Lai; lai.khinwee@um.edu.my

Received 5 February 2021; Revised 19 April 2021; Accepted 19 May 2021; Published 7 June 2021

Academic Editor: Alireza Jolfaei

Pneumonia is an infamous life-threatening lung bacterial or viral infection. The latest viral infection endangering the lives of many people worldwide is the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), which causes COVID-19. This paper is aimed at detecting and differentiating viral pneumonia and COVID-19 disease using digital X-ray images. The current practices include tedious conventional processes that solely rely on the radiologist or medical consultant's technical expertise that are limited, time-consuming, inefficient, and outdated. The implementation is easily prone to human errors of being misdiagnosed. The development of deep learning and technology improvement allows medical scientists and researchers to venture into various neural networks and algorithms to develop applications, tools, and instruments that can further support medical radiologists. This paper presents an overview of deep learning techniques made in the chest radiography on COVID-19 and pneumonia cases.

1. Introduction

Pneumonia is life-threatening and one of the top diseases, which causes most deaths worldwide. It was projected that 1.4 million children die of pneumonia every year, in which 18% of the total children who died are below five years of age. In December 2019, at the epicentre in Wuhan, China, a novel coronavirus, severe acute respiratory syndrome–coronavirus-2 (SARS-CoV-2), causing COVID-19, emerged and is now a worldwide pandemic. As of 29th September 2020, COVID-19 has been confirmed in 215 countries and territories, involving 33,558,131 cases with 1,006,471 deaths globally, which is a 3% mortality rate [1]. Most reported infections were in the USA, Brazil, India, Russia, South Africa, Mexico, Peru, Colombia, Chile, Spain, and many others [1]. Countries have declared emergencies and national lockdown while cases have been reported to increase at an alarming rate [2].

Pneumonia is the inflammation of the alveoli inside the lungs [3]. The inflammation will build up fluid and pus that subsequently causes breathing difficulties. The patient will show symptoms such as shortness of breath, cough, fever, chest pains, chills, or fatigue. Antibiotics and antiviral drugs can treat bacterial and viral pneumonia. COVID-19 was originally called novel coronavirus-infected pneumonia (NCIP) [3]. The symptoms are similar to other variations of viral pneumonia and more [4] of which include rapid heartbeat, breathlessness, rapid breathing—also known as acute respiratory distress syndrome (ARDS), dizziness, and heavy perspiration [3]. COVID-19 damages the cells and tissues that line the air sacs in the lungs [3]. The damaged cells and tissues can disintegrate and clot the lungs causing

difficulties in breathing [3]. Nevertheless, an immediate diagnosis of COVID-19 and the consequent application of medication and treatment can significantly aid and prevent the deterioration of the patient's condition, which eventually can lead to death [5].

Hence, it is a challenge to diagnose a patient with COVID-19 via medical imaging. Deep learning models mimic human-level accuracy and precision in analysing and segmenting a medical image without human error [6]. However, deep learning cannot substitute medical professionals like physicians, clinicians, and radiologists in medical diagnosis [6], but it can assist medical experts in the field in executing and processing time-consuming works, such as determining chest radiographs for the signs of pneumonia and distinguishing the types of pneumonia and its severity [6].

2. Background of COVID-19

Coronaviruses are single-stranded ribonucleic acid (RNA) viruses, with the size of the virus approximately 26 to 32 kilobases. In late December 2019, a new (novel) coronavirus was identified in China, causing severe respiratory disease, including pneumonia. US Department of Health and Human Services/Centers for Disease Control and Prevention (CDC) reported that Chinese authorities declared an outbreak caused by a novel coronavirus, SARS-CoV-2 [7]. The coronavirus can cause mild to severe respiratory illness, known as Coronavirus Disease 2019 (COVID-19). The outbreak began in Wuhan, Hubei Province, China, and has spread to many countries worldwide—including Malaysia. World Health Organisation (WHO) declared COVID-19 a pandemic on 11 March 2020. CDC also stated the coronavirus could be spread mainly through close contact from person to person, face to face, and physically near each other within 6 feet [8].

The SARS-CoV-2 spreads more efficiently than influenza but not as efficiently as measles, one of the most contagious viruses. The respiratory ailment spreads throughout droplets of air. The infection is transmitted primarily via close contact and infects through respiratory droplets distributed in the air when a person coughs or sneezes. When a person contaminated with SARS-CoV-2 coughs, sneezes, sings, talks, or breathes, he or she produces respiratory droplets which range in size from large droplets visible to the human eye to smaller droplets. The tiny droplets can also form particles as they dry very quickly in the airstream [8]. Breathing difficulty is an indication of plausible pneumonia and requires prompt clinical deliberation and care. Research indicates that people suffering from COVID-19 often show hyperthermia and breathing problems [9]. Currently, there are no antibodies or definitive treatment for COVID-19 patients available to the public. The US Food and Drug Administration (FDA) had no authorised or approved vaccine to prevent COVID-19 [8] until 12 December 2020, when the Pfizer-BioNTech coronavirus vaccine, which offers up to 95% protection against COVID-19, has been authorised as safe, effective, and only for emergency use [10]. However, the World Health Organisation (WHO) encouraged the facilitation of vaccines by public persuasion instead of making the injections mandatory [11].

Early diagnosis of COVID-19 is critical to prevent human transmission of the virus to maintain a healthy population. Reverse transcription-polymerase chain reaction (RT-PCR) test is used to detect COVID-19 disease. It shows high specificity but is inconsistent with sensitivity in sensing the existence of the disease [12]. It demonstrates a certain proportion of false-negative results. However, when the pathological load is high during the symptomatic phase, the test is more accurate. The (RT-PCR) test kits are also limited in some geographical regions, especially third-world countries [13]. The turnaround time is 24 hours in major cities and is even longer in rural regions [9]. There is an urgency to explore other possibilities to distinguish the ailment and enable immediate referrals for the SARS-CoV-2-infected patient [9]. The chest X-ray plays a crucial role and is the first imaging technique to diagnose COVID-19 [14]. The virus presents on the Chest X-ray as ground-glass opacities, with peripheral, bilateral, and primary basal distribution [12]. These presentations seem comparable to those resulting from non-SARS-CoV-2-related viral, bacterial, fungal pneumonia [9, 12].

Furthermore, researchers found it problematic to differentiate viral pneumonia from other forms of bacterial and fungal pathogens [15]. Both chest X-ray and CT scan are not encouraged to be used as the primary diagnostic tool to screen/confirm and evaluate respiratory damage in COVID-19 because of the high risk and rapid increase in disease transmission [9, 13]. CT scans are discovered to be less explicit than RT-PCR but highly sensitive in sensing COVID-19 and can act as a fundamental role in disease analysis/treatment [13]. Nevertheless, the American College of Radiology has endorsed CT scans' practice as a primary-line assessment [16]. There are further concerns in using CT scans as a first-line test for the augmented risk of transmission, access, and cost, contributing to the recommendation [9]. As the pandemic became calamitous, radiological imaging is considered compulsory where portable chest X-rays are a useful and practical alternative [12]. However, the images' valuation placed a severe responsibility for radiological know-how, which is frequently lacking in regions with limited resources. Therefore, automated decision-making tools could be essential to appease some of this problem and to quantify and identify disease development [9].

2.1. Background on Deep Learning (DL). Artificial intelligence (AI) is a computer science branch that allows machines to execute human intelligence tasks. With the evolution of AI and Internet-of-Things, medical equipment has rapidly changed, which provides many possibilities in medical radiology. Machine learning (ML) techniques can achieve the objective of AI. It is the subset of AI to allow computer systems with the learning ability and implement tasks with the data automatically without manual programming. Deep learning (DL) is a subset of machine learning related to methods simulating the neurons of the human brain [17, 18]. The implementation of ML is to apply DL as an essential subject with its technology in classification, recognition, and identification of images or videos. The algorithm instructs the information to process patterns impersonating the human neural system. DL is currently an essential subject

with its technology in classification, recognition, and identification of images or videos. DL functions on algorithms for cognitive method simulation and data mining developing concepts [19]. DL maps input data consisting of hidden deep layers required to be labelled and analyzed concealed patterns within the complex data [20]. Between ML and DL, DL can automatically classify features and provide accurate results with high-end GPU help whereas ML requires a wider range of data to be preprocessed as the features need to be extracted manually. ML integrates various computational models and algorithms to mimic the human neural system whereas the DL-based network is more profound and is created with many hidden layers compared to conventional ANN. DL algorithms do not require many feature classifications and acquire directly from the data to display their higher problem-solving aptitudes. DL can interpret data and extract a wide range of dimensional features, notwithstanding if the features are visible or invisible to the naked human eye. This diminishes manual data preprocessing such as segmentation. DL can handle complex data representations and mimic trained physicians by identifying and detecting the features to make clinical decisions. DL architectures are applied in medical X-ray detection and various areas such as image processing and computer vision in medicine [17]. DL progresses in the medical sector to comprehend higher results, expand disease possibility, and execute valid real-time medical image [21, 22] in disease recognition systems [23]. Table 1 shows the neural network's significant contributions to deep learning [23, 24].

Figure 1 below shows the mind map of the types of machine learning and deep learning techniques created [25].

Convolutional neural network (CNN) most often apply to image processing problems where a computer identifies the object in an image. CNN can also be used in natural language processing projects as well. CNN modelling is adequate for processing and classifying images. A regular neural network has three layers: an input layer, a hidden layer, and an output layer. The input layer has different forms, whereas the hidden layer performs calculations on these inputs. The output layer delivers the outcome of the calculations and extractions. Each of the layers contains neurons and has its weight connected to the neurons in the previous layer. Hence, the data that is provided in the network does not produce assumptions via the output layer. However, the regular neural network cannot be applied if the data consists of images or languages. This is where convolutional neural network (CNN) comes in. CNN treats data as spatial data. Unlike regular neural network, the CNN neurons are not connected to every layer from the input layers to the hidden layers, and finally, the output layers only choose the neurons closest to it with the same weight. CNN upholds the spatial aspect of the dataset, which means that it undergoes a filtering process that simplifies complex images to better-processed images that are understood. The CNN is made up of many layers, consisting of several individual layers known as the convolutional layer, the pooling layer, and a fully connected layer. Inside, the layer of the CNN also consists of the rectified linear unit layer (ReLU). The ReLU layer activates the function to ensure nonlinearity as the data progresses through each layer in the network. Without ReLU, the data that is provided at the input layer would lose the dimensionality that is required in the network. The fully connected layer performs classification on the datasets. The CNN works by placing a filter over an array of image pixels and creating a convolved feature map. The analogy is like looking at an image through a window allowing specific features within the image to be seen. This is also known as the typical 2D convolutional neural network. The pooling layer reduces the sample size of the particular feature map, which speeds up the process by reducing the parameters the network needs. The output is the pool featured map, consisting of two execution methods, i.e., max pooling and average pooling. Max pooling takes the maximum input of a particular convolved feature, whereas the average pooling takes the convolved feature's average. The next step is feature extractions, whereby the network creates a picture of the image data based on its mathematical rules. The images' classification requires the network to move into the fully connected layer by flattening and simplifying the images. A complex set of neural network connections can only process linear data. If the data is unlabelled, unsupervised learning methods can be applied by using autoencoders to compile the data in a low dimension space performing calculations, then, additional layers are reconstructed to upsample the existing data.

Table 1: Significant contributions of the neural network to deep learning [23, 24].

Milestone/contribution	Year
McCulloch-Pitts neuron	1943
Perceptron	1958
Backpropagation	1974
Neocognition	1980
Boltzmann machine	1985
Restricted Boltzmann machine	1986
Recurrent neural networks	1986
Autoencoders	1987
LeNet	1990
LSTM	1997
Deep belief networks	2006
Deep Boltzmann machine	2009

CNN is the reason DL is so well known, but it has limitations and fundamental drawbacks. The max-pooling or successive convolutional layers lose valuable information. CNN needs a large amount of data to work, and it loses information in the pooling area, which in turn reduces spatial resolution, resulting their outputs to be invariant to small changes in the inputs. Currently, the issue is addressed by building complex architectures around CNNs to recover the lost information.

Generative adversarial network (GAN) trains two networks which comprise the artificial data samples that resemble data in the training set and the discriminative network that distinguishes the artificial and the original model: in simple means, GAN has a generator data, and the other

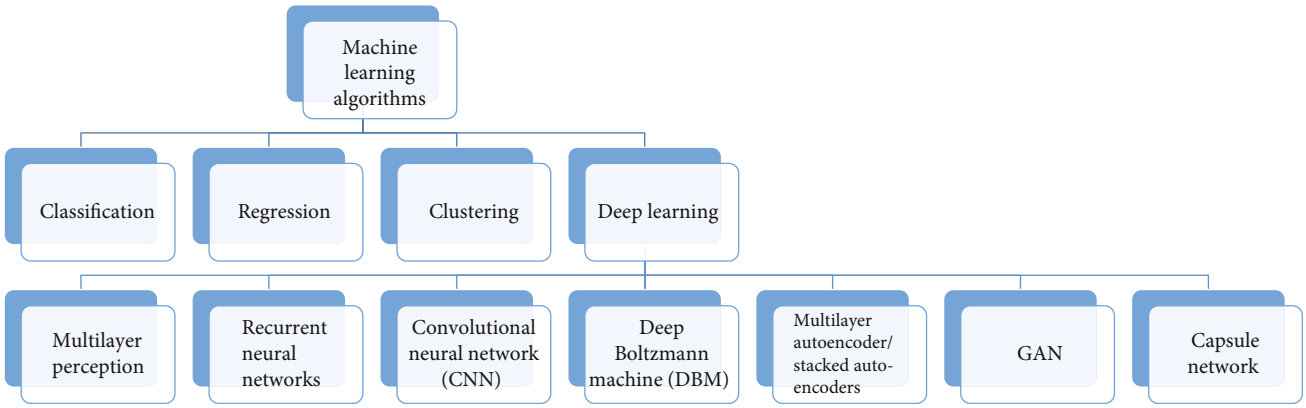


Figure 1: Mindmap of machine learning (ML) algorithm created by Robert Herman from mind meister [25].

is the discriminative data. The generator data is the counterfeiter that consistently produces artificial data, and the discriminator will try to expose the counterfeiter. Each time the discriminator manages to identify the image as a counterfeit, the generator will keep improving it until it is as accurate as possible.

Capsule networks is an artificial neural network that is significantly new. It is a network that applies local capsules in an artificial neural network that consists of complicated internal computations on the inputs and encapsulates these computations' results into a small vector of highly informative outputs. CapsNet architecture reached state-of-the-art performance on MNIST and had better performance than CNNs on MultiMNIST [26].

3. Radiology Perspective of Coronavirus Disease 2019 (COVID-19)

In December 2019, a lower respiratory tract feverish illness of unfamiliar derivation was informed in a cluster of patients in Wuhan City, Hubei Province, China. Coronavirus disease 2019 (COVID-19) is accountable for this epidemic to date. Other corresponding pulmonic conditions have been documented as being triggered by other strains of the coronavirus family. The most notable instances are the severe acute respiratory syndrome (SARS) and the Middle East respiratory syndrome (MERS). The SARS epidemic was under control with no human contaminations reported since 2003 whereas minor MERS occurrences continue to be stated. Hence, imaging is an essential analytical procedure tool observing disease development and coronavirus-related pulmonary syndrome [27]. Imaging structures in critical and chronic phases of SARS and MERS are inconsistent and inexplicit [28]. The first accounts of imaging discoveries of COVID-19 have also been described as inconclusive [29–31]. Researchers are conducting various studies to distinguish further and identify the imaging features of this new coronavirus syndrome, but the information is still inadequate.

The incident of COVID-19 intensified beyond human beings comprehension; more clusters and incidences are reported daily by the several ten thousand in some parts of the world. The disorder's etiologic and medical structures are comparable to SARS and MERS; the knowledge and aptitude from those pulmonary syndromes can support handling the sharp increase of COVID-19 eruption. This review segment will allow us to be familiar with the radiologist and imaging spectrum of coronavirus syndromes and discuss the reported imaging features of COVID-19.

SARS was discovered in 2003 as the first epidemic of the new era in Guangdong Province, China, which its clinical discovery presented as novel viral pneumonia. The clinical disease-infested 8,422 individuals demanded 916 lives before it was confined, and no occurrence has been reported ever since [32]. MERS was revealed in Saudi Arabia, where a patient's sputum consisted of the novel coronavirus in 2012 [32]. The disease has infected 2,492 individuals worldwide, and 858 human lives were lost, as the latest discovery was reported in December 2019 [32].

There are various imaging features of SARS and MERS that share similarity to one another, but some differences are shown in Table 2. The analysis of COVID-19 is hypothesised on the foundation of indications of pneumonia (e.g., dry cough, lethargy, myalgia, malaise, and dyspnea similar to symptoms of SARS and MERS) as well as past travelling activities to China or acquaintance with a COVID-19 patient. The development of the diseases and their severity rely on chest imaging to acquire valuation, discovery, and identification. A portable chest X-ray (CXR) is used as the first-line modality for COVID-19 patients instead of CT scans, as CT scans are applied in specific situations. Portable chest X-ray (CXR) has the benefit of discarding patients' need to travel from one location to another and diminish the use of personal protective equipment (PPE). The arrangement is to avoid nonessential imaging and transportations to the radiology department. Czawlytco et al. discovered that chest X-ray is insensitive in the early detection of COVID-19 with a sensitivity of only 59% [33]. Chest X-ray is not recommended for patients with flu/influenza-like symptoms. It is also not recommended to be used on confirmed COVID-19 patients with mild symptoms. Therefore, chest X-ray is designated for COVID-19 patients with acute respiratory status or COVID-19 patients with mild symptoms but has high-risk factors for developing severe disease. Chest radiography and tomography cannot be used as first-line screening or diagnosis in COVID-19, even with a normal chest X-ray and CT images, the possibility of COVID-19 cannot be ruled

Table 2: Comparison of clinical and radiological features of COVID-19, SARS, and MERS [32].

Feature	COVID-19	SARS	MERS
Clinical sign or symptom			
Fever or chills	Yes	Yes	Yes
Dyspnea	Yes	Yes	Yes
Malaise	Yes	Yes	Yes
Myalgia	Yes	Yes	Yes
Headache	Yes	Yes	Yes
Cough	Dry	Dry	Dry or productive
Diarrhoea	Uncommon	Yes	Yes
Nausea or vomiting	Uncommon	Yes	Yes
Sore throat	Uncommon	Yes	Yes
Arthralgia		Yes	Uncommon
Imaging finding			
Acute phase			
Initial imaging			
Normal	15–20% of patients	15–20% of patients	17% of patients
Abnormalities			
Common	Peripheral multifocal airspace opacities (GGO, consolidation, or both) on chest radiography and CT.	Peripheral multifocal airspace opacities (GGO, consolidation, or both) on chest radiography and CT.	Peripheral multifocal airspace opacities (GGO, consolidation, or both) on chest radiography and CT.
Rare	Pneumothorax	Pneumothorax	Pneumothorax
Not seen	Cavitation or lymphadenopathy	Cavitation or lymphadenopathy	Cavitation or lymphadenopathy
Appearance	Bilateral, multifocal, basal airspace; normal chest radiography findings (15%)	Bilateral, multifocal basal airspace on chest radiography or CT (80%); isolated unilateral (20%)	Unilateral, focal (50%); multifocal (40%); diffuse (10%)
Follow-up imaging appearance	Persistent or progressive airspace opacities	Unilateral, focal (25%); progressive (most common, can be unilateral and multifocal or bilateral with multifocal consolidation)	Extension into upper lobes or perihilar areas, pleural effusion (33%), interlobular septal thickening (26%)
Indications of poor prognosis	Consolidation (vs. GGO)	Bilateral (like ARDS), for or more lung zones, progressive involvement after 12 d	Greater involvement of the lungs, pleural effusion, pneumothorax
Chronic phase	Unknown, but pleural effusion and interlobar septal thickening have not yet been reported		
Transient reticular opacities		Yes	Yes
Air trapping		Common (usually persistent)	
Fibrosis	More than one-third of patients	Rare	One-third of patients

Note: SARS: severe acute respiratory syndrome; MERS: Middle East respiratory syndrome; COVID-19: coronavirus disease 2019; GGO: ground-glass opacity; ARDS: acute respiratory distress syndrome [32].

out as a patient might be asymptomatic, and the lung condition maintains to be expected. However, information of COVID-19 patients initially declared hostile on the virus using the real-time reverse transcriptase-polymerase chain reaction (RT-PCR) was discovered to have COVID-19 via early CT findings [32]. In the meantime, initial findings in imaging may show normal conditions of the lungs. Hence, standard chest imaging does not rule out the possibility of being infected with SARS-CoV-2 [32].

3.1. Artificial Intelligence on Chest X-Ray (CXR) and CT Scans. With the struggle against the SARS-CoV-2 rapid infection, active screening and immediate medical response for the infected patients are desperately needed. RT-PCR is a common screening application which is manual, time-consuming, intricate, and arduous with only a 63% positivity rate [34, 35]. Research regarding early identification of COVID-19 by using CXR and other imaging modalities is still in development. The Guardian reported information shared by a respiratory physician that SARS-CoV-2 pneumonia is different from common viral pneumonia cases [36]. However, the images of several viral cases of pneumonia are comparable with other infectious and inflammatory lung diseases [34]. The COVID-19 symptoms being similar to other viral

pneumonia can result in wrong diagnosis and prognosis in many hospitals, especially in the emergency department which is overloaded and understaff [34].

Today, many biomedical problems and complications such as brain tumour detection, lung disease detection, breast cancer detection, and other oncological emergencies are using artificial intelligence (AI) solutions [34]. Convolutional neural network (CNN), a deep learning technique, has been advantageous in revealing image features that are not obvious in the original image [34]. The accuracy of the deep learning algorithm relies on imaging quality, and CNN can improve imaging quality in low-light images from a high-speed video endoscopy, discover pulmonary nodules through CT images, identify paediatric pneumonia from CXR images, and automatically labelling of polyps in a colonoscopy and cystoscopic image analysis from videos [34]. Hence, only confirmed positive COVID-19 patients' images were selected. Wang et al. (2017) have shown to accumulate datasets that allow significant developments in medical imaging tools to progress in the prediction of various pneumonia and the outcome towards the infected patient [37, 38]. Rajpurkar et al. (2017) and Cohen et al. (2019) works on both organised models to foresee various pneumonia [37, 39, 40]. Deep learning models and algorithms are tools that can be developed for triage cases during the shortage of physical tests, particularly RT-PCR [37, 41, 42]. The American College Radiology (ACR) only recommended portable CXR in an ambulant care facility when required and strongly discourage CT to apply and inform decisions on a suspected COVID-19 patient and whether or not to conduct RT-PCR test, admit the patient, provide other treatment, and dissuade the patient from being quarantines or others [33]. However, deep learning models and algorithms should predict patient outcomes and permitting the physician to immediately facilitate care and management [37, 43]. COVID-19 can be considered in extraordinary extreme situations, where physicians could be faced with decisions to select which patient to assign for which healthcare resources based on the severity level [43]. The tools would serve to monitor the development of SARS-CoV-2 positive patients' ailment evolution [37].

3.2. Approached Techniques and Convolutional Neural Network Architecture. Deep learning (DL) is a subsection of machine learning, and a convolutional neural network is a type of deep learning commonly applied in the computer vision domain. Examples of CNN architectures are LeNet, AlexNet, GoogLeNet, Visual Geometry Group (VGG) Net, ResNet, and others [44]. The goal is to apply deep learning neural network architectures to create practical applications to improve diagnosis and prognosis performance [44].

Deep CNN was created with LeNet designed to recognise handwritten digits. However, LeNet has limitations, and thus, its successor AlexNet was the first deep CNN that accomplished outstanding results for the organisation and recognition tasks on the image. Due to hardware limitations in early 2000, deep CNN architectures' learning capacity was restricted to small sample size images. AlexNet was made applicable to all types of images—its depth was extended from LeNet's five layers to eight layers: five convolutional layers, two fully connected hidden layers, and one fully connected output layer generalised for different image resolutions. However, it caused overfitting issues. The overfitting issue was fixed with the dropout algorithm, which arbitrarily eliminated some transformational units during the training process. DenseNet is a modern CNN architecture that requires fewer visual object recognition parameters. It is the product of the previous layer that combines with the output of a future layer. The objective of DenseNet is to recognise visual objects by densely connecting all the layers. ResNet is known as the residual net, which divides a layer into two branches, where one branch does nothing to the signal, and the other processes ResNet adds the previous layer with the future layers. Usually, a deep neural network tends to randomly overfit and sometimes produce more preliminary results than a network with a few layers.

CNN is based on biological processes of the visual cortex of the human and the animal brain. CNN consists of multiple layers where a higher layer is connected to a lower layer to study abstract features of the images by considering the spatial relationships between the receptive fields. This allows CNN to recognise patterns and identify images within the layers of images. Various CNN models apply different layers, number of neurons, and receptive fields in the respective layers and algorithm [44]. Integrating transfer learning into the technique modifies the CNN models applied to pretrain many radiology image datasets to diagnose COVID-19 problems [44]. This technique bypasses the hassle to train all the images from scratch everytime new cases or images are identified. However, this method is not valid with the amount of radiology images dataset available for the public.

Based on the studies in Table 3, several studies use deep learning for COVID-19 diagnosis using radiology images.

Table 3 includes some research conducted with deep learning models using two types of medical images, i.e., chest-X-ray (CXR) and CT images. Based on the table, the majority of the researchers used CXR images because of their availability. The CXR requires low memory space and high results performance which reassure researchers to apply these images into the respective deep learning models. There are a total of 52 researches using various deep learning methods to achieve results. Out of the 52 journals mentioned above, 34 of the studies used CXR images, 17 studies used CT images, and 4 of the studies used CT and CXR images. More CXR images from COVID-19 patients found in the public databases encouraged researchers to study deep learning utilising these images. Journals from the medical field often mentioned that CT images show higher accuracy performance, but these accuracies were debunked because it was not explicitly shown in the deep learning-based CAD systems. The nature of the CT images that produce many cross-sections just for one patient result in high memory usage for the facility to handle. In general, CT images were previously deemed more accurate than CXR images because the CT images' cross-section images are individually labelled. Hence, studies that utilised the combination of CT and CXR images show promising results. However, studies with 3D data have lower performance than 2D data, mainly because there are primarily 2D data available for the public to use.

Table 3: Summary of deep learning methods and CNN architectures for COVID-19 using radiology images. CT images are computer tomography images, and CXR images are chest X-ray images.

No.	Papers	Data	Types of images	AI methods to establish the algorithm	CNN architecture	Results for detecting COVID
1	[44, 45]	Total images: 4,356, COVID-19 images: 1,296, pneumonia images: 1,735, nonpneumonia images: 1,325	CT	3D deep learning	ResNet-50 and COVNet	Area under the curve (AUC): 0.96
2	[44, 46]	Total images: 618, COVID-19 images: 219, influenza-A (H1N1, H3N2, H5N1, H7N9, and others), images: 224, normal healthy lungs images: 175	CT	3D CNN model for segmentation	Location-attention network and ResNet-18	Accuracy of 86.7%, average time: 30 s
3	[44, 47]	(PA) posterior-anterior images: 5,941, normal images: 1,583, bacterial pneumonia images: 2,786, non-COVID-19 viral pneumonia images: 1,804, COVID-19 images: 68	CXR	Drop weights based Bayesian CNNs	Bayesian ResNet50V2	Accuracy of 89.92%
4	[44, 48]	COVID-19 images: 453, training images: 217	CT	Inception migration-learning model		Internal validation: accuracy: 82.9%, specificity: 80.5%, sensitivity: 84%; External testing dataset: accuracy: 73.1%, specificity: 67%, sensitivity: 74%
5	[44, 48]	Total images: 1,065, COVID-19 images: 325; viral pneumonia images: 740	CT	Modified inception transfer-learning model		Accuracy: 79.30%, specificity: 0.83, sensitive: 0.67
6	[44, 49]	Total patients: 133, severe/critical patients: 54, nonsevere/critical patients: 79	CT	Multilayer perception and long short term memory (LSTM)		Area under the curve (AUC): 0.954
7	[44, 50]	Total images: 4,266, COVID-19 images: 2,529, CAP images: 1,338, influenza A/B images: 135, standard images: 258, total patients: 3,177, COVID-19 patients: 1,502, influenza A/B patients: 83, CAP patients: 1,334, healthy subjects: 258	CT	2D deep learning CNN	ResNet 152	Accuracy: 94.98%, AUC 97.71%, sensitivity: 90.19%, specificity: 95.76%, the average time is taken to read: 2.73 s
8	[44, 51]	Total 1,136 cases from 5 hospitals, COVID-19 images: 723, non-COVID-19 images: 413	CT	3D deep learning method	UNet ++ & ResNet-50	Specificity: 0.922, sensitive: 0.974
9	[44, 52],	COVID-19 patients: 50, ordinary people: 50,	CXR	5 pretrained CNN	ResNet-50, ResNet101, ResNet52, InceptionV3, and inception-ResNetV2	ResNet-50: accuracy: 98.0%

Table 3: Continued.

No.	Papers	Data	Types of images	AI methods to establish the algorithm	CNN architecture	Results for detecting COVID
10	[44, 53]	Total images:13,975, total patients:13,870	CXR	Deep learning CNN	COVID-net	Accuracy: 92.4%
11	[44, 54]	Total patients: 157	CT	CNN	ResNet-50	Area under the curve (AUC): 0.996
12	[34, 44]	Normal images: 1,341, viral pneumonia images: 1,345, COVID-19 images: 190	CXR	CNN	AlexNet, ResNet-18, DenseNet-201, SqueezeNet	Accuracy: 98.3%
13	[44, 55]	Total COVID-19 images: 531, CXR images: 170, CT images: 361	CT and CXR	CNN with transfer learning	Pretrained AlexNet	Accuracy: CXR images: 98.3%, CT image: 94.1%
14	[6]	Total images: 5,232, normal images: 1,346, bacterial pneumonia images: 2,538, viral pneumonia images: 1,345	CXR	Deep learning framework using transfer learning	Pretrained on ImageNet, trained using AlexNet, ResNet18, inception V3, DenseNet121, GoogLeNet, and ensemble model	Ensemble model: accuracy: 96.4%, recall: 99.62% (unseen data)
15	[5]	Total images: 5,247, bacterial pneumonia images: 2,561, viral pneumonia images 1,345, normal images: 1,341	CXR	Pretrained deep CNN and used for transfer learning	AlexNet, ResNet18, DenseNet201, and SqueezeNet	DenseNet201 accuracy: normal and pneumonia: 98%, normal images, bacterial, and viral pneumonia: 93.3%, bacterial and viral pneumonia: 95%
16	[17]	Total images: 306, COVID-19 images: 69, normal images: 79, bacterial pneumonia images: 79, viral pneumonia images: 79. The dataset number increases to 8,100 images after using the GAN network.	CXR	Deep transfer learning: using GAN network to generate more images to help detect the virus. Three deep transfer models.	AlexNet, GoogLeNet, Restnet18 with performance measures in different scenario and classes	GoogLeNet accuracy: 80.56%
17	[56]	Dataset was collected from medRix and bioRxiv; COVID-19 images: 349, total patients: 216	CT	Multitask learning and self-supervised	DenseNet-169, ResNet-50	F1 score: 0.90, AUC: 0.98, accuracy: 0.89
18	[36]	Total images: 2,200, COVID-19 images: 800, viral pneumonia images: 600	CT	Machine learning technique using Microsoft Azure	ResNet	High accuracy: 91%, overall accuracy: 87.6%
19	[57]	Total images: 15,495, normal images: 12,544, COVID-19 image: 2,951	CXR	CNN model	UNet, UNet++, DLA, DenseNet-121, CheXNet; inception-v3, ResNet-50	F1 score: 85.81%, sensitivity: 98.37%, specificity: 99.16%
20	[58]	Diverse datasets from a different source	CT	Deep fully convolutional networks (FCN)	UNet, ResDense FCN	DSC: 0.780, sensitivity: 0.822, specificity: 0.951
21	[59]	Total images: 954, COVID-19 images: 308, normal images: 323, pneumonia images: 323 images	CXR	Deep learning modules using stacked architecture concept	DenseNet; GoogleNet	Sensitivity: 0.91, specificity: 0.95, F1 score: 0.91, AUC: 0.97

Table 3: Continued.

No.	Papers	Data	Types of images	AI methods to establish the algorithm	CNN architecture	Results for detecting COVID
22	[52]	Total images: 7,406, COVID-19 images: 341, normal images: 2,800, viral pneumonia images: 1,493, bacterial pneumonia images: 2,772	CXR	2D five pretrained CNN based models	ResNet50, ResNet101, ResNet152, InceptionV3, and inception-ResNetV2	COVID-19 and normal: accuracy: 96.1%, COVID-19 and pneumonia accuracy: 99.5%, COVID-19 and bacterial accuracy: 99.7%
23	[60]	Total images (COVID-19, pneumonia, and normal): 1,266, COVID-19 images: 924	CT	3D pretrained the deep learning system and validate it.	DNN	Sensitivity (train): 78.93%, specificity (train): 89.93%, sensitivity (val): 80.39%, specificity (val): 81.16%
24	[61]	Total images (COVID-19, bacterial, and normal): 275, COVID-19 images: 88	CT	2D pretrained ResNet 50 using the feature pyramid network (FPN)	DRE-net	Sensitivity: 93%, specificity: 96%, accuracy: 99%
25	[62]	Total images: 624, COVID-19 images: 50	CXR	2D GAN + TL	AlexNet, GoogLeNet, ResNet18, SqueezeNet	Accuracy: 99%
26	[63]	Total images (COVID-19, bacterial, and normal): 1,427, COVID-19 images: 224, bacterial and viral pneumonia images: 714	CXR	2D transfer learning (TL)	VGG19, MobileNet, Inception, Xception, Inception ResNet v2.	Sensitivity: 98.66%, specificity: 96.46%, accuracy: 94.72%
27	[64]	Total images (COVID-19, pneumonia, normal): 6,008, COVID-19 images: 184	CXR	2D transfer learning (TL)	Three ResNet models	Accuracy: 93.9%
28	[65]	Total images (COVID-19, pneumonia, and normal): 8,850, COVID-19 images: 498	CXR	2D convolutional autoencoder (CAE)	AE: COVIDomaly	Accuracy: 76.52%
29	[66]	Total images (COVID-19, pneumonia, and normal): 2,905, COVID-19 images: 219	CXR	2D	CNN + k-NN + SVM	Accuracy: 98.70%
30	[67]	Total images (COVID-19, pneumonia, and normal): 2,905, COVID-19 images: 219	CXR	2D using hyperparameters Bayesian optimisation algorithm	ANN + AlexNet	Sensitivity: 89.39%, specificity: 99.75%, accuracy: 98.97%, F-score: 96.72%
31	[68]	Total images (COVID-19, pneumonia, and normal): 502, COVID-19 images: 180	CXR	2D patch-based convolutional neural network	ResNet-18	Sensitivity: 76.90%, specificity: 100.00%
32	[69]	Total images (COVID-19, pneumonia, and normal): 2,905, COVID-19 images: 219	CXR	2D	Ensemble: Resnet50 and VGG16	Sensitivity: 91.24%, specificity: 99.82%
33	[70]	Total images (COVID-19 and normal): 2,492, COVID-19 images: 1,262	CT	2D	TL and DenseNet201	Accuracy: 99.82%
34	[71]	COVID-19, pneumonia, and normal images from Cohen et al. [37]	CXR	2D	Xception	Sensitivity: 97.09%, specificity: 97.29%, accuracy: 97.40%

Table 3: Continued.

No.	Papers	Data	Types of images	AI methods to establish the algorithm	CNN architecture	Results for detecting COVID
35	[72]	Total images (COVID-19 and normal): 380, COVID-19: 180	CXR	2D	5 pretrained models+ SVM	Accuracy: 94.7%
36	[73]	Total images (COVID-19, pneumonia, normal, and non-COVID-19): 2,905, COVID-19 images: 219	CXR	2D pretrained models such as ResNet101, Xception, InceptionV3, MobileNet, and NASNet	InstaCovNet-19	Accuracy: 99.08%, accuracy: 99.53%
37	[74]	Datasets contain COVID-19, pneumonia and normal images.	CXR	2D	5 pretrained CNNs	Accuracy: 95.00%
38	[75]	Datasets contain bacterial pneumonia, non-COVID viral pneumonia, and COVID-19 images.	CXR	2D	5 COVID-CAPS	Sensitivity: 90%, specificity: 95.8%, accuracy: 95.7%
39	[76]	Total images (COVID-19 and normal): 5,000, COVID-19 images: 184	CXR	2D	5 TL + pretrained models	Sensitivity: 100%, specificity: 98.38%
40	[55]	Total images (COVID-19 and normal): 526, COVID-19 images: 238	CXR + CT	2D	TL + AlexNet model	Sensitivity: 72%, specificity: 100%, accuracy: 94.1%
41	[77]	Total images (COVID-19 and normal): 320, COVID-19 images: 160	CXR + CT	2D Apache spark framework	TL + inceptionV3 & ResNet5	Sensitivity: 72%, specificity: 100%, accuracy: 99.01%
42	[78]	Total images (COVID-19, pneumonia, and normal): 4,575, COVID-19 images: 1,525	CXR	2D CNN used for deep feature extraction, and LSTM is used for detection using the extracted feature	LSTM+CNN	Sensitivity: 99.2%, specificity: 99.9%, accuracy: 99.4%
43	[79]	Dataset 1 images (COVID-19, pneumonia, and normal): 4,448, COVID-19 images: 2,479, dataset 2 images (COVID-19, pneumonia, and normal): 101, COVID-19 images: 52	CXR	2D	3D inception V1	Dataset 1: accuracy: 99.4%; dataset 2: sensitivity: 98.08%, specificity: 91.30%, accuracy: 93.3%
44	[80]	Total images (COVID-19, pneumonia, and normal): 1,343, COVID-19 images: 446	CXR	2D	Conditional GAN: LightCovidNet	Accuracy: 97.28%
45	[81]	Total images (COVID-19 and normal): 8,504, COVID-19 images: 445	CXR	2D	TL VGG-16 model	Sensitivity: 98.0%, specificity: 100.00%, accuracy: 94.5%
46	[82]	Total images (COVID-19 and normal): 746, COVID-19 images: 349	CT	2D	TL+ ensemble of 15 pretrained models: EfficientNets(B0-B5), NasNetLarge, NasNetMobile, InceptionV3, ResNet-50, SeResnet 50, Xception, DenseNet121, ResNext50, and Inception_resnet_v2	Accuracy: 85.0%

Table 3: Continued.

No.	Papers	Data	Types of images	AI methods to establish the algorithm	CNN architecture	Results for detecting COVID
47	[83]	Total images (COVID-19 and normal): 2,482, COVID-19 images: 1,252	CT	2D	AE + random forest	Specificity: 98.77%, accuracy: 97.87%
48	[84]	Total images (COVID-19 and normal1): 50, COVID-19 images: 25	CXR	3D	COVIDX-net	Sensitivity: 100.00%, specificity: 80.00%
49	[85]	Total images (COVID-19 and Normal): 800, COVID-19 images: 400	CXR	2D using modern and traditional machine learning methods: (ANN), (SVM), linear kernel and (RBF), k-nearest neighbor (k-NN), decision tree (DT), and CN 2 rule inducer techniques	Deep learning models: MobileNets V2, ResNet50, GoogleNet, DarkNet, and Xception	ResNet50 accuracy: 98.8%
50	[86]	Total images (COVID-19 and Normal): 800, COVID-19 images: 400	CXR	2D CLAHE and Butterworth bandpass filter was applied to enhance the contrast and eliminate the noise.	The hybrid multimodal deep learning system COVID-deep net system.	Sensitivity: 99.9%, specificity: 100.0%, accuracy: 99.3%
51	[87]	Datasets from Cohen et al. [37]. Total images (COVID-19 and normal): 800, COVID-19 images: 400	CXR	2D benchmarking and diagnostic models: decision matrix that embedded a mix of 10 evaluation criteria and 12 diagnostic models, also known as multicriteria decision making (MCDM)	TOPSIS is applied for benchmarking and ranking purpose, while entropy is used to calculate the criteria's weights. SVM is selected as the best diagnosis model	Coefficient value: 0.9899
52	[88]	Total images (COVID-19 and normal): 800, COVID-19 images: 400	CXR	2D hybrid deep learning framework, pretrained deep learning models incorporating of a ResNet34, and high-resolution network model	COVID-CheXNet system	Sensitivity: 99.98%, specificity: 100.0%, accuracy: 99.99%

The table also shows that deep learning models produced more stable results with more data.

3.3. COVID-19 Radiology Data Sources for Potential Modelling. This section describes the radiology imaging data source available for researchers to exploit the capabilities of deep learning techniques using CNN architectures to overcome COVID-19. The variability of the data requires different AI methods to study. Radiology images like CXR and CT images are high-dimensional data requiring CNN-based models to process the images like LeNet, AlexNet, GoogLeNet, VGGNet, and ResNet [44]. AlexNet is a category of CNN designed by Alex Krizhevsky in 2012. It is a popular CNN that sets the essential milestones to its incomers like network-in network [89] by Lin et al. [90], VGGNet [91] by Simonyan et al., and GoogLeNet (Inception v-1) by Szegedy et al.

CNN architecture application requires a large dataset for training, testing, and validating. Table 4 describes the available data sources for COVID-19 radiology images, mainly CXR and CT images.

The data sources depicted in Table 4 are the standard open-source radiology images available for the public to access, study, and characterise using CNN architectures. However, based on the table, there are minimal COVID-19 data to comprehensively utilise AI techniques to conduct an intensive study. This creates concerns and difficulties when utilising these techniques in real-world practice with a limited number of datasets available.

4. Challenges in the Interpretation and Application of Imaging Features of COVID-19 and Suggestions to Overcome

In theory, utilising AI is to eliminate fake news that can be found on the worldwide web and various social media platforms to ensure authenticity, responsible, and dependable

Table 4: Available data sources about COVID-19 radiology images for both chest X-ray and CT images.

No.	Sources	Data type	No. of images	Image type	Types of images	Links
1	J. P. Cohen's GitHub	Viral pneumonia (SARS, varicella, influenza) and COVID-19, bacterial pneumonia (Streptococcus spp., Klebsiella spp., Escherichia coli, Mycoplasma spp., Legionella spp., unknown, Chlamydophilla spp.) and COVID-19, fungal (Pneumocystis spp., lipoid) and COVID-19	Raw images: 910, annotated images: 210	jpg and png	CXR	https://github.com/ieee8023/covid-chestxray-dataset
2	European Society of Radiology	Total cases or images unknown	N/A	pdf	CXR and CT	https://www.eurorad.org/advanced-search?search=COVID
4	Kaggle	Posterior-anterior (PA), anterior-posterior (AP) lateral for X-rays and axial or coronal for CT scans	Normal images: 1,576, pneumonia ARDS images: 2, viral pneumonia images: 1,493, COVID-19 images: 58, SARS images: 4, bacterial pneumonia images: 2,772, bacterial Streptococcus images: 5	png, jpg, jpeg, and others	CXR and CT	https://www.kaggle.com/bachrr/covid-chest-xray
5	UCSD-AI4H	Total: 349 images from 216 patients	COVID-19 images: 349, non-COVID-19 images: 397	jpg and png	CT	https://github.com/UCSD-AI4H/COVID-CT
6	MedSeg	Images were segmented by a radiologist using 3 labels: ground-glass (mask value = 1), consolidation (=2), and pleural effusion (=3).	Image volumes—9 volumes, a total of >800 slices, COVID-19 masks 350 annotated slices. Lung masks > 700 annotated slices	jpg	CT	http://medicalsegmentation.com/covid19/
7	COVID-19 Radiography Database		COVID-19 images: 219, normal images: 1,341, viral pneumonia images: 1,345	png	CXR	https://www.kaggle.com/tawsifurrahman/covid19-radiographydatabase

information about the pandemic. However, scientists face many challenges and limitations shown in Table 5 below to produce ethical and reliable results for the public.

When implementing a DL model, test and train images from the same goal are used to distribute data and predict the medical images into their respective categories. The idea is impossible to achieve due to limited data availability or weak labels [9]. Despite many cases happening worldwide, we have very limited COVID-19 CXR or CT image data publicly available. Therefore, it is difficult to train the DL models and distinguish the images between COVID-19-related CXR and CT and non-SARS-CoV-2 viral, bacterial, and other pathogen-related CXR and CT images. The Radiological Society of North America (RSNA) [97] and Imaging COVID-19 AI Initiative in Europe [98] are aimed at providing easily accessible data to the public. These data allow various features across categories to enhance interclass variance, leading to better DL performance. Due to lack of data, the model will overfit and produce weakly generalised results [99]. Hence, data augmentation has been proven to be effective in training discriminative DL models. Examples of data augmentation techniques are flipping, rotating, colour jittering, random cropping, elastic distortions, and generative adversarial network- (GAN-) based synthetic data generation [100]. Medical images found in ImageNet have different visual characteristics showing high interclass similarities [101]. Thus, traditional augmentation methods that perform simple image alterations are less effective [102]. GAN refers to the specialised algorithms and cavernous learning systems towards the compelling predictions and transformation of data from one to another that produce dynamic data and images so that better recognition and analysis can be done. GAN-based DL models are applied to generate data artificially. Therefore, to overcome the data-scarce situation, GAN is used to develop effective data augmentation strategies for medical visual recognition.

Based on the journal written by Afshar et al., CNNs that were applied to identify positive COVID-19 CXR images are prone to lose spatial information between image instances and require a large dataset to compensate for the loss. Capsule networks, a.k.a COVID-CAPS, is an alternative modelling framework capable of handling small datasets.

Table 5: Challenges of radiology imaging addresses and AI applications.

No	Applications	Type of data	Challenges	AI methods	Sources
1	COVID-19 early detection using radiology images. Typically CXR and CT images	CXR images	Limited availability of annotated medical images and medical image classification remains the biggest challenge in medical diagnosis.	DeTraC deep convolution neural network	[14]
2		CXR images	Finding optimal parameters for the SVM classifier can be seen as a challenge. Finding optimal parameters for the SVM classifier can be seen as a challenge. Finding optimal values for the relief algorithm can be seen as another limitation of the study	COVIDetectioNet	[92]
3		CT images	Redundant data such as interferential vessels can be misdiagnosed as pathology. Radiologists have difficulty differentiating COVID-19 and other atypical and viral pneumonia diseases, which are the same in CT imagery and have similar symptoms.	AlexNet, VGG-16, VGG-19, SqueezeNet, GoogleNet, MobileNet-V2, ResNet-18, ResNet-50, ResNet-101, Xception	[93]
4		CXR images	Due to the sudden existence and infectious nature of COVID-19, systematic collection of the extensive data set for CNN training is formidable. Biomarkers found in the CXR images can be misleading.	Patch-based convolutional network	[68]
5		CXR images	The research is dealing with images taken directly from patients with severe COVID-19 or some form of pneumonia. However, in the real world, more people are unaffected by pneumonia. The limited number of data available provides a limitation to provide feasible results.	Multiclass classification and hierarchical classification, using texture descriptors and also pretrained CNN model	[94]
6		CXR images	Insufficient pulmonary diseases data limit us to conduct verification techniques.	Localise the areas in CXR symptomatic of the COVID-19 presence	[95]
7		CT images	Shortage of radiology image labelled "data"	Segmentation deep network (Inf-net)	[96]

Capsule network consists of capsules in the convolutional layers. It has the potential to improve further diagnosis capabilities. Hence, using capsule network to pretrain the images is expected to improve the accuracy where each capsule in the convolutional layers represents a specific image instance at a specific location through several neurons. The routing agreement in the capsule network helps CNN models to identify spatial relations.

5. Conclusion and Future Works

COVID-19 has disrupted the lives of people worldwide. The number of casualties related to the disease cannot be contained and has increased by the thousands daily. AI technologies have existed to help us live comfortably and have many successes and contributions in streamlining processes and procedures. However, the spread of COVID-19 is exceptionally lethal as it transmits faster and broader than ever. The coronavirus is also continuously revolutionising with new spikes, and protein mutations have been reported in countries like Malaysia, United Kingdom, South America, Australia, the Netherlands, and Singapore. The clinical impact of this discovery and its infectivity or aggressiveness is still unknown. Whether or not the mutation will affect the development of radiography imaging is also still a mystery.

Based on the worldmeter website: https://www.worldometers.info/coronavirus/, some countries failed to respond to the disease, some are barely tackling the situation, and some are handling the situation much more successfully. Hence, a country that managed to have the situation under control might experience a spike increase overnight if society became lenient in taking proper measures.

Although many researchers have published their works, the number of contributions and AI applications towards tackling COVID-19 is rudimentary. With the petrifying number of deaths and infected patients discovered daily and the virus's mutation undergoing speedily and unknowingly, we are nowhere near applying AI on radiography imaging to identify that the patient is infected with SARS-CoV-2. The development of AI and radiography imaging is slow due to the limited availability of COVID-19 datasets. With the number of people affected worldwide, AI methods require massive data and several computational models and CNN architectures to learn and acquire knowledge. The

current data that most researchers acquired from open-source websites is insufficient. Even with the best available data, it is far from perfect; as the data alone cannot explain the pandemic's whole situation. Therefore, for future research and development, in terms of acquiring radiography imaging data, the best way is to have access to reliable, global, open data, and research to build an infrastructure that allows researchers who are experts in the field of radiology, artificial intelligence, deep learning, and imaging to navigate and understand this data and its development.

Most of the COVID-19 radiography image datasets are stored in different formats, standards, sizes, and quality, which are obstacles for scientists to speed up development for COVID-19-related AI research. Therefore, in future development, COVID-19 radiography images should have standard operating procedures to allow researchers or scientists and anyone who has the passion and are interested to contribute and utilise the information freely. A future study on deep learning models identifying and distinguishing the difference between COVID-19 images and viral pneumonia is essential. The study would help radiologists and physicians understand the virus and evaluate future coronaviruses using CT and CXR images more efficiently and effectively.

Data Availability

Data analyzed in this study were a reanalysis of existing data, which are openly available at locations cited in the reference section.

Conflicts of Interest

All the authors declared that they have no conflict of interest.

Acknowledgments

This work is supported by RU Geran University of Malaya (ST014-2019).

References

[1] American Library Association, "Worldmeter," August 2020, https://www.worldometers.info/coronavirus/.

[2] Centers for Disease Control and Prevention, "Coronavirus disease 2019 (COVID-19)," August 2020, https://www.cdc.gov/coronavirus/2019-ncov/cases-in-us.html.

[3] WebMD, *Coronavirus and pneumonia*, WebMD LLC, 2020, August 2020, https://www.webmd.com/lung/covid-and-pneumonia#1.

[4] Healthline, "How is COVID-19 pneumonia different from regular pneumonia?," August 2020, https://www.healthline.com/health/coronavirus-pneumonia#vs-regular-pneumonia.

[5] T. Rahman, M. E. H. Chowdhury, A. Khandakar et al., "Transfer learning with deep convolutional neural network (CNN) for pneumonia detection using chest X-ray," *Applied Sciences*, vol. 10, no. 9, p. 3233, 2020.

[6] V. Chouhan, S. K. Singh, A. Khamparia et al., "A novel transfer learning based approach for pneumonia detection in chest X-ray images," *Applied Sciences*, vol. 10, no. 2, p. 559, 2020.

[7] Occupational Safety and Health Administration, *COVID-19*, United States Department of Labor, United States, 2020.

[8] Centers for Disease Control and Prevention, "Centers for disease control and prevention," December 2020, https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/how-covid-spreads.html.

[9] S. Rajaraman and S. Antani, "Weakly labeled data augmentation for deep learning: a study on COVID-19 detection in chest X-rays," *Diagnostics*, vol. 10, no. 6, p. 358, 2020.

[10] BBC News, *Covid: FDA Approves Pfizer Vaccine for Emergency Use in US*, BBC, 2020, https://www.bbc.com/news/world-us-canada-55265477.

[11] New Straits Times, *New Straits Times*, New Straits Times, 2020, December 2020, https://www.nst.com.my/world/world/2020/12/647620/who-against-mandatory-covid-19-vaccines.

[12] G. Rubin, C. Ryerson, L. Haramati et al., "The role of chest imaging in patient management during the COVID-19 pandemic: a multinational consensus statement from the Fleischner Society," *Radiology*, vol. 296, no. 1, pp. 172–180, 2020.

[13] H. X. Bai, B. Hsieh, Z. Xiong et al., "Performance of radiologists in differentiating COVID-19 from non-COVID-19 viral pneumonia at chest CT," *Radiology*, vol. 296, no. 2, pp. E46–E54, 2020.

[14] A. Abbas, M. M. Abdelsamea, and M. M. Gaber, "Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network," *Applied Intelligence*, vol. 51, no. 2, pp. 854–864, 2021.

[15] D. Kermany, M. Goldbaum, W. Cai et al., "Identifying medical diagnoses and treatable diseases by image-based deep learning," *Cell*, vol. 172, no. 5, pp. 1122–1131.e9, 2018.

[16] ACR, "ACR recommendations for the use of chest radiography and computed tomography (CT) for suspected COVID-19 infection," *American College of Radiology*, 2020, December 2020, https://www.acr.org/Advocacy-and-Economics/ACR-Position-Statements/Recommendations-for-Chest-Radiography-and-CT-for-Suspected-COVID19-Infection.

[17] M. Loey, F. Smarandache, and N. E. M. Khalifa, "Within the lack of chest COVID-19 X-ray dataset: a novel detection model based on GAN and deep transfer learning," *Symmetry*, vol. 12, no. 4, p. 651, 2020.

[18] D. Rong, L. Xie, and Y. Ying, "Computer vision detection of foreign objects in walnuts using deep learning," *Computer Electronic Ariculture*, vol. 162, pp. 1001–1010, 2019.

[19] G. Eraslan, Ž. Avsec, J. Gagneur, and F. Theis, "Deep learning: new computational modelling techniques for genomics," *Nature Reviews Genetics*, vol. 37, pp. 389–403, 2019.

[20] J. Riordon, D. Sovilj, S. Sanner, D. Sinton, and E. Young, "Deep learning with microfluidics for biotechnology," *Trends in Biotechnology*, vol. 37, no. 3, pp. 310–324, 2019.

[21] A. Lundervold and A. Lundervold, "An overview of deep learning in medical imaging focusing on MRI," *Zeitschrift für Medizinische Physik*, vol. 29, no. 2, pp. 102–127, 2019.

[22] A. Maier, C. Syben, T. Lasser, and C. Riess, "A gentle introduction to deep learning in medical image processing," *Zeitschrift für Medizinische Physik*, vol. 29, no. 2, pp. 86–101, 2019.

[23] A. Shrestha and A. Mahmood, "Review of deep learning algorithms and architectures," *IEEE*, vol. 7, pp. 53040–53065, 2019.

[24] S. Pouyanfar, S. Sadiq, Y. Yan et al., "Multimedia big data analytics," *ACM Computing Surveys*, vol. 51, no. 1, pp. 1–34, 2018.

[25] R. Herman, "Mindmeister," May 2021, https://www.mindmeister.com/969402186/machine-learning-algorithms?fullscreen=1.

[26] Ao Visualization, Director, *Capsule Networks - Introduction*, Art of Visualization, Toronto, Canada, 2018.

[27] J. Al-Tawfiq, A. Zumla, and Z. Memish, "Coronaviruses," *Current Opinion in Infectious Diseases*, vol. 27, no. 5, pp. 411–417, 2014.

[28] L. Ketai, N. Paul, and K. Wong, "Radiology of severe acute respiratory syndrome (SARS)," *Journal of Thoracic Imaging*, vol. 21, no. 4, pp. 276–283, 2006.

[29] M. Chung, A. Bernheim, X. Mei et al., "CT imaging features of 2019 novel coronavirus (2019-nCoV)," *Radiology*, vol. 295, no. 1, pp. 202–207, 2020.

[30] J. Chan, S. Yuan, K. H. Kok et al., "A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster," *The Lancet*, vol. 395, no. 10223, pp. 514–523, 2020.

[31] P. Liu and X. Tan, "2019 novel coronavirus (2019-nCoV) pneumonia," *Radiology*, vol. 295, no. 1, p. 19, 2020.

[32] M. Hosseiny, S. Kooraki, A. Gholamrezanezhad, S. Reddy, and L. Myers, "Radiology perspective of coronavirus disease 2019 (COVID-19): lessons from severe acute respiratory syndrome and Middle East respiratory syndrome," *American Journal of Roentgenology*, vol. 214, no. 5, pp. 1078–1082, 2020.

[33] C. Czawlytko, R. Hossain, and C. S. White, "COVID-19 diagnostic imaging recommendations," *Applied Radiology*, vol. 49, no. 3, pp. 10–15, 2020.

[34] M. E. H. Chowdhury, T. Rahman, A. Khandakar et al., "Can AI help in screening viral and COVID-19 pneumonia?," *IEEE Access*, vol. 8, pp. 132665–132676, 2020.

[35] W. Wang, "Detection of SARS-CoV-2 in different types of clinical specimens," *JAMA*, vol. 323, no. 18, pp. 1843-1844, 2020.

[36] S. Sharma, "Drawing insights from COVID-19-infected patients using CT scan images and machine learning techniques: a study on 200 patients," *Environmental Science and Pollution Research*, vol. 27, no. 29, pp. 37155–37163, 2020.

[37] C. Joseph Paul, M. Paul, and L. Dao, "Covid-19 image data collection," *Image and Video Processing*, vol. eessIV, pp. 1–4, 2020.

[38] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, and R. M. Summers, "ChestX-Ray8: hospital-scale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 369–373, Honolulu, HI, USA, 2017.

[39] J. P. Cohen, P. Bertin, and V. Frappier, "Chester: a web delivered locally computed chest X-ray disease prediction system," 2019, https://arxiv.org/abs/1901.11210.

[40] P. Rajpurkar, A. Y. Hannun, M. Haghpanahi, C. Bourn, and A. Y. Ng, "Cardiologist-level arrhythmia detection with convolutional neural networks," 2017, https://arxiv.org/abs/1707.01836.

[41] M. Satyanarayana, "Shortage of RNA extraction kits hampers efforts to ramp up COVID-19 coronavirus testing," vol. 3, 2020.

[42] M. Kelly Geraldine, *Interviewee, Testing Backlog Linked to Shortage of Chemicals Needed for COVID-19 Test*, The Canadian Press, 2020.

[43] M. Yascha, *Coronavirus: Extraordinary Decisions For Italian Doctors-The Atlantic*, The Atlantic, 2020.

[44] T. T. Nguyen, "Artificial intelligence in the battle against coronavirus (COVID-19): a survey and future research directions," 2020, https://arxiv.org/abs/2008.07343.

[45] L. Lin, Q. Lixin, X. Zeguo et al., "Artificial intelligence distinguishes COVID-19 from community acquired pneumonia on chest CT," *Radiology*, vol. 296, no. 2, Article 200905, 2020.

[46] B. Charmaine, G. Jagpal, C. David, and A. B. Benson, "Deep learning system to screen coronavirus disease 2019 pneumonia," *Applied Intelligence*, pp. 1–7, 2020.

[47] G. Biraja and T. Allan, "Estimating uncertainty and interpretability in deep learning for coronavirus (COVID-19) detection," 2020, https://arxiv.org/abs/2003.10769.

[48] W. Shuai, K. Bo, M. Jinlu et al., "A deep learning algorithm using CT images to screen for corona virus disease (COVID-19)," *European Radiology*, vol. 31, no. 6, 2021.

[49] B. Xiang, F. Cong, Z. Yu et al., "Predicting COVID-19 malignant progression with AI techniques," *SSRN Electronic Journal*, 2020.

[50] J. Cheng, C. Weixiang, C. Yukun et al., "Development and evaluation of an AI system for COVID-19 diagnosis," *Nature Communications*, vol. 11, article 5088, 2020.

[51] J. Shuo, W. Bo, X. Haibo et al., "AI-assisted CT imaging analysis for COVID-19 screening: building and deploying a medical AI system in four weeks," *Applied Soft Computing*, vol. 98, 2020.

[52] N. Ali, K. Ceren, and P. Ziynet, "Automatic detection of coronavirus disease (COVID-19) using X-ray images and deep convolutional neural networks," *Pattern Analysis and Applications*, vol. 24, no. 2, 2021.

[53] L. Wang, Z. Q. Lin, and A. Wong, "COVID-net: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images," *Nature*, vol. 10, no. 1, article 19549, 2020.

[54] O. Gozes, M. Frid-Adar, H. Greenspan et al., "Rapid AI development cycle for the coronavirus (COVID-19) pandemic: initial results for automated detection & patient monitoring using deep learning CT image analysis," 2020, https://arxiv.org/abs/2003.05037.

[55] H. Maghdid, A. T. Asaad, K. Z. G. Ghafoor, A. S. Sadiq, S. Mirjalili, and M. K. K. Khan, "Diagnosing COVID-19 pneumonia from X-ray and CT images using deep learning and transfer learning algorithms," in *Multimodal Image Exploitation and Learning 2021*, 2021.

[56] Y. Xingyi, H. Xuehai, Z. Jinyu, Z. Yichen, Z. Shanghang, and X. Pengtao, "COVID-CT-dataset: a CT scan dataset about COVID-19," 2020, https://arxiv.org/abs/2003.13865.

[57] A. Degerli, M. Ahishali, M. Yamac et al., "COVID-19 infection map generation and detection from chest X-ray images," *Health Information Science and Systems*, vol. 9, no. 1, p. 15, 2021.

[58] A. Omar Ibrahim, "Automatic deep learning system for COVID-19 infection quantification in chest CT," 2020, https://arxiv.org/abs/2010.01982.

[59] A. Ridhi, B. Vipul, B. Himanshu et al., "AI-based diagnosis of COVID-19 patients using X-ray scans with stochastic ensemble of CNNs," 2020.

[60] W. Shuo, Z. Yunfei, L. Weimin et al., "A fully automatic deep learning system for COVID-19 diagnostic and prognostic analysis," *European Respiratory Journal*, vol. 56, no. 2, Article 2000775, 2020.

[61] Y. Song, Z. Shuangjia, L. Liang et al., "Deep learning enables accurate diagnosis of novel coronavirus (COVID-19) with CT images," *Computational Biology and Bioinformatics*, vol. 18, no. 2, p. 1, 2020.

[62] N. E. M. Khalifa, M. H. N. Taha, A. E. Hassanien, and S. Elghamrawy, "Detection of coronavirus (COVID-19) associated pneumonia based on generative adversarial networks and a fine-tuned deep transfer learning model using chest X-ray dataset," 2020, https://arxiv.org/abs/2004.01184.

[63] I. D. Apostolopoulos and T. A. Mpesiana, "Covid-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks," *Physical and Engineering Sciences in Medicine*, vol. 43, no. 2, pp. 635–640, 2020.

[64] M. Sampa, J. Seungwan, L. Seiyon, M. Ravi, and K. Chulhong, "Multi-channel transfer learning of chest X-ray images for screening of COVID-19," 2020, https://arxiv.org/abs/2005.05576.

[65] S. S. Khan, F. Khoshbakhtian, and A. B. Ashraf, "Anomaly detection approach to identify early cases in a pandemic using chest X-rays," https://arxiv.org/abs/2010.02814.

[66] M. Nour, Z. Cömert, and K. Polat, "A novel medical diagnosis model for COVID-19 infection detection based on deep features and Bayesian optimization," *Applied Soft Computing*, vol. 97, article 106580, 2020.

[67] M. F. Aslan, M. F. Unlersen, K. Sabanci, and A. Durdu, "CNN-based transfer learning-BiLSTM network: a novel approach for COVID-19 infection detection," *Applied Soft Computing*, vol. 98, p. 106912, 2021.

[68] Y. Oh, S. Park, and J. C. Ye, "Deep learning COVID-19 features on CXR using limited training data sets," *Transactions on Medical Imaging*, vol. 39, no. 8, pp. 2688–2700, 2020.

[69] L. O. Hall, R. Paul, D. B. Goldgof, and G. M. Goldgof, "Finding COVID-19 from chest X-rays using deep learning on a small dataset," 2020, https://arxiv.org/abs/2004.02060.

[70] A. Jaiswal, N. Gianchandani, D. Singh, V. Kumar, and M. Kaur, "Classification of the COVID-19 infected patients using DenseNet201 based deep transfer learning," *Journal of Interdisciplinary Mathematics*, vol. 24, no. 2, pp. 381–409, 2020.

[71] N. Narayan Das, N. Kumar, M. Kaur, V. Kumar, and D. Singh, "Automated deep transfer learning-based approach for detection of COVID-19 infection in chest X-rays," *IRBM*, vol. 42, no. 3, 2020.

[72] T. Ozturk, M. Talo, E. A. Yildirim, U. B. Baloglu, O. Yildirim, and U. Rajendra Acharya, "Automated detection of COVID-19 cases using deep neural networks with X-ray images," *Computers in Biology and Medicine*, vol. 121, article 103792, 2020.

[73] A. Gupta, S. G. Anjum, and R. Katarya, "InstaCovNet-19: a deep learning classification model for the detection of COVID-19 patients using chest X-ray," *Applied Soft Computing Journal*, vol. 99, article 106859, 2021.

[74] M. Antonios, K. Ioannis, and T. Konstantinos, "COVID-19 detection from chest X-Ray images using deep learning and convolutional neural networks," in *11th Hellenic Conference on Artificial Intelligence*, 2020.

[75] P. Afshar, S. Heidarian, F. Naderkhani, A. Oikonomou, K. N. Plataniotis, and A. Mohammadi, "COVID-CAPS: a capsule network-based framework for identification of COVID-19 cases from X-ray images," *Pattern Recognition Letters*, vol. 138, pp. 638–643, 2020.

[76] S. Minaee, R. Kafieh, M. Sonka, S. Yazdani, and G. Jamalipour Soufi, "Deep-COVID: predicting COVID-19 from chest X-ray images using deep transfer learning," *Medical Image Analysis*, vol. 65, article 101794, 2020.

[77] B. Houssam, H. Hana, and A. Aouatif, "Deep learning transfer with apache spark to detect COVID-19 in chest X-ray images," *Romanian Journal of Information Science and Technology*, vol. 23, pp. S117–S129, 2020.

[78] M. Z. Islam, M. M. Islam, and A. Asraf, "A combined deep CNN-LSTM network for the detection of novel coronavirus (COVID-19) using X-ray images," *Informatics in Medicine Unlocked*, vol. 20, p. 100412, 2020.

[79] Y. Kun, L. Xinfeng, Y. Yingli et al., "End-to-end COVID-19 screening with 3D deep learning on chest computed tomography," *Research Square*, pp. 1–14, 2020.

[80] M. A. Zulkifley, S. R. Abdani, and N. H. Zulkifley, "COVID-19 screening using a lightweight convolutional neural network with generative adversarial network data augmentation," *Computer and Engineering Science and Symmetry*, vol. 12, no. 9, p. 1530, 2020.

[81] M. Heidari, S. Mirniaharikandehei, A. Z. Khuzani, G. Danala, Y. Qiu, and B. Zheng, "Improving the performance of CNN to predict the likelihood of COVID-19 using chest X-ray images with preprocessing algorithms104284," *International Journal of Medical Informatics*, vol. 144, 2020.

[82] P. Gifani, A. Shalbaf, and M. Vafaeezadeh, "Automated detection of COVID-19 using ensemble of transfer learning with deep convolutional neural network based on CT scans," *International Journal of Computer Assisted Radiology and Surgery*, vol. 16, no. 1, pp. 115–123, 2021.

[83] G. Chirag, K. Abhimanyu, D. Satish Kumar, and S. Vishal, "Efficient deep network architecture for COVID-19 detection using computed tomography images," 2020.

[84] E. E.-D. Hemdan, M. A. Shouman, and M. E. Karar, "COVIDX-Net: a framework of deep learning classifiers to diagnose COVID-19 in X-ray images," 2020, https://arxiv.org/abs/2003.11055.

[85] M. Abed Mohammed, K. Hameed Abdulkareem, B. Garcia-Zapirain et al., "A comprehensive investigation of machine learning feature extraction and classification methods for automated diagnosis of COVID-19 based on X-ray images," *Computers, Materials & Continua*, vol. 66, no. 3, pp. 3289–3310, 2021.

[86] A. S. al-Waisy, M. Abed Mohammed, S. al-Fahdawi et al., "COVID-DeepNet: hybrid multimodal deep learning system for improving COVID-19 pneumonia detection in chest X-ray images," *Computers, Materials & Continua*, vol. 67, no. 2, pp. 2409–2429, 2021.

[87] M. A. Mohammed, K. H. Abdulkareem, A. S. al-Waisy et al., "Benchmarking methodology for selection of optimal COVID-19 diagnostic model based on entropy and TOPSIS methods," *IEEE Access*, vol. 8, pp. 99115–99131, 2020.

[88] A. S. Al-Waisy, S. Al-Fahdawi, M. A. Mohammed et al., "COVID-CheXNet: hybrid deep learning framework for identifying COVID-19 virus in chest X-rays images," *Soft Computing*, vol. 24, no. 24, 2020.

[89] S. Christian, I. Sergey, V. Vincent, and A. Alex, "Inception-v4 inception-ResNet and the impact of residual connections on learning," 2016.

[90] M. Lin, Q. Chen, and S. Yan, "Network in network," 2014, https://arxiv.org/abs/1312.4400.

[91] S. Karen and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," https://arxiv.org/abs/1409.1556.

[92] M. Turkoglu, "COVIDetectioNet: COVID-19 diagnosis system based on X-ray images using features selected from pre-learned deep features ensemble," *Applied Intelligence*, vol. 51, no. 3, pp. 1213–1226, 2021.

[93] A. A. Ardakani, A. R. Kanafi, U. R. Acharya, N. Khadem, and A. Mohammadi, "Application of deep learning technique to manage COVID-19 in routine clinical practice using CT images: results of 10 convolutional neural networks," *Computers in Biology and Medicine*, vol. 121, article 103795, 2020.

[94] R. M. Pereira, D. Bertolini, L. O. Teixeira, C. N. Silla Jr., and Y. M. G. Costa, "COVID-19 identification in chest X-ray images on flat and hierarchical classification scenarios," *Computer Methods and Programs in Biomedicine*, vol. 194, article 105532, 2020.

[95] L. Brunese, F. Mercaldo, A. Reginelli, and A. Santone, "Explainable deep learning for pulmonary disease and coronavirus COVID-19 detection from X-rays," *Computer Methods and Programs in Biomedicine*, vol. 196, p. 105608, 2020.

[96] D. P. Fan, T. Zhou, G. P. Ji et al., "Inf-Net: automatic COVID-19 lung infection segmentation from CT images," *IEEE Transactions on Medical Imaging*, vol. 39, no. 8, pp. 2626–2637, 2020.

[97] Radiological Society of North America (RSNA), *RSNA Announces COVID-19 Imaging Data Repository*, Radiological Society of North America (RSNA), 2020, March 2021, https://press.rsna.org/timssnet/media/pressreleases/14_pr_target.cfm?ID=2167.

[98] Imaging COVID-19 AI initiative, *A European Initiative for Automated Diagnosis and Quantitative Analysis of COVID-19 on Imaging*, Imaging COVID-19 AI initiative, 2020, March 2021, https://imagingcovid19ai.eu/.

[99] W. Jason and P. Luis, "The effectiveness of data augmentation in image classification using deep learning," 2017, https://arxiv.org/abs/1712.04621.

[100] G. Prasanth, R. Sivaramakrishnan, and L. Rodney, "Assessment of data augmentation strategies toward performance improvement of abnormality classification in chest radiographs," in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, p. 4, Berlin, Germany, 2019.

[101] D. Jia, D. Wei, S. Richard, L. Li-Jia, and L. L. F.-F. Kai, "ImageNet: a large-scale hierarchical image database," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, p. 8, Miami, FL, USA, 2009.

[102] B.-C. Avi, K. Eyal, M. A. Michal, G. Jacob, and G. Hayit, "Anatomical data augmentation for CNN based pixel-wise classification," in *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, pp. 4–7, Washington, DC, USA, 2018.

Hindawi
Computational and Mathematical Methods in Medicine
Volume 2021, Article ID 6676987, 6 pages
https://doi.org/10.1155/2021/6676987

Research Article

Clinical Characteristics and Early Interventional Responses in Patients with Severe COVID-19 Pneumonia

Susu He,[1] Lina Fang,[2] Lingzhen Xia,[3] Shuangxiang Lin,[4] Junhui Ye,[5] and Dinghai Luo[6]

[1]*Department of Respiratory Medicine, Taizhou Hospital of Zhejiang Province Affiliated to Wenzhou Medical University, Linhai, Zhejiang, China*
[2]*Endoscopy Center, Taizhou Hospital of Zhejiang Province Affiliated to Wenzhou Medical University, Linhai, Taizhou, China*
[3]*General Practice, Taizhou Central Hospital, Taizhou, Zhejiang, China*
[4]*Department of Radiology, Taizhou Hospital of Zhejiang Province Affiliated to Wenzhou Medical University, Linhai, Zhejiang, China*
[5]*Department of Respiratory Medicine, The People's Hospital of Sanmen, Taizhou, Zhejiang, China*
[6]*Department of Gastroenterology, Taizhou Hospital of Zhejiang Province Affiliated to Wenzhou Medical University, Linhai, Taizhou, China*

Correspondence should be addressed to Junhui Ye; smyejh@hotmail.com and Dinghai Luo; luodh@enzemed.com

Received 19 December 2020; Revised 13 January 2021; Accepted 11 April 2021; Published 2 June 2021

Academic Editor: Kaijian Xia

Progressive acute respiratory distress syndrome (ARDS) is the most lethal cause in patients with severe COVID-19 pneumonia due to uncontrolled inflammatory reaction, for which we found that early intervention of combined treatment with methylprednisolone and human immunoglobulin is a highly effective therapy to improve the prognosis of COVID-19-induced pneumonia patients. *Objective*. Herein, we have demonstrated the clinical manifestations, laboratory, and radiological characteristics of patients with severe Coronavirus Disease-2019 (COVID-19) pneumonia, as well as measures to ensure early diagnosis and intervention for improving clinical outcomes of COVID-19 patients. *Summary Background Data*. The COVID-19 is a new infection caused by a severe acute respiratory syndrome- (SARS-) like coronavirus that emerged in China in December 2019 and has claimed millions of lives. *Methods*. We included 37 severe COVID-19 pneumonia patients who were hospitalized at Taizhou Public Health Medical Center in Zhejiang province from January 17, 2020, to February 18, 2020. Demographic, clinical, and laboratory features; imaging characteristics; treatment history; and clinical outcomes of all patients were collected from electronic medical records. *Results*. The patients' mean age was 54 years (interquartile range, 43–64), with a slightly higher male preponderance (57%). The most common clinical features of COVID-19 pneumonia were fever (29 (78%)), dry cough (28 (76%)), dyspnea (9 (24%)), and fatigue (9 (24%)). Serum interleukin (IL)-6 and IL-10 were elevated in 35 (95%) and 19 (51%) patients, respectively. Chest computerized tomography scan revealed bilateral pneumonia in 35 (95%) patients. Early intervention with a combination of methylprednisolone and human immunoglobulin was highly effective in improving the prognosis of these patients. *Conclusions*. Progressive acute respiratory distress syndrome is the most common cause of death in patients with severe COVID-19 pneumonia owing to an uncontrolled inflammatory response. Early intervention with methylprednisolone and human immunoglobulin was highly effective in improving their prognosis.

1. Introduction

In December 2019, a novel coronavirus was identified in patients with viral pneumonia in Wuhan, which was later named Coronavirus Disease-2019 (COVID-19) by the World Health Organization (WHO) on January 11, 2020 [1, 2]. Given the possibility of airborne transmission in humans, the virus has infected more than 100 million people and spread worldwide into hundreds of countries [3–5]. Most coronaviruses usually cause mild illness, but there are two β-coronaviruses, severe acute respiratory syndrome (SARS)-coronavirus and Middle-East respiratory syndrome (MERS)-coronavirus, which can lead to severe acute respiratory syndrome with a mortality rate of 10% and 37%,

respectively [6, 7]. Recent studies have revealed that COVID-19 causes severe pneumonia, respiratory distress syndrome, and multiple organ failure, leading to a high mortality rate [8–10].

This study is aimed at describing the clinical, laboratory, and radiological characteristics of critically ill patients with COVID-19, as well as the effective outcomes after early detection and intervention. These findings could provide a useful medicinal strategy for the treatment of severely ill patients with COVID-19 pneumonia.

2. Materials and Methods

2.1. Ethics. This study was approved by the ethics committee of the Enze Hospital of the Zhejiang Enze Medical Center (group). Data collected from enrolled cases were shared with the WHO. Written informed consent was obtained from all patients before data collection.

2.2. Clinical Records and Data Collection. Thirty-seven adult patients with severe COVID-19 pneumonia, diagnosed following the WHO interim guidelines [1, 11], were hospitalized at Taizhou Public Health Medical Center in Taizhou City, Zhejiang Province from January 17, 2020, to March 11, 2020. The medical records obtained from patients were analyzed by the respiratory research team of Zhejiang Enze Medical Center (group) Taizhou Hospital and Enze Hospital. The clinical features; laboratory investigations; chest imaging characteristics; treatment history, including antiviral therapy, hormone therapy, human immunoglobulin therapy, and respiratory support; and recovery data were retrieved from electronic medical records. The source data included demographic data; symptoms; comorbidities; laboratory results; chest computerized tomography (CT) scan; and treatment measures, including laboratory results, such as complete blood count, blood clotting parameters, liver and kidney functions, serum electrolytes, erythrocyte sedimentation rate, C-reactive protein (CRP) level, procalcitonin (PCT) level, creatine kinase level, blood gas analysis, troponin level, and cytokine levels. All data were examined by two physicians.

Table 1: Baseline characteristics of patients infected with COVID-19. Data are median (IQR) or n/N (%), where N is the total number of patients with available data.

Patients ($n = 37$)	
Age, years	
Median (IQR)	54 (43-64)
Range	27-86
<40	6 (16%)
40-70	30 (81%)
>70	1 (3%)
Sex	
Male	21 (57%)
Female	16 (43%)
Current smoking	3 (8%)
Comorbidity	9 (24%)
Hypertension	6 (16%)
Diabetes	2 (5%)
Chronic obstructive pulmonary disease	2 (5%)
Hypothyroidism	2 (5%)
Signs and symptoms at admission	
Fever	29 (78%)
Cough	28 (76%)
Sputum production	6 (16%)
Dyspnea	9 (24%)
Myalgia	3 (8%)
Fatigue	8 (22%)
Diarrhea	3 (8%)
Headache	4 (11%)
Dizzy	5 (14%)
Pharyngula	3 (8%)
More than one sign or symptom	32 (86%)
Clinical outcome	
Discharged	37 (100%)
Died	0 (0%)

2.3. Laboratory Investigations. Real-time fluorescent polymerase chain reaction (RT-PCR) was used to confirm novel coronavirus infection by detecting viral nuclei from respiratory tract specimens. Diagnosing severe respiratory symptoms requires fulfilling an oxygenation index criterion (PaO_2/FIO_2 of <300 mmHg). Clinical stability was defined as an oxygenation index of ≥300 mmHg with stable clinical symptoms. The time of symptom remission was defined as the time from the diagnosis of severe illness to the time of symptom relief. The time of oxygenation improvement was the time of diagnosis of severe disease to the time of oxygenation index of ≥300 mmHg. The apparent absorption time of the internal lesion was regarded as the time from the diagnosis of severe disease till obvious absorption of the pulmonary lesion on CT. The discharge criteria were defined as normal body temperature for more >3 days, significant improvement of respiratory symptoms, pulmonary imaging showing significant inflammatory absorption, and two consecutive negative respiratory nucleic acid test results (sampling interval of at least >24 h).

2.4. Medical Treatment. Laboratory tests and chest CT were performed on the patients on admission. Next, 5 million units of aerosolized α-interferon in combination with oral lopinavir/ritonavir (500 mg) were administered twice daily to reduce viral activity. During hospitalization, vital signs and blood oxygen saturation were closely monitored. Blood gas analysis was performed promptly to determine the oxygenation index and the development of severe or new symptoms (chest tightness) were monitored. Respiratory support (high flow oxygen inhalation through a nasal catheter) was immediately administered to severely ill patients along with intravenous methylprednisolone (0.5 – 1 mg/kg daily), human immunoglobulin (0.3 – 0.5 g/kg daily), and fluids. Albendazole tablets (200 mg thrice daily) were administered to increase the efficacy of antiviral drugs. Human

Table 2: Laboratory findings of patients infected with COVID-19.

Blood routine	Patients ($n = 37$)
Leucocytes (×109 per L; normal range 3.5–9.5)	6 (4.3-7.6)
Increased	5 (14%)
Decreased	2 (5%)
Neutrophils (×109 per L; normal range 1.8–6.3)	4.5 (2.7-6.6)
Increased	9 (24%)
Lymphocytes (×109 per L; normal range 1.1–3.2)	0.8 (0.6-1.0)
Decreased	28 (76%)
Platelets (×109 per L; normal range 125.0–350.0)	200 (141-258)
Increased	1 (3%)
Decreased	7 (19%)
Coagulation function	
Activated partial thromboplastin time (normal range 23.5–36.0)	30.2 (28.1-32.7)
Increased	1 (3%)
Decreased	0 (0%)
Prothrombin time (s; normal range 12.5–14.0)	11.9 (11.3-12.5)
Increased	2 (5%)
Decreased	28 (76%)
D-dimer (μg/L; normal range 0.0–0.55)	0.32 (0.22-0.71)
Increased	9 (24%)
Blood biochemistry	
Alanine aminotransferase (U/L; normal range 9.0–50.0) ALT	22 (16.5-36)
Increased	7 (19%)
Aspartate aminotransferase (U/L; normal range 15.0–40.0) AST	28 (21.5-39.5)
Increased	9 (24%)
Serum creatinine (μmol/L; normal range 62.0–97.0)	42 (29.5-54.5)
Increased	5 (14%)
Creatine kinase (U/L; normal range 38.0–174.0)	88 (66.5-161.5)
Increased	8 (22%)
Troponin(ng/mL; normal range 0.00-0.08)	0.01 (0.01)
Increased	0 (0%)
Infection-related biomarkers	
C-reactive protein (mg/L; normal range 0.0–5.0)	19.1 (10.0-50.8)
Increased	32 (86%)
Erythrocyte sedimentation rate (mm/h; normal range 0.0–20.0)	42 (29.5-54.5)
Increased	31 (84%)
Procalcitonin (ng/mL; normal range 0.0–0.05)	0.05 (0.03-0.07)
Increased	14 (38%)
Interleukin-6 (pg/mL; normal range 0.1–2.9)	12.85 (6.3-27.7)
Increased	35 (95%)
Interleukin-10 (pg/mL; normal range 0.1–5.0)	5.1 (3.4-9.2)
Increased	19 (51%)

immunoglobulin was stopped based on the patients' condition, and the dose of methylprednisolone was gradually reduced accordingly. Empirical antibiotics were administered to patients who showed elevated white blood cell count and/or significantly elevated C-reactive protein (CRP) level. Critically ill patients were evaluated for oxygenation index at least once daily until clinical signs and symptoms were relieved.

2.5. Statistical Analyses. The categorical variables were described as frequency and percentages, and the continuous variables were described using mean, median, and interquartile range (IQR). All statistical analyses were performed using SPSS (version 20.0) software (IBM, Armonk, NY, USA).

3. Results

The study population consisted of 37 critically ill inpatients diagnosed with COVID-19 pneumonia. The mean age of the patients with severe COVID-19 pneumonia was 54 years (IQR, 43–64), with a slightly higher male preponderance (57%). Nine (24%) patients had one or more comorbidities. The most common clinical features of COVID-19 pneumonia were fever (29 (78%)), dry cough (28 (76%)), dyspnea (9 (24%)), and fatigue (9 (24%)). As of March 11, 2020, all patients were discharged (Table 1).

At admission, patients had absolute lymphocytosis (28 (76%)) and thrombocytopenia (7 (19%)), but leukopenia was rare (2 (5%)). Among the coagulation parameters, only one (3%) patient showed a slightly increased activated partial prothrombin time, two (5%) showed a decreased prothrombin time, 28 (76%) showed a decreased D-dimer level, and a few (9 (24%)) showed an increased D-dimer level (IQR, 0.22–0.71 μg/mL). Liver function was impaired to varying degrees in nine (24%) patients with elevated alanine aminotransferase/aspartate aminotransferase levels, and renal function was impaired in five (14%) patients with elevated creatinine levels. A slightly increased creatine kinase level was found in eight (22%) patients. The troponin levels were in the normal range in all patients.

A vast majority of patients (32 [86%]) showed an increased CRP level (IQR, 10.0–50.8 mg/L). Thirty-one (84%) patients had an increased erythrocyte sedimentation rate (IQR, 29.5–54.5 mm/h). Serum PCT levels were partially increased in 14 (38%) patients, although it was not significant (IQR, 0.03–0.07 ng/mL). Serum interleukin (IL)-6 and IL-10

Table 3: Radiographic findings of patients infected with COVID-19.

Patients ($n = 37$)	
Ground-glass opacity	27 (73%)
Patch shadow	37 (100%)
Interstitial abnormalities	25 (68%)
Consolidation	22 (59%)
Nodule	10 (27%)
Normal	0 (0%)
Local patchy shadowing	2 (5%)
Bilateral patchy shadowing	35 (95%)

Data are n/N (%), where N is the total number of patients with available data.

Table 4: Treatment of patients infected with COVID-19.

	n/N (%)	Median (IQR)
No. of patients	37	37
Oxygen therapy		
Nasal cannula	32 (86%)	NA
High-flow nasal cannula	5 (14%)	NA
Antiviral treatment		NA
α-Interferon	37 (100%)	NA
Lopinavir/litonavir	37 (100%)	NA
Arbidol	23 (62%)	NA
Glucocorticoids	30 (81%)	7 (4-11)
Intravenous immunoglobulin therapy	25 (68%)	4 (3-5.5)
Antibiotic treatment	9 (24%)	7 (4.5-7.5)

Data are median (IQR) or n/N (%), where N is the total number of patients with available data.

levels were elevated in 35 (95%) and 19 (51%) patients, respectively (Table 2).

Abnormal CT findings were detected in all patients. Only two (5%) patients had unilateral lobar lesions, while the remaining 35 (95%) had bilateral lobar lesions. Chest CT in most patients showed multiple manifestations, the most common being a patch shadow, which was found in all patients. The other findings were ground glass shadows (27 (73%)), interstitial changes (25 (68%)), consolidation shadows (22 (59%)), and nodular shadows without pleural effusion (10 (27%)) (Table 3).

All patients received oxygen; 32 (86%) and five (14%) patients received high-flow oxygen through a nasal catheter without invasive ventilation. All patients were treated with 5 million units of aerosolized α-interferon and lopinavir/ritonavir therapy (500 mg, twice daily). Albendazole tablets (200 mg thrice daily) were administered to 23 (62%) patients. Based on the white blood cell count and CRP level, nine (24%) patients were treated with antibiotics, including penicillin, macrocyclic lipids, and quinolones. The duration of antibiotic treatment was 2–11 days (median, 7 days; IQR, 4.5–7.4). Thirty (81%) patients received intravenous methylprednisolone intravenous infusion (0.5–1 mg/kg daily) for was 3–18 days (median, 7 days; IQR, 4–11 days), depending on the clinical condition. Twenty-five (68%) patients were treated with human immunoglobulin (0.3–0.5 g/kg daily), and 24 (65%) of them were treated for 3–11 days (median, 4 days; IQR, 3–5.5) (Table 4).

The median duration from symptom onset to diagnosis of severe illness was 7.5 days (IQR, 4.5–7.5), the median duration for symptom remission was 5 days (IQR, 2.5–11), and the median duration for saturation improvement was 5 days (IQR, 3–9). Significant absorption of pulmonary lesions was observed in 28 (76%) patients within a median duration of 9 days (IQR, 6–11) (Table 5).

4. Discussion

Among the 37 laboratory-confirmed severe COVID-19 patients, nine (24%) had underlying diseases, mainly hypertension, but a significant number of previously healthy patients had worsened clinical status during treatment. The main symptoms of patients with severe infection are fever, dry cough, chest tightness, and weakness. A few patients may have muscle pain, pharyngeal pain, dizziness, headache, loss of appetite, etc. COVID-19 patients rarely have gastrointestinal symptoms, such as diarrhea, which is different from the symptoms of SARS-coronavirus and MERS-coronavirus infections [11, 12].

Similar to recent reports [13–15], a reduced absolute lymphocyte count was found in 76% of critically ill patients. Reduced IL-6 and IL-10 levels were found in 95% and 19% of patients, respectively, suggesting that COVID-19 may have a role in cellular immune deficiency. The virus spreads through the mucous membrane along the respiratory tract, which provokes a series of immune reactions leading to cytokine storm, eventually causing a change in the peripheral blood lymphocyte count. IL-6 may play a proinflammatory role in pulmonary inflammation, and a significant increase in IL-6 level was found in the majority of patients, suggesting that severe COVID-19 may cause a significant exudation of IL-6 into the lungs. A decrease in the absolute lymphocyte count and increased IL-6 and IL-10 levels may indicate disease worsening.

In our study, PCT level, D-dimer level, clotting parameters, and troponin level were not much altered or slightly elevated, unlike in previous reports involving patients infected with SARS-CoV and MERS-CoV [16]. We speculate that our patients were not critically ill enough to require intensive care and did not develop coagulation disorders, liver and kidney dysfunctions, and myocardial injuries. Furthermore, it also suggests that PCT level, coagulation parameters, troponin level, and other common indicators cannot predict disease worsening.

Chest CT of patients with severe COVID-19 mainly shows bilateral patchy and frosted glass shadows. In our study, the involvement of multiple lobes was found in most patients, some of them showed consolidations, interstitial changes, and nodule shadows. No patient showed pleural effusion. The oxygenation index is a sensitive indicator of the progress of COVID-19. Patients diagnosed with COVID-19 were closely monitored for vital signs immediately after admission,

Table 5: Clinical outcome of patients infected with COVID-19.

	Patients, *n*	Median (IQR), days
Time of severe	37	7.5 (4.5-7.5)
Time of symptom remission	37	5 (3-10)
The time of oxygenation improvement	37	5 (3-9)
Obvious absorption time of intrapulmonary lesion	28	9 (6-11)

Data are median (IQR). Time of severe was onset of symptoms to the diagnosis of severe, symptoms diagnosed with severe time represent the time to relieve symptoms, improve oxygenation time expressed as a diagnosis of severe oxygenation index of 300 mmHg or to time, and lung lesions significantly absorbed within the said time view demonstrate the focal area of the lungs to the CT diagnosis of severe time, one of the most obvious time to absorb the time obviously. One patient is still in a critical stage.

and oxygen saturation was measured twice daily. Blood gas analysis was also performed periodically.

We evaluated the pattern of changes in symptoms. Disease aggravation was reflected by chest tightness. To detect disease progression early, blood gas analysis should be performed immediately to determine the oxygenation index. Because some patients cannot withstand imaging due to advanced disease, and imaging findings may be detected much late (advanced disease), it is not recommended to intervene after the appearance of chest CT features [17, 18].

In patients diagnosed with COVID-19, we immediately administered α-interferon along with lopinavir/ritonavir. However, a considerable number of patients were in the critical stage, suggesting that the antiviral drugs may not be as efficacious as expected. Many patients had diarrhea and other adverse effects; hence, the administration of antiviral drugs remains to be discussed further.

For the diagnosis of severe COVID-19, inflammatory mediators produced by viral infections need to be assessed [19, 20]. Considering cytokine storm, short-term use of a small dose of corticosteroid can reduce the risk of acute respiratory distress syndrome [21]. Therefore, methylprednisolone (0.5−1 mg/kg) was administered to most patients in addition to human immunoglobulin in order to prevent cytokine storm. After patients showed a good response, the oxygenation index returned to ≥300 mmHg. The intervention continued for an average of 9 days after review. Timely administration of glucocorticoids and human immunoglobulin could effectively inhibit the progression of COVID-19 and reverse the condition [22, 23]. However, it needs further study whether glucocorticoids prolong the time of virus excretion from the body. In the treatment of severely ill COVID-19 patients, we added albendazole, whose efficacy against COVID-19 needs further evaluation.

Severe COVID-19 patients were subjected to early intervention under close monitoring. All patients improved and could be discharged alive, which is notably different from previous reports [21, 24]. We believe that the severity of the disease is independent of its prognosis [25, 26], suggesting that symptoms, clinical signs, laboratory results, and chest imaging findings are more reliable for assessing the prognosis of COVID-19 pneumonia.

Progressive acute respiratory distress syndrome is the most common cause of death in patients with severe COVID-19 pneumonia owing to an uncontrolled inflammatory response. Early isolation, early diagnosis, and early treatment with methylprednisolone and human immunoglobulin can reduce the mortality rate of patients with COVID-19 pneumonia.

Data Availability

The data included in this paper are available without any restriction.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] WHO, *Clinical Management of Severe Acute Respiratory Infection when Novel Coronavirus (nCoV) Infection is Suspected: Interim Guidance*, 2020, https://apps.who.int/iris/handle/10665/330854.

[2] C. Huang, Y. Wang, X. Li et al., "Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China," *The Lancet*, vol. 395, no. 10223, pp. 497–506, 2020.

[3] S. I. Alsharidah, M. Ayed, R. M. Ameen et al., "COVID-19 convalescent plasma treatment of moderate and severe cases of SARS-CoV-2 infection in Kuwait: a multicenter interventional study," *SSRN Electronic Journal*, p. 31, 2020.

[4] D. W. Wang, B. Hu, C. Hu et al., "Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus–infected pneumonia in Wuhan, China," *JAMA*, vol. 323, no. 11, pp. 1061–1069, 2020.

[5] WHO, "Summary of probable SARS cases with onset of illness from 1 November 2002 to 31 July 2003," 2003, https://www.who.int/csr/sars/country/table2004_04_21/en.

[6] WHO, "Middle East respiratory syndrome coronavirus (MERS-CoV)," 2019, http://www.who.int/emergencies/mers-cov/en.

[7] L. T. Phan, T. V. Nguyen, Q. C. Luong et al., "Importation and human-to-human transmission of a novel coronavirus in Vietnam," *The New England Journal of Medicine*, vol. 382, no. 9, pp. 872–874, 2020.

[8] D. Ray, M. Salvatore, R. Bhattacharyya et al., "Predictions, role of interventions and effects of a historic national lockdown in India's response to the COVID-19 pandemic: data science call to arms," 2020.

[9] S. Marchini, E. Zaurino, J. Bouziotis, N. Brondino, V. Delvenne, and M. Delhaye, "Study of resilience and loneliness in youth (18–25 years old) during the COVID-19

pandemic lockdown measures," *Journal of Community Psychology*, vol. 49, no. 2, pp. 468–480, 2021.

[10] W. K. Leung, K. F. To, P. K. S. Chan et al., "Enteric involvement of severe acute respiratory syndrome-associated coronavirus infection," *Gastroenterology*, vol. 125, no. 4, pp. 1011–1017, 2003.

[11] M. H. Jang, M. J. Shin, and Y. B. Shin, "Pulmonary and physical rehabilitation in critically ill patients," *Acute and Critical Care*, vol. 34, no. 1, pp. 1–13, 2019.

[12] D. Wu, D. Ellis, and S. Datta, "COVID-19, reduced lung function, and increased psycho-emotional stress," *Bioinformation*, vol. 16, no. 4, pp. 293–296, 2020.

[13] A. Assiri, A. McGeer, T. M. Perl et al., "Hospital outbreak of Middle East respiratory syndrome coronavirus," *The New England Journal of Medicine*, vol. 369, no. 5, pp. 407–416, 2013.

[14] S. Wan, Q. Yi, S. Fan et al., "Characteristics of lymphocyte subsets and cytokines in peripheral blood of 123 hospitalized patients with 2019 novel coronavirus pneumonia (NCP)," *MedRxiv*, 2020.

[15] J. H. Wu, X. Li, B. Huang et al., "Pathological changes of fatal coronavirus disease 2019 (COVID-19) in the lungs: report of 10 cases by postmortem needle autopsy," *Zhonghua Bing Li Xue Za Zhi = Chinese Journal of Pathology*, vol. 49, no. 6, pp. 568–575, 2020.

[16] J. C. Leemans, M. J. B. M. Vervoordeldonk, S. Florquin, K. P. van Kessel, and T. van der Poll, "Differential role of interleukin-6 in lung inflammation induced by lipoteichoic acid and peptidoglycan from Staphylococcus aureus," *American Journal of Respiratory and Critical Care Medicine*, vol. 165, no. 10, pp. 1445–1450, 2002.

[17] Z. Xu, L. Shi, Y. Wang et al., "Pathological findings of COVID-19 associated with acute respiratory distress syndrome," *The Lancet Respiratory Medicine*, vol. 8, no. 4, pp. 420–422, 2020.

[18] P. Zhang, J. Li, H. Liu et al., "Long-term bone and lung consequences associated with hospital-acquired severe acute respiratory syndrome: a 15-year follow-up from a prospective cohort study," *Bone Res*, vol. 8, no. 1, p. 8, 2020.

[19] J. T. Wu, K. Leung, and G. M. Leung, "Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study," *The Lancet*, vol. 395, no. 10225, pp. 689–697, 2020.

[20] N. Chen, M. Zhou, X. Dong et al., "Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study," *Lancet*, vol. 395, no. 10223, pp. 507–513, 2020.

[21] J.-Y. Wang, S.-Y. Chang, Y.-W. Huang, and S. C. Chang, "Serology-positive but minimally symptomatic COVID-19 may still cause lung injury and lung function impairment," *The International Journal of Tuberculosis and Lung Disease*, vol. 24, no. 6, pp. 568-569, 2020.

[22] Y. D. Wan, T. W. Sun, Z. Q. Liu, S. G. Zhang, L. X. Wang, and Q. C. Kan, "Efficacy and safety of corticosteroids for community-acquired pneumonia: a systematic review and meta-analysis," *Chest*, vol. 149, no. 1, pp. 209–219, 2016.

[23] X. Li, C. Wang, S. Kou, P. Luo, M. Zhao, and K. Yu, "Lung ventilation function characteristics of survivors from severe COVID-19: a prospective study," *Critical Care*, vol. 24, no. 1, p. 300, 2020.

[24] J. H. Hull, J. K. Lloyd, and B. G. Cooper, "Lung function testing in the COVID-19 endemic," *The Lancet Respiratory Medicine*, vol. 8, no. 7, pp. 666-667, 2020.

[25] D. S. Hui, K. T. Wong, F. W. Ko et al., "The 1-year impact of severe acute respiratory syndrome on pulmonary function, exercise capacity, and quality of life in a cohort of survivors," *Chest*, vol. 128, no. 4, pp. 2247–2261, 2005.

[26] W.-j. Guan, Z.-y. Ni, Y. Hu et al., "Clinical characteristics of 2019 novel coronavirus infection in China," *med Rxiv*, 2020.

Hindawi
Computational and Mathematical Methods in Medicine
Volume 2021, Article ID 9940148, 10 pages
https://doi.org/10.1155/2021/9940148

Research Article

Expression-EEG Bimodal Fusion Emotion Recognition Method Based on Deep Learning

Yu Lu,[1] **Hua Zhang**,[1] **Lei Shi**,[1] **Fei Yang**,[1] **and Jing Li**[2]

[1]*Fuyang Vocational and Technical College, Fuyang, Anhui 236031, China*
[2]*Department of Electrical & Information Engineering, Sichuan Engineering Technical College, Deyang, Sichuan 618000, China*

Correspondence should be addressed to Jing Li; lj@scetc.edu.cn

Received 26 March 2021; Revised 23 April 2021; Accepted 10 May 2021; Published 26 May 2021

Academic Editor: Kaijian Xia

As one of the key issues in the field of emotional computing, emotion recognition has rich application scenarios and important research value. However, the single biometric recognition in the actual scene has the problem of low accuracy of emotion recognition classification due to its own limitations. In response to this problem, this paper combines deep neural networks to propose a deep learning-based expression-EEG bimodal fusion emotion recognition method. This method is based on the improved VGG-FACE network model to realize the rapid extraction of facial expression features and shorten the training time of the network model. The wavelet soft threshold algorithm is used to remove artifacts from EEG signals to extract high-quality EEG signal features. Then, based on the long- and short-term memory network models and the decision fusion method, the model is built and trained using the signal feature data extracted under the expression-EEG bimodality to realize the final bimodal fusion emotion classification and identification research. Finally, the proposed method is verified based on the MAHNOB-HCI data set. Experimental results show that the proposed model can achieve a high recognition accuracy of 0.89, which can increase the accuracy of 8.51% compared with the traditional LSTM model. In terms of the running time of the identification method, the proposed method can effectively be shortened by about 20 s compared with the traditional method.

1. Introduction

Emotional computing is widely used in the fields of games, mental health, learning, and education. The goal is to develop a computing system that can carry out "emotional communication." It is required that the algorithm's type recognition, degree judgment, and feedback speed of emotions should be as close as possible to real human emotional responses [1–4]. As one of the key issues in the field of emotional computing, emotion recognition has rich application scenarios and important research value and has attracted widespread attention in recent years [5–7]. Taking medical work as an example, real-time and accurate grasp of the physical and psychological conditions of unconscious patients play an important role in the recovery of patients. Therefore, it is of great significance to use the existing pattern recognition technology to study difficult problems such as the physical condition of unconscious patients.

Emotion recognition methods can be divided into two categories: single-modality and multimodality [8–10]. Single-modal data acquisition is easier, and analysis methods are more mature. Therefore, most previous researches focused on monomodal emotion recognition, that is, emotion recognition for a pattern of data, such as traditional audio, facial expressions, physiological signals, or a certain form of data in text and body movements [11].

Multimodal emotion recognition can use all the characteristics of different signals such as human expressions and brain signals, so that the complementary characteristics of multiple modalities in expressing emotions are reflected in the final algorithm results, thereby improving the recognition accuracy [12, 13].

As the product of current artificial intelligence technology and big data collection, deep neural network combines multimodal emotion recognition methods with deep learning networks. Through multilayer network model iterative

training and learning [14, 15], the signal data can be extracted and calculated effectively. And based on the continuous learning of its own network, the network model parameters can also be adjusted in time. The problem of feature redundancy and lack of key features is solved, and the corresponding emotion recognition performance is improved.

The rest of this article is organized as follows. The second chapter introduces the related research in recent years. The third chapter introduces the bimodal emotion recognition method based on deep learning, including facial expression feature extraction and EEG signal features. The fourth chapter introduces the experimental simulation analysis of the feasibility and optimality of the method proposed in this paper based on the MAHNOB-HCI data set. The fifth chapter is the conclusion of this article.

2. Related Works

In recent years, researchers have conducted extensive research on various types of modal information that can express emotions. Studies have found that changes in human emotions can cause changes in expression, behavior, psychology, and physiology. Among them, facial expressions, postures, and physiological signals can independently express certain emotions [16–18].

In the research of monomodal emotion recognition, video, speech, text, and physiological signals all have certain expressions of emotion. Reference [19] uses ConvNet and DBNs to obtain information from videos, which has good performance on some emotions. Reference [20] summarized the methods and achievements of emotion recognition using electroencephalogram (EEG) in recent years. Reference [21] uses the convolutional-recurrent neural network (CRNN) to perform emotion recognition on multichannel EEG data and achieves ideal results.

But it needs to be pointed out that these types of emotional information are comprehensively displayed in the process of people communicating with each other. At present, researchers found that unimodal data has certain restrictions on the expression of emotions, and each mode has different sensitivities to different emotions [22]. Since each modal has a certain expression of emotion, some researchers have begun to conduct multimodal fusion emotion recognition research. Reference [23] uses a dual-mode autoencoder to study the emotional performance of EEG and eye movement signals. Experiments show that compared with the fusion of the two, the effect of identifying EEG features and eye movement features separately is poor. Reference [24] introduced a new method of modeling spatiotemporal information using three-dimensional convolutional neural networks (C3D) and combined it with a multimode deep belief network (MMDBN), which can represent audio and video stream cascades. Experiments on the eNTERFACE multimodal emotion database show that this method improves the performance of multimodal emotion recognition and is significantly better than the latest research scheme. Reference [25], based on the radial basis function and support vector machine network model, proposed a multimodal emotion recognition metric learning (MERML); a unified analysis of audio and video has a good performance in emotion recognition. This type of method usually only combines features by simple splicing, which easily causes feature redundancy. For video signals containing very large sample data, this will cause unnecessary experimental costs.

At the specific algorithm level of multimodal emotion recognition, with the breakthrough of deep learning methods in the computer field, neural network models are gradually applied to emotion recognition tasks [26]. The recurrent neural network (RNN) model has received extensive research and attention due to its obvious advantages in processing sequence tasks. Although recurrent neural networks can rely on cyclic connections to capture certain sequence context information, RNNs have the problem of vanishing gradients during back propagation. That is, as the number of neural network layers increases, the amount of values transferred is small and cannot cause parameter disturbances. An effective technology to overcome the problem of RNN gradient disappearance is adopted, namely, the Long Short-Term Memory (LSTM) network [27]. The LSTM network structure selectively "forgets" some inputs and "shields" some outputs through the "gate" structure so as not to affect the weight update of the next layer, so that the LSTM network can learn the best timing information related to the classification task [28]. In emotion recognition tasks, because of the continuity of facial expressions and EEG signals, emotional expressions are highly correlated in time series. However, single-point facial expression pictures and EEG signal data are often in the process of facial expression changes, which are prone to misjudgment [29, 30]. Therefore, for sequence emotion recognition tasks, LSTM's processing of sequences is similar to the processing method of the human brain on emotion recognition tasks, and the algorithm has the advantage of natural adaptability.

In view of the existing research work on emotion recognition, this paper proposes a deep learning-based expression-EEG bimodal fusion emotion recognition method. The main contributions are as follows:

(1) Aiming at the problem of low accuracy of single-modal emotion recognition, combined with the advantages of human expression signal recognition and EEG signal emotion recognition, the accuracy of emotion recognition is improved, and the six emotions of anger, disgust, fear, happiness, sadness, and surprise are realized via accurate classification and identification

(2) Facing the recognition accuracy and real-time requirements of emotion recognition models, it is based on the improved VGG-FACE network to realize the extraction of expression features and EEG features. First, the self-attention mechanism is introduced between the hierarchical networks to better distinguish each training layer and enhance the robustness of the system. The penalty term is introduced in the loss function to further improve the network and realize the diversification of the state vector of each layer. At the same time, it reduces the time of model training and learning and has a good recognition effect

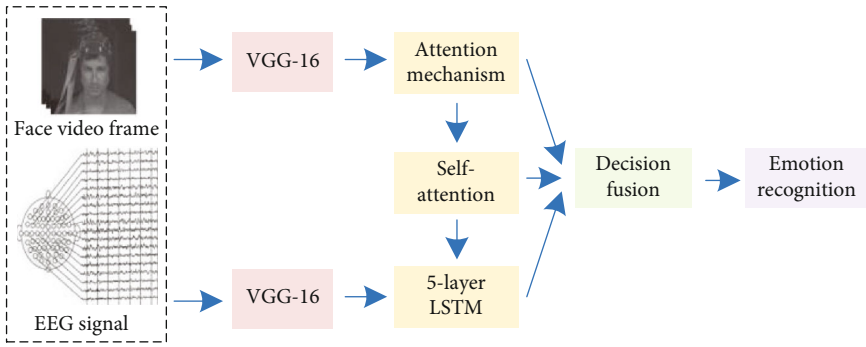


Figure 1: Recognition system framework based on VGG-LSTM network model.

3. Bimodal Emotion Recognition Method Based on Deep Learning

3.1. Expression-EEG Interactive Emotion Recognition Model. The system framework is shown in Figure 1. The LSTM emotion recognition model proposed in this paper for interactive collaboration between EEG signals and face video mainly includes two stages: feature extraction and interaction collaboration. In the feature extraction stage, first select key signal frames that need to be focused for data preprocessing and then extract features with strong expression and generalization capabilities. In the interactive collaboration stage, the features of the two modalities are first fused and learned. The special feature is that this article will also use the spatial frequency band attention mechanism to calculate the importance of the visual images of the α, β, and γ waves in the EEG signal. Reinforcement learning (RL) is performed through the time-domain attention mechanism to calculate the key signal frame time information that needs to be focused at the next time point and feed it back to the feature extraction stage. Finally, the emotion recognition result is outputted by the emotion classifier. Under this model, a closed loop is formed between the input signal and the model action—a process of selectively and repeatedly focusing on the multimodal signals of human emotions for emotion recognition.

3.2. Facial Expression Extraction. This paper uses fine-tuning to complete the retraining of the pretrained network. The advantage of fine-tuning is that you can use limited data to make the model achieve the desired effect. This paper uses the face data set FER2013 (the Facial Expression Recognition 2013 Dataset) to fine-tune the existing VGG-FACE network.

VGG-FACE is a 16-layer or 19-layer CNN architecture developed by the Visual Geometry Group (VGG) of Oxford University, which performs well in face recognition tasks [31]. Unlike VGG trained on the ImageNet data set, VGG-FACE is trained on a data set that only contains face data. And a deep convolution neural network model (DCNN) without pretraining as a baseline for experimental comparison was also introduced.

All the above network models are fine-tuned using the FER2013 data set. It was verified on the FER2013 test set and SFEW validation set to observe its performance. The experimental results are shown in Table 1.

As can be seen from the data in the above table, most of the pretrained network experimental results are better than the DCNN without pretraining, because the pretrained network has better initialization model parameters. Whether it is on FER2013 or SFEW, the best results are the VGG-FACE network pretrained on the face data set. The network reached an accuracy of 89.21% on the FER2013 test set and an accuracy rate of 78.24% on the SFEW test set. According to the experimental results, this paper finally uses a pretrained 16-layer VGG-FACE network. The network is fine-tuned on FER2013, and the acquired features are outputted to the LSTM unit to identify timing features.

The attention mechanism can be introduced between the input and output of the model, so that the performance of the model can be improved [32]. As shown in Figure 2, the main working principle of the attention mechanism is as follows: imagine the elements in the source as a series of $<K, V>$ data pairs, determine the element Q, calculate the correlation between Q and each K, obtain the weight coefficient of each K corresponding to V, and then perform weighting on V and get the final attention value Y_{att}:

$$Y_{\text{att}} = \sum_{i=1}^{l} \text{Similarity}(Q, K_i) \cdot V_i, \tag{1}$$

where l represents the length of source, and the meaning of the formula is as described above. The self-attention mechanism does not refer to the attention mechanism between the target and the source but occurs between the internal elements of the source or target. The attention mechanism can be understood as the situation of $K = V = Q$. The self-attention mechanism can more easily capture the long-distance interdependent features in the input sequence.

In the applied stacked LSTM network, three LSTM stacks are used to ensure that the model can learn higher-level temporal feature representation. Sequence data operations based on LSTM mean that the addition of layers increases the abstraction level of the input observation time and has better expressive capabilities.

In order to make each layer of LSTM in the stacked LSTM network have different proportions, the network model in this paper is further improved, and a self-attention mechanism is introduced between each layer of the LSTM network. It is worth noting that, unlike the

Table 1: Experimental simulation results of fine tuning of each model.

Network model	Pretraining data set	FER 2013 test set results (%)	SFEW test set results
Baseline DCNN	Null	56.21	54.35
GoogLeNet	ImageNet	61.23	59.21
CaffeNet	ImageNet	67.41	56.22
Residual network	ImageNet	72.34	68.34
VGG	ImageNet	76.21	71.23
VGG-FACE	Faces	89.21	78.24

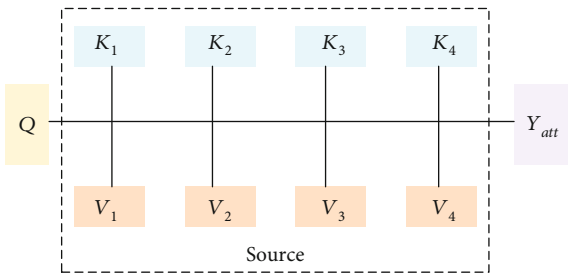


Figure 2: Schematic diagram of attention mechanism.

attention mechanism, it can be updated iteratively through its own information. The flowchart of this part of the method is shown in Figure 3. This network model is mainly composed of a stacked LSTM network embedded with a self-attention mechanism. The hidden state and unit state of the stacked LSTM are used as the input of the self-attention mechanism module, and the output is the corresponding weight vector.

$$\mathbf{u}^t = \mathbf{v}^T \tanh\left(\mathbf{W}_s \mathbf{X}_t + \mathbf{b}\right), \tag{2}$$

$$\mathbf{a}^t = \text{Softmax}\left(\mathbf{u}^t\right), \tag{3}$$

where the dimension of vector $\mathbf{X}_t$ is $n \times r$, the dimension of vector $\mathbf{W}_s$ is $r \times d_a$, $\mathbf{b}$ and $\mathbf{v}^T$ are vectors of dimension d_a. $\mathbf{W}_s$, $\mathbf{b}$, $\mathbf{v}^T$ are the parameters of the network model, and $\mathbf{X}_t$ is the input of the self-attention mechanism module, which represents the hidden state $\mathbf{Y}_t$ or the unit state $\mathbf{Z}_t$ of a certain layer in the stacked LSTM.

$$\mathbf{Y}_t = \left(y_t^{(1)}, y_t^{(2)}, \cdots, y_t^{(l)}\right), \tag{4}$$

$$\mathbf{Z}_t = \left(z_t^{(1)}, z_t^{(2)}, \cdots, z_t^{(l)}\right). \tag{5}$$

Dot multiplying the weight vector $\mathbf{a}^t$ with the state value of LSTM, we can get

$$\mathbf{G}_t = \mathbf{a}^t \mathbf{X}_t, \tag{6}$$

where $\mathbf{G}_t$ is the weighted vector $\mathbf{Y}_t'$ or $\mathbf{Z}_t'$ obtained after the stack LSTM is updated. After calculation by the self-attention mechanism, different weights can be assigned to each layer of the network in the stack LSTM according to their importance. The network has been optimized to a certain extent, and the expression ability of the hierarchical features is improved.

Since the self-attention mechanism between adjacent time steps tends to assign similar weights, this paper adds a penalty term to prevent this problem from occurring and makes the weight vectors of different levels more diverse. While optimizing the weight, the penalty term not only reduces the redundant feature information but also makes the hierarchical relationship in the stacked LSTM more differentiated. This paper uses the statistical variance method to optimize the network.

$$P = \frac{1}{T}\sum_{t}^{T}\sum_{i}^{L}\left(\left(\alpha_{ti} - \mu\right)^2 + \left(\beta_{ti} - \eta\right)^2\right), \tag{7}$$

$$\mu = \frac{1}{L}\sum_{i}^{L}\alpha_{ti}, \eta = \frac{1}{L}\sum_{i}^{L}\beta_{ti}. \tag{8}$$

In the formula, α_{ti} and β_{ti}, respectively, represent the attention weight of the hidden state and the unit state at different time steps and levels: minimize it together with the original loss function.

$$L_d = -\log\left(p(y|a)\right) - P, \tag{9}$$

where $-\log\left(p(y|a)\right)$ represents the cross-entropy loss function, a represents the actual output of the model, and y represents the sample label.

3.3. EEG Feature Extraction

3.3.1. EEG Feature Collection. EEG signals are the distribution of potentials on the scalp produced by brain neuron activity and are usually obtained by using an EEG device. The electrodes placed on the scalp transmit the electrical signals generated by the brain to the signal collector and then perform preamplification and electronic filtering (such as a 50 Hz notch filter). Then, through the power amplifier and A/D converter, the analog signal is converted into a digital signal that can be processed by the computer and then transmitted to the computer for relevant analysis and processing.

According to the recommendations of the International Electroencephalography Society, the current electrode placement for EEG acquisition generally adopts the international 10/20 system standard, as shown in Figure 4. Divide the connecting lines of the root of the nasion, vertex, and inion in equal proportions of 10%, and then divide the connecting lines of the nasion, external, ear hole, and inion into 10 equal parts. The electrode position is determined according to the intersection of the concentric circle centered on the vertex and the radius, and most of them are placed at the position of an integer multiple of 10% or 20% of the connecting line, so it is called a 10/20 system. There are a total of 21 electrodes, of which A1 and A2 are reference electrodes, as

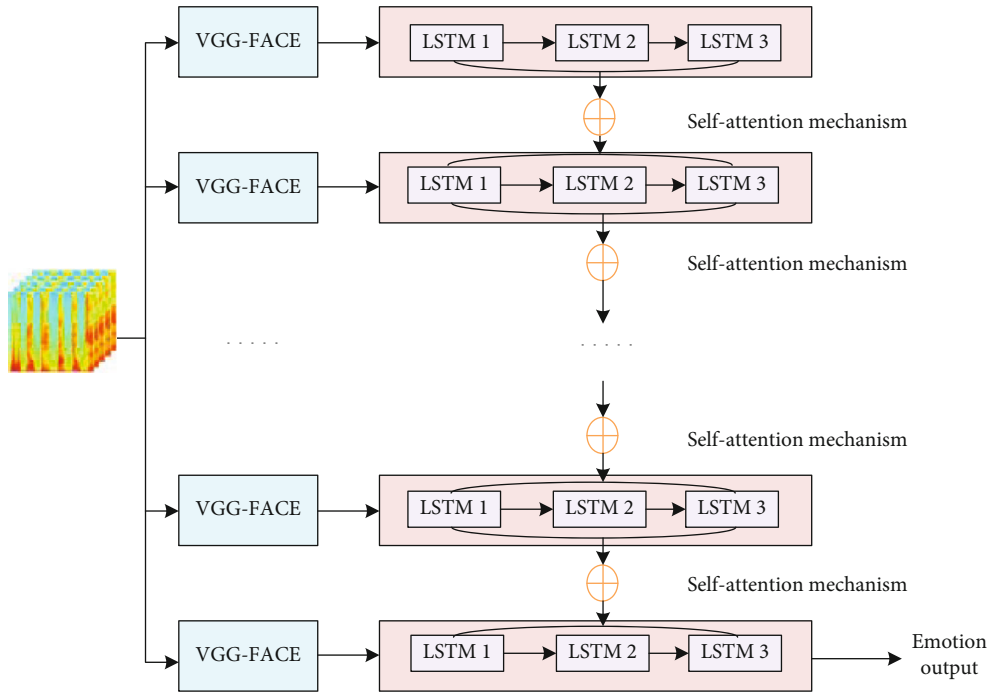
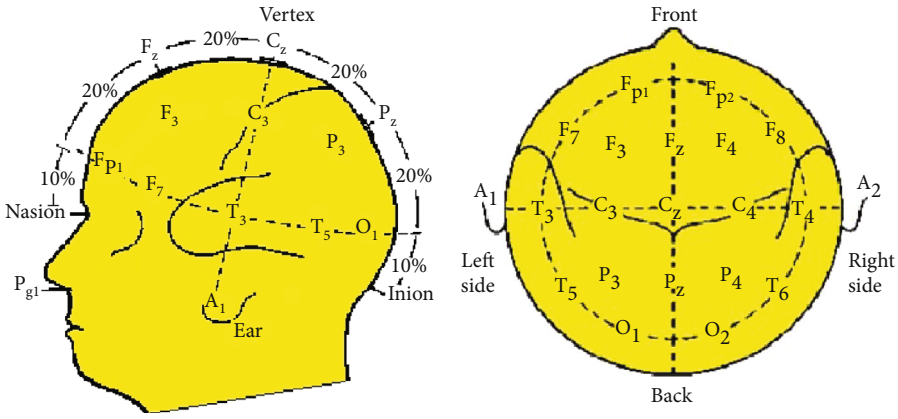


Figure 3: Schematic diagram of hybrid model.



(a) Side view of electrode placement

(b) Top view of electrode placement

Figure 4: Schematic diagram of piecewise linear transformation.

shown in Figure 4(b). The beginning of each electrode name uses one or two letters to indicate its area, as shown in Table 2. After the electrode name, a number or letter is used to indicate the distance from the center. An odd number means the left brain, and an even number means the right brain: the larger the number, the farther away from the center line. The position of the center line uses the mark "z" to represent the number 0 to distinguish it from the letter O. Modern 32-lead or 64-lead electrode caps are also based on the 10/20 system expansion. However, it should be noted that different EEG systems often have different names for similarly located electrodes.

Table 2: Correspondence of electrode code.

Symbol	Name
Fp	Front pole
F	Frontal
C	Central
P	Parietal
O	Occipital
T	Temporal

3.3.2. Data Preprocessing. The EEG signal will be affected by the experimental equipment and the breathing movement of the collected person during the collection process. Noise may interfere with the EEG signal, making the measurement result of the original signal unreliable. The purpose of preprocessing is to improve the overall recognition quality of EEG signals for more accurate analysis and measurement.

The main categories of noise are low frequency baseband drift (BW) caused by breathing and body movement, high frequency random noise caused by power system interference (50 or 60 Hz), muscle movement, and random offset caused by poor electrode contact with the muscle interference.

Figure 5(a) shows the original EEG signals collected. In the filtering process, a 35 Hz Butterworth filter and a 50 Hz power interference removal filter are designed to eliminate power frequency interference, myoelectric interference, and electromagnetic interference in most power systems. Then, the signal is filtered by wavelet packet decomposition to remove baseband drift, and the EEG signal that removes interference and baseband drift is shown in Figure 5(b).

3.3.3. Feature Extraction Process. Face video is the facial activity signal of experimental participants collected by an ordinary camera, which belongs to the visual signal. The method of collecting EEG signals is to allow experiment participants to wear electrode EEG caps while watching emotion-inducing videos, so as to obtain EEG signals from 32 different positions on the human cerebral cortex. It is difficult to directly merge two heterogeneous signals. For this reason, this paper proposes to extract features with strong expression ability and generalization ability and at the same time make the features of the two modalities effectively interact and cooperate. For face videos, facial expression features are extracted based on VGG-FACE. The feature extraction process of the face video is as follows: first, the face area in the video frame is detected by the VGG-FACE model. Then, use the VGG-FACE model to extract features from the face area. Finally, use the fully connected layer to process the features and output the final feature vector $\mathbf{x}_{v,n}$.

The feature extraction of the EEG signal is more complicated: Firstly, the original EEG signal is removed by the wavelet soft threshold algorithm to remove artifacts, thereby obtaining a relatively pure signal. Then, the EEG signal is divided into segments with a duration of T. Next, extract the spectral energy information of the three brainwave frequency bands of α wave, β wave, and γ wave from the t^{th} segment data and visualize it on the 32 electrodes of the corresponding electrode caps to obtain the three frequency bands of EEG images. The rising β wave of human emotion activation will be significantly enhanced in the forehead. Finally, CNN is used to extract the layer features $e_{\alpha,n}$, $e_{\beta,n}$, and $e_{\theta,n}$ of the EEG images of the three frequency bands to fuse, as shown in equations (10) and (11).

In the calculation, the spatial frequency band attention mechanism is used to calculate the importance e_n' of the three groups of features, and finally, the fully connected layer is used to process the e_n' output feature vector $x_{e,n}$.

$$e_n' = e_{\alpha,n}\theta_{en,1} + e_{\beta,n}\theta_{en,2} + e_{\theta,n}\theta_{en,3}. \tag{10}$$

In the formula, $\theta_{en,1}$, $\theta_{en,2}$, and $\theta_{en,3}$ represent the importance assigned to $e_{\alpha,n}$, $e_{\beta,n}$, and $e_{\theta,n}$respectively:

$$\theta_{en,i} = \frac{\exp\left(W_{h,i}h_{n-1} + b_{n,i}\right)}{\sum_{j=1}^{3}\exp\left(W_{h,j}h_{n-1} + b_{n,j}\right)}, \quad i = 1, 2, 3. \tag{11}$$

In the formula, $W_{h,i}$ represents the weight matrix to be learned, $b_{n,i}$ represents the deviation, and h_{n-1} represents the hidden state of the multilayer LSTM at a time point $n-1$.

3.4. Expression-EEG Bimodal Fusion Emotion Recognition. This paper integrates facial expressions and speech signals for emotion recognition and uses the decision fusion method [33] to solve the fusion problem of two different modalities. The purpose of the decision fusion is to deal with the categories generated by each model and use specific criteria for redifferentiation. In the realization of this article, both facial expression recognition and speech emotion recognition use the Softmax function for classification. Their outputs are defined as

$$S^{\text{face}} = \left\{S_1^{\text{face}}, S_2^{\text{face}}, S_3^{\text{face}}, \cdots, S_k^{\text{face}}\right\}, \tag{12}$$

$$S^{\text{speech}} = \left\{S_1^{\text{speech}}, S_2^{\text{speech}}, S_3^{\text{speech}}, \cdots, S_k^{\text{speech}}\right\}, \tag{13}$$

where k is the number of emotional categories, and the weighted decision fusion calculation is

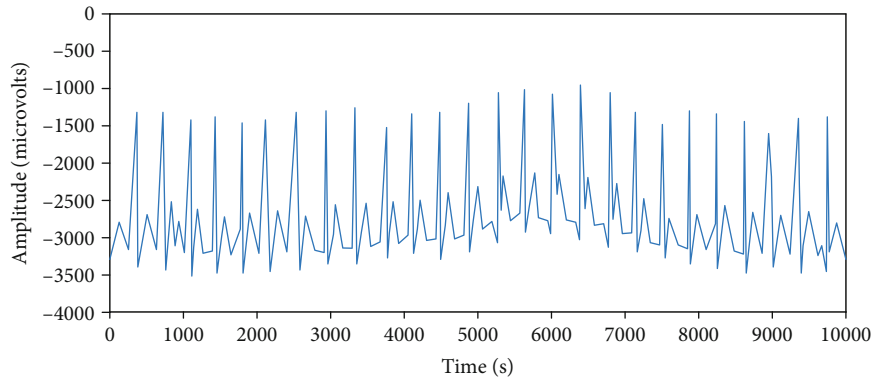
$$S = w_0 S^{\text{face}} + w_1 S^{\text{speech}}, \quad w_0 + w_1 = 1. \tag{14}$$

In the formula, w_0 and w_1, respectively, represent the weights assigned by the two modes.

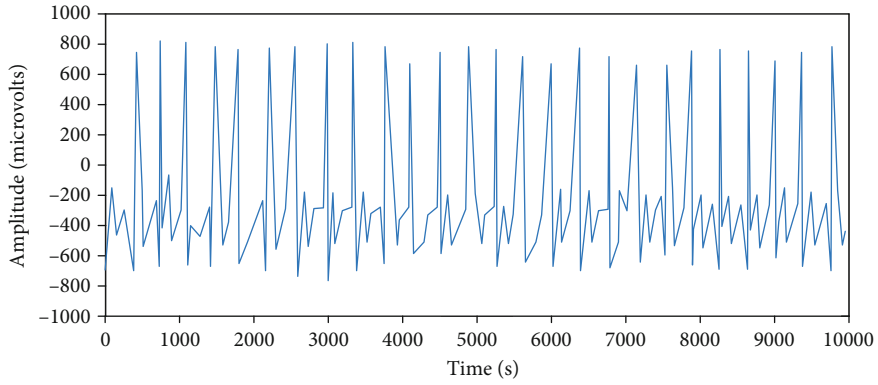
4. Experimental Scheme

In order to verify the feasibility and practicability of the method mentioned above, the experimental simulation robot hardware environment is a Lenovo ThinkPad E14, AMD Ryzen 7 4700U 8-core processor, 16 GB RAM, and integrated graphics. The software environment is operating system Chinese Windows 10 and English version software Microsoft Visual Studio 2012.

This paper uses the Caffe deep learning framework to implement model training and testing on the MAHNOB-HCI data set. The data of 35 experimental participants in the MAHNOB-HCI data set is divided into training set A, validation set A1, and test set B at a ratio of 5 : 1 : 1. In the process of data preprocessing, the face video of the data set is downsampled to 8 fps. At the same time, the face image in the video is detected and cropped, and the image size is rescaled to 300 × 300. In the training process, this paper uses the Adam method [24] to update the parameters. The sample set used for each update is obtained by extracting a minibatch = 12 samples from the training set A through the experience playback mechanism. In order to prevent the model from overfitting, the value of the dropout is set to 0.5. Set the value of the maximum time step Nmax to 30. In



(a) Primitive EEG signal



(b) Original EEG signal after processing

Figure 5: Schematic diagram of comparison before and after EEG signal processing.

addition, all fine-tuned VGG-16 networks used in this article have fixed parameters and are only used to extract features. In the experiment, rotation, flip, color distortion, and image transformation are used to expand the data. The whole data set was initially trained for 100 cycles with a batch size of 50. The initial learning rate of the model is 0.015, which is set to 0.001 after 10000 iterations. Set the weight decay and momentum to 0.00015 and 0.87, respectively. It is worth noting that the deep emotion recognition model is trained using a stochastic gradient descent scheme.

4.1. Sentiment Recognition Model Optimization and Analysis

4.1.1. The Effect of LSTM Stacking Layers on System Recognition Rate. In order to explore whether the number of LSTM layers will improve the experimental results accordingly, this paper is based on the baseline model to conduct comparative experiments under different layers of LSTM.

Figure 6 shows the effect of LSTMs with different layers on the recognition rate of the system. Experimental data shows that, compared with a single-layer network, a multilayer LSTM has a better recognition effect and can better extract abstract features in a sequence. When $L = 5$, the recognition effect on the selected data set is the best, and the recognition rate can reach 0.89. When $L > 5$, the displayed effect gradually decreases. Therefore, the number of LSTM layers selected in this paper is 5.

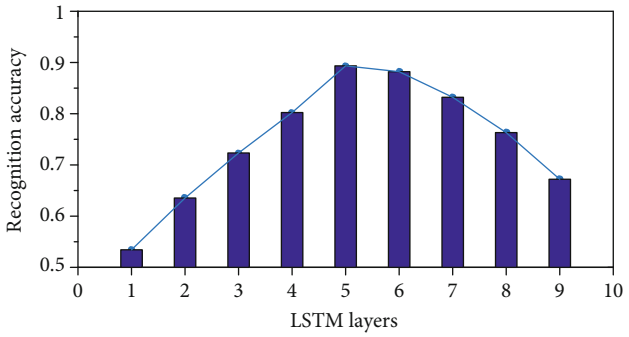
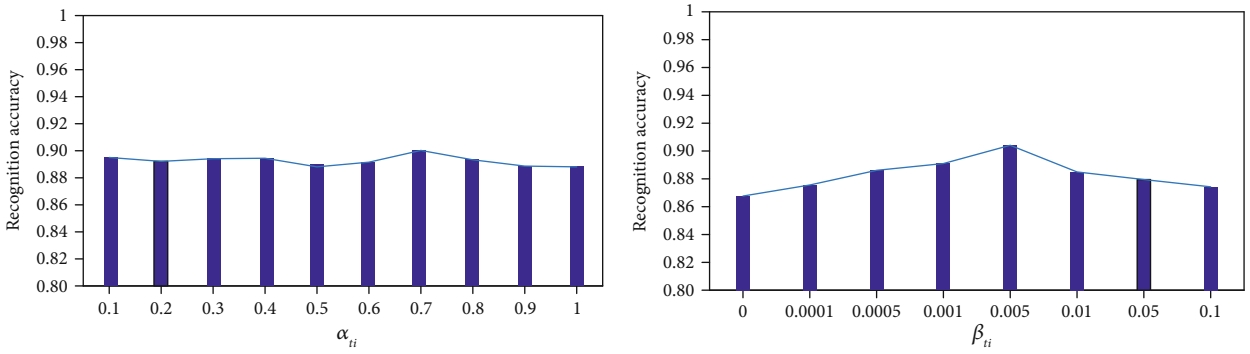


Figure 6: The recognition accuracy of LSTM under different layers.

Table 3: Identification accuracy rate under different processing conditions.

Treatment conditions	5-layer LSTM	
	Recognition accuracy	Running time (s)
Ordinary	0.82	121.34
Add attention mechanism	0.89	102.31

4.1.2. The Effect of Hierarchical Attention Mechanism on System Recognition Rate. After introducing the attention mechanism, different levels can be selectively paid attention



(a) Fixed β_{ti} value and variable α_{ti} value

(b) Fixed α_{ti} value and variable β_{ti} value

Figure 7: Model parameter sensitivity analysis.

to at each time step. In order to study whether the attention mechanism has a certain influence on the improvement of the network, this paper designs a corresponding comparative experiment.

As shown in Table 3, the introduction of the attention mechanism has improved the recognition effect of the model. After introducing the attention mechanism, the proposed model can achieve a high recognition accuracy of 0.89, which can increase the accuracy of 8.51% compared with the traditional LSTM model. In terms of the running time of the identification method, the proposed method can effectively be shortened by about 20 s compared with the traditional method. The attention mechanism is used to assign different proportions to each layer in the stacked LSTM, which is conducive to the network to filter out more useful information and improve the level of expression of the model and is more conducive to the extraction of image abstract features. Experimental results show that the introduction of the attention mechanism can improve the recognition effect.

4.1.3. The Impact of Penalty Items on System Recognition Rate. In the attention mechanism, the weight coefficient of attention is used to improve the recognition effect. Among them, the addition of penalty terms α_{ti} and β_{ti} can be used to update the weight coefficient, and the recognition models obtained by different weight coefficients are different. By introducing variance, the difference between different weight coefficients is obtained, and then, the back propagation algorithm is used to maximize the variance.

In order to analyze the sensitivity of different penalty terms α_{ti} and β_{ti} in the proposed model to optimize the model parameters, the MAHNOB-HCI data set recognition task was verified experimentally.

The optimal values of model parameters are analyzed in the self-collected data set, as shown in Figure 7. In the first experiment (a), α_{ti} was fixed at 0.001 according to the setting of multiple experiments, and β_{ti} was changed in [0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, and 1.0] to learn different models. It can be observed that the accuracy of the model generally increases first, reaches the maximum value at 0.7, and then decreases. The results show that when α_{ti} is set to 0.7, the network model has no effect on parameter selection.

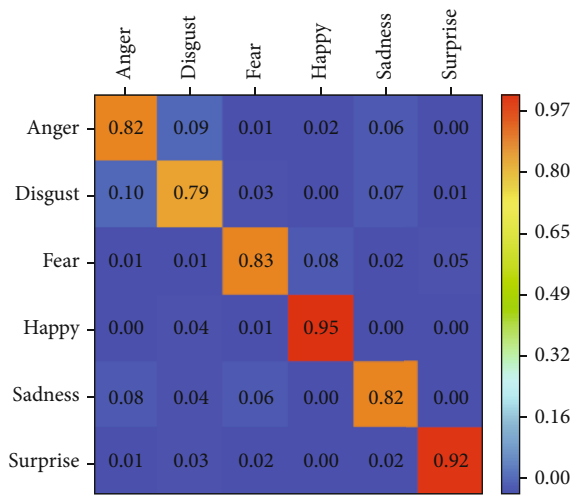


Figure 8: Schematic diagram of the confusion matrix of the MAHNOB-HCI data set.

Experiment (b) fixed the value of α_{ti} to 0.7 and changed β_{ti} in the set [0, 0.0001, 0.0005, 0.001, 0.005, 0.01, 0.05, and 0.1] for comparison experiments. The results show that the recognition performance is very sensitive to the value of the parameter β_{ti}, and $\beta_{ti} = 0.005$ guarantees the excellent recognition performance of deep learning features.

4.2. MAHNOB-HCI Data Set Identification Analysis. Based on the above determination of the model structure and related parameters, this paper uses the MAHNOB-HCI data set to perform expression testing on the final fusion network. The confusion matrix of the final identification result of the test set is shown in Figure 8. Each row represents the category to which the video really belongs, and each column represents the category given by the fusion network.

It can be seen from Figure 8 that the expression-EEG interaction model proposed in this paper performs very well in identifying "happy" and "surprise" samples, and the recognition accuracy can reach 0.95 and 0.92, respectively. In addition, it can be noticed that the fusion network has a low ability to recognize expressions of "disgust," but the recognition accuracy rate also reaches 0.79. As can be seen from the above figure, most "disgust" samples are mistaken for

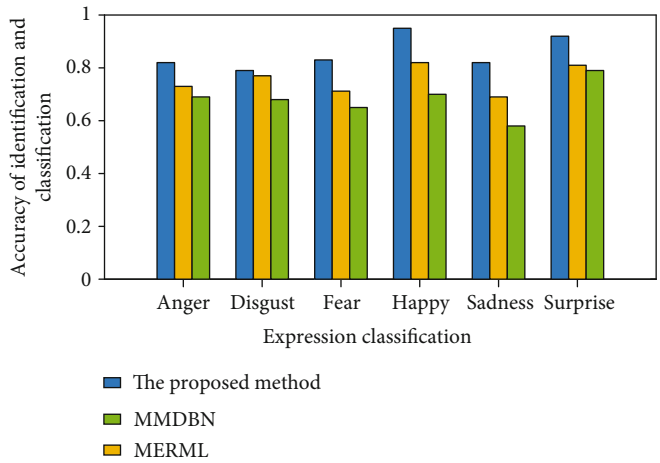


Figure 9: Emotion recognition accuracy rate under different methods.

"anger," "happy," and "sadness," while most "sadness" samples are classified as similar emotions, such as "anger," "disgust," and "fear."

4.3. Comparison of Facial Expression Recognition Classification Algorithms in MAHNOB-HCI Data Set. The MAHNOB-HCI data set contains expression signals and EEG signals. In order to ensure that the signal characteristics in each modal can better reflect the emotional information of the modal, the characteristics of the two signals are, respectively, fused to obtain the fusion characteristics representing each modal. The MMDDN method in reference [24], the MERML method in reference [25], and the method proposed in this paper have different methods for fusing multimodal features. The two method models are classified and identified under the MAHNOB-HCI data set, and the results are shown in Figure 9.

As shown in Figure 9, the proposed method has a higher accuracy in the classification and recognition of various emotions than the comparison method. The recognition accuracy rates of "anger," "disgust," "fear," "happy," "sadness," and "surprise" were 0.82, 0.79, 0.83, 0.95, 0.82, and 0.92, respectively.

Based on the above analysis, compared with other methods, the multimodal fusion feature obtained by separately fusing each modal feature in this paper has better performance in emotion recognition. It shows that with the reduction of the cost of multimodal feature selection, the classification performance of each emotion is also improved to a certain extent.

5. Conclusion

Multimodal emotion recognition is an important and challenging research problem in human-computer interaction. Facing the accuracy and real-time requirements of emotion recognition, this paper proposes a deep learning-based expression-EEG bimodal fusion emotion recognition method. This method is based on the improved VGG-FACE network model to realize the rapid extraction of facial expression features and shorten the training and learning time of the network model. The wavelet soft threshold algorithm is used to remove artifacts from EEG signals to extract high-quality EEG signal features. Then, the signal features extracted in the expression-brain electrical bimodal state are based on the long and short-term memory network model and the decision fusion method to realize the final bimodal fusion emotion classification and identification research. In terms of the running time of the identification method, the proposed method can effectively be shortened by about 20 s compared with the traditional method. The attention mechanism is used to assign different proportions to each layer in the stacked LSTM, which is conducive to the network to filter out more useful information and improve the level of expression of the model and is more conducive to the extraction of image abstract features. Experimental results show that the introduction of the attention mechanism can improve the recognition effect.

Analysis of the experimental results shows that the proposed method can reduce the model emotion recognition time by about 20 s compared with the traditional method, and the accuracy of the six typical emotion recognition can be maintained above 0.79. The focus of future research will be to explore the platformization of the proposed method and strive to realize the commercialization of the proposed method.

Data Availability

The data included in this paper are available without any restriction.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the Natural Science Research Project of Anhui Province Department of Education (No. KJ2019A1031).

References

[1] Q. Zhou, "Multi-layer affective computing model based on emotional psychology," *Electronic Commerce Research*, vol. 18, no. 1, pp. 109–124, 2018.

[2] J. D. Schwark, "Toward a taxonomy of affective computing," *International Journal of Human Computer Interaction*, vol. 31, no. 11, pp. 761–768, 2015.

[3] R. D. Ward and P. H. Marsden, "Affective computing: problems, reactions and intentions," *Interacting with Computers*, vol. 16, no. 4, pp. 707–713, 2004.

[4] C. Xu, Z. Feng, and Z. Meng, "Affective experience modeling based on interactive synergetic dependence in big data," *Future Generation Computer Systems*, vol. 54, no. 1, pp. 507–517, 2016.

[5] H. N. Wang, S. Q. Sun, T. Shu, and J. F. Wu, "Research on affective recognition with physiological signals based on IAGA and wkNN classification," *Advanced Materials Research*, vol. 143-144, no. 1, pp. 677–681, 2010.

[6] J. Granatyr, N. Osman, J. Dias et al., "The need for affective trust applied to trust and reputation models," *ACM Computing Surveys (CSUR)*, vol. 50, no. 4, pp. 1–36, 2017.

[7] Y. Guo, Y. Xia, J. Wang, H. Yu, and R. C. Chen, "Real-time facial affective computing on mobile devices," *Sensors (Basel, Switzerland)*, vol. 20, no. 3, pp. 1–15, 2020.

[8] L. Bozhkov, P. Georgieva, I. Santos, C. Silva, and A. Pereira, "EEG-based subject independent affective computing models," *Procedia Computer Science*, vol. 53, no. 1, pp. 375–382, 2015.

[9] M. Q. Yang, L. Lin, and S. Milekic, "Affective image classification based on user eye movement and EEG experience information," *Interacting with Computers*, vol. 30, no. 5, pp. 417–432, 2018.

[10] F. Lingenfelser, J. Wagner, J. Deng, R. Brueckner, B. Schuller, and E. Andre, "Asynchronous and event-based fusion systems for affect recognition on naturalistic data in comparison to conventional approaches," *IEEE Transactions on Affective Computing*, vol. 9, no. 4, pp. 410–423, 2018.

[11] W. J. Li, W. J. Luo, Y. W. Li, C. Y. Su, Y. H. Chen, and Y. Cao, "Speech emotion recognition based on separable convolution and LSTM," *Information technology*, vol. 44, no. 10, pp. 61–66, 2020.

[12] X. Li, G. M. Lu, J. J. Yan, and Z. Y. Zhang, "A review of multimodal dimension emotion prediction," *Acta Automatica Sinica*, vol. 44, no. 12, pp. 2142–2159, 2018.

[13] P. Lakhan, N. Banluesombatkul, V. Changniam et al., "Consumer grade brain sensing for emotion recognition," *IEEE Sensors Journal*, vol. 19, no. 21, pp. 9896–9907, 2019.

[14] A. T. Lopes, E. de Aguiar, A. F. de Souza, and T. Oliveira-Santos, "Facial expression recognition with convolutional neural networks: coping with few data and the training sample order," *Pattern Recognition*, vol. 61, no. 1, pp. 610–628, 2017.

[15] Q. Xu and S. Bo, "Expression recognition model based on deep learning and evidence theory," *Computer engineering and science*, vol. 29, no. 2, pp. 1–10, 2020.

[16] K. R. Scherer, M. Mortillaro, I. Rotondi, S. Trznadel, and I. Sergi, "Appraisal-driven facial actions as building blocks for emotion inference," *Journal of Personality & Social Psychology*, vol. 114, no. 3, pp. 358–379, 2018.

[17] C. Navarretta, "Feedback facial expressions and emotions," *Journal on Multimodal User Interfaces*, vol. 8, no. 2, pp. 135–141, 2014.

[18] S. L. Li, T. R. Zhang, N. Liu, W. Zhang, D. Tao, and Z. Wang, "Drivers' attitudes, preference, and acceptance of in-vehicle anger intervention systems and their relationships to demographic and personality characteristics," *International Journal of Industrial Ergonomics*, vol. 75, no. 1, article 102899, 2020.

[19] S. E. Kahou, X. Bouthillier, P. Lamblin et al., "EmoNets: multimodal deep learning approaches for emotion recognition in video," *Journal on Multimodal User Interfaces*, vol. 10, no. 2, pp. 1–13, 2015.

[20] S. M. Alarcao and M. J. Fonseca, "Emotions recognition using EEG signals: a survey," *IEEE Transactions on Affective Computing*, vol. 10, no. 3, pp. 374–393, 2019.

[21] X. Li, D. Song, P. Zhang, G. Yu, Y. Hou, and B. Hu, "Emotion recognition from multi-channel EEG data through convolutional recurrent neural network," in *IEEE international conference on Bioinformatics & Biomedicine*, pp. 352–359, Shen zhen, China, 2017.

[22] G. Zhang, T. Luo, W. Pedrycz, M. A. el-Meligy, M. A. F. Sharaf, and Z. Li, "Outlier processing in multimodal emotion recognition," *IEEE Access*, vol. 8, no. 1, pp. 55688–55701, 2020.

[23] W. Liu, W. L. Zheng, and B. L. Lu, "Emotion recognition using multimodal deep learning," in *international conference on neural information processing*, pp. 521–529, Springer International Publishing, 2016.

[24] D. Nguyen, K. Nguyen, S. Sridharan, A. Ghasemi, D. Dean, and C. Fookes, "Deep spatio-temporal features for multimodal emotion recognition," in *2017 IEEE Winter Conference on Applications of Computer Vision*, pp. 1215–1223, Santa Rosa, CA, 2017.

[25] E. Ghaleb, M. Popa, and S. Asteriadis, "Metric learning-based multimodal audio-visual emotion recognition," *IEEE Multimedia*, vol. 27, no. 1, pp. 37–48, 2020.

[26] J. Zhao, X. Mao, and L. Chen, "Learning deep features to recognise speech emotion using merged deep CNN," *IET Electric Power Applications*, vol. 12, no. 6, pp. 713–721, 2018.

[27] Y. Xie, R. Liang, Z. Liang, C. Huang, C. Zou, and B. Schuller, "Speech emotion classification using attention-based LSTM," in *IEEE/ACM transactions on audio*, vol. 27, no. 11pp. 1675–1685, Speech and Language Processing (TASLP), 2019.

[28] L. I. YJ, H. U. A. N. G. JJ, W. A. N. G. HY, and Z. H. O. N. G. N, "Study of emotion recognition based on fusion multi-modal bio-signal with SAE and LSTM recurrent neural network," *Journal on Communications*, vol. 38, no. 12, pp. 109–120, 2017.

[29] F. Shen, G. Dai, G. Lin, J. Zhang, W. Kong, and H. Zeng, "EEG-based emotion recognition using 4D convolutional recurrent neural network," *Cognitive Neurodynamics*, vol. 14, no. 6, pp. 815–828, 2020.

[30] D. Y. Choi and B. C. Song, "Semi-supervised learning for facial expression-based emotion recognition in the continuous domain," *Multimedia Tools and Applications*, vol. 79, no. 37-38, pp. 28169–28187, 2020.

[31] T. H. Kim, C. Yu, and S. W. Lee, "Facial expression recognition using feature additive pooling and progressive fine-tuning of CNN," *Electronics Letters*, vol. 54, no. 23, pp. 1326–1328, 2018.

[32] P. Barros, G. I. Parisi, C. Weber, and S. Wermter, "Emotion-modulated attention improves expression recognition: a deep learning model," *Neurocomputing*, vol. 253, no. 30, pp. 104–114, 2017.

[33] B. Sun, L. Li, X. Wu et al., "Combining feature-level and decision-level fusion in a hierarchical classifier for emotion recognition in the wild," *Journal on Multimodal User Interfaces*, vol. 10, no. 2, pp. 125–137, 2016.

Hindawi
Computational and Mathematical Methods in Medicine
Volume 2021, Article ID 9967592, 12 pages
https://doi.org/10.1155/2021/9967592

Research Article

Expression EEG Multimodal Emotion Recognition Method Based on the Bidirectional LSTM and Attention Mechanism

Yifeng Zhao and Deyun Chen

School of Computer Science and Technology, Harbin University of Science and Technology, Harbin, Heilongjiang 150080, China

Correspondence should be addressed to Deyun Chen; chendeyun@hrbust.edu.cn

Received 15 March 2021; Revised 15 April 2021; Accepted 30 April 2021; Published 11 May 2021

Academic Editor: Kaijian Xia

Due to the complexity of human emotions, there are some similarities between different emotion features. The existing emotion recognition method has the problems of difficulty of character extraction and low accuracy, so the bidirectional LSTM and attention mechanism based on the expression EEG multimodal emotion recognition method are proposed. Firstly, facial expression features are extracted based on the bilinear convolution network (BCN), and EEG signals are transformed into three groups of frequency band image sequences, and BCN is used to fuse the image features to obtain the multimodal emotion features of expression EEG. Then, through the LSTM with the attention mechanism, important data is extracted in the process of timing modeling, which effectively avoids the randomness or blindness of sampling methods. Finally, a feature fusion network with a three-layer bidirectional LSTM structure is designed to fuse the expression and EEG features, which is helpful to improve the accuracy of emotion recognition. On the MAHNOB-HCI and DEAP datasets, the proposed method is tested based on the MATLAB simulation platform. Experimental results show that the attention mechanism can enhance the visual effect of the image, and compared with other methods, the proposed method can extract emotion features from expressions and EEG signals more effectively, and the accuracy of emotion recognition is higher.

1. Introduction

With the rapid development of computer and artificial intelligence technology, computer intelligence is paid more and more attention. Emotion recognition is one of the important research issues of human-computer interaction and the key technology of computer intelligence [1]. The purpose of emotion recognition research is to build a robot system that can recognize human emotions and give correct feedback so as to make the human-computer interaction process more friendly [2]. Emotion recognition is the detection of human emotional states by computer technology such as classification recognition; that is, humans hope to make computers feel, understand, and even express their own emotions very well through intelligent technology and, at the same time, hope to complete relevant instructions through human emotional manipulation computers and realize the computer intelligence of emotion calculation [3].

The commonly used emotional signals at present include external intuitive emotional signals such as voice and posture and internal bioelectrical signals such as electroencephalogram and electrocardiogram [4]. Among them, the voice signal is the most convenient and direct way to get each other's emotional state from many external emotional signals [5]. But for the speech process, deliberately concealed information cannot make an accurate judgment. For the intrinsic bioelectric signals such as EEG and ECG, although there is no intuitive expression of emotional information such as voice signals, the emotion recognition system built by bioelectric signals is more reliable because the bioelectric signals are not easy to be modified by subjective manipulation. Among the many bioelectrical signals, the EEG can reflect the activity of hundreds of millions of neurons in the cerebral cortex and has the characteristics of convenient collection, simple cost, and high time resolution [6]. Studies have shown that different emotional states reflect different EEG signals in

the cortex of the brain. Therefore, EEG signals can be used as effective intrinsic bioelectrical signals to express human emotional information [7]. How to effectively process the EEG signal and extract the emotional information contained in the EEG signal is currently the main research issue of the emotion recognition of the EEG signal.

In order to build an emotion recognition system that is more in line with the human emotional processing mechanism, breaking through the bottleneck of multimodal emotion recognition technology which combines different physiological signals to assess emotional states has become a research hotspot [8]. The multimodal emotion recognition system can combine multimodal information and reduce the redundancy between different modes to make the emotional information complementary to each other so as to achieve a better emotion recognition effect [9]. Although some progress has been made in multimodal affective recognition technology, the current effect of affective recognition has not reached an ideal state due to the short research period. Therefore, how to construct a multimodal affective recognition system will be another key issue in affective computing research [10].

Based on the above analysis, an expression electroencephalogram multimodal affective recognition method based on the bidirectional LSTM and attention mechanism is proposed to solve the problems of the single mode, which are difficulty of deeply extracting signal characteristics and low accuracy of affective recognition.

(1) The bilinear convolution network (BCN) has designed a new projection layer and uses a one-dimensional filter to extract the optimal feature space. Therefore, the proposed strategy based on BCN extracts facial expression features, converts the EEG signal into three bands of image sequences, and uses BCN to fuse its image features to obtain more accurate expression EEG multimodal affective features

(2) In order to avoid the randomness or blindness of sampling methods, the attention mechanism is introduced in the LSTM network, and a characteristic fusion network with a three-layer bidirectional LSTM structure is designed to combine emoticons and EEG characteristics to improve the accuracy of emotion recognition

2. Related Works

Emotion is the role of the human cognitive process and conscious process, and it is an indispensable part of the human physiological process. Speech emotion recognition is relatively mature at present, and speech emotion recognition has been applied in different industries. The construction of a speech emotion recognition system mainly includes three parts: the establishment of a database, the extraction of emotion features, and the construction of an emotion classification model. In order to evaluate emotions more effectively, the research of the speech recognition system has been widely used by relevant researchers [11]. The establishment of an effective database is the basis of building a speech affective recognition system. Obtaining a subset of affective features that effectively represent the degree of emotional difference is the key to building a speech affective recognition system. The construction of a speech affective classification model is the key to achieving speech affective recognition. Reference [12] identifies three basic elements of the database, feature extraction, and classification method in the voice emotion recognition system, discusses the performance of the voice emotion recognition system, and uses the HMM (hidden Markov model), GMM (Gaussian mix model), and SVM (support vector machine) to achieve emotion recognition. A speech emotion recognition algorithm based on an improved stack kernel sparse depth model is proposed in Reference [13]. The algorithm is improved based on an automatic encoder, denoising automatic encoder, and sparse automatic encoder. The evaluation results on the test dataset show that the algorithm is better than the existing new algorithm in speech emotion recognition accuracy, but it is easy to be affected by the external environment in practical application, and there are many restrictions.

Among many bioelectrical signals, the EEG signal, as the product of brain neural activity, is more used in emotion recognition than other bioelectrical signals [14]. EEG signals are obtained by recording the potential changes in the cerebral cortex, which can be affected by external stimuli or human emotions, consciousness, and other activities. Reference [15] uses a classifier model to identify human emotional states or emotions through EEG signals and uses a classifier based on DBN (deep belief network) to classify and identify emotions using extracted features. Finally, the feasibility of the proposed method is analyzed from the indicators of accuracy, sensitivity, and specificity, and the results show that the proposed method can accurately identify emotions. Reference [16] provides an effective method of emotion recognition based on FAWT (flexible analytic wavelet transform). FAWT decomposes the EEG signal into different subband signals, uses the information potential to extract the features from the decomposition subband signal of the brain electrical signal, and then enters it smoothly into the random forest and support vector machine classifier that classifies the emotion. Compared with the existing method, the method shows better performance in human emotion classification. In the emotion recognition of electro-electrical signals, the current research difficulties are the establishment of effective databases and the extraction of EEG emotional characteristics; the above methods are based on a single modal database; although it has better popularity, the accuracy is not high.

At present, emotion recognition through a single signal has achieved a good recognition effect, but emotion recognition through a single signal is often disturbed by other signals, making the emotion recognition system in the process of emotion recognition have certain limitations. Therefore, in order to overcome the problem of the low recognition rate of the single modal emotional signal and the mutual influence between emotional signals, multimodal emotion recognition has become a research hotspot in emotion computing [17]. Multimodality is the combination of a variety of emotional signals, and various emotional signals complement

each other. From the perspective of human perception of the emotional mechanism and processing mode, multimodal emotion recognition can more truly realize the simulation and reconstruction of a human emotional processing mechanism. So, the construction of a multimodal emotion recognition system is of great significance for the realization of emotion computing [18]. Reference [19] proposes a multimodal emotion recognition method that integrates speech and brain electrical signals, induces joy, sadness, anger, and neutral emotion through sound stimulation, collects the corresponding speech and brain electrical signals, extracts the nonlinear geometric and attribute characteristics of brainpower, constructs the characteristic fusion algorithm based on the constraint Boltzmann machine, and constructs the multimodal emotion recognition system by using the secondary decision algorithm through decision fusion. The results showed that the overall recognition rate of the multimodal emotion recognition system constructed by feature fusion was 1.08% and 2.75% higher than the single modal voice signal and brain electrical signal, respectively. However, the proposed method cannot quickly locate the emotional key information from the multimodal signals containing a large amount of redundant information, so the overall efficiency of the model needs to be improved. In Reference [20], a novel visual expression feature extraction and classification method based on the displacement of specific continuous landmarks in the continuous frame of voice is proposed. The discrete wavelet transform is used to analyze the displacement signal of the landmark, and a variety of dimensionality reduction schemes are used to reduce the complexity of the derived model and improve the efficiency. The experimental results on the SAVEE, RML, and eNTERFACE05 databases show the effectiveness of the proposed method of visual feature extraction, but the improvement of the interactive collaborative fusion of heterogeneous multimodal signals of human emotion can further improve the accuracy of recognition. The basic idea of the traditional multimodal emotion recognition method is to manually design and extract the features of each mode, then fuse the multimodal signals, and finally train the pattern classifier using the marked dataset. However, this kind of method has low efficiency in processing large-scale human emotion data [21]. In recent years, popular deep learning methods have a strong ability of feature expression.

At present, most of the emotion recognition methods based on EEG and face videos regard the signals of two modes as time series and construct the LSTM emotion recognition model for the two modes, respectively, to learn the recognition results of each sequence and finally fuse the recognition results at the decision level. Reference [22] established a psychophysiological database, which classified the EEG, GSR, and heart rate of 30 participants exposed to the affective virtual environment. 743 features were extracted from physiological signals. Then, by using feature selection techniques, the dimensions of the feature space are reduced to a smaller subset of only 30 features. Using KNN, SVM, distinguished analysis (DA), and classification tree (four classification techniques), the emotional psychophysiological database is classified into four emotional clusters and eight emotional tags. The experimental results show that physiological signals can be used to classify emotional experiences with high accuracy.

Although more and more researchers pay attention to multimodal emotion recognition technology, the overall emotion recognition rate is relatively low, which is not enough to be applied in real life. Therefore, multimodal emotion recognition will be the key research part in emotion computing [23, 24].

3. Feature Extraction Based on the Bilinear Convolution Network

The input signals for emotion recognition are facial expressions and electroencephalogram signals collected by the participants when they watch the emotion-induced video. Among them, the facial expression is the facial activity signal collected by the ordinary camera and belongs to the visual signal [25]. The EEG (electroencephalogram) signal refers to the electrophysiological signal generated by the spontaneous and rhythmic movement of brain neurons recorded in the scalp surface in chronological order, and it belongs to the physiological signal. EEG signals are collected by allowing participants to wear electrode brain caps while watching emotion-induced videos to obtain EEG signals from 32 different locations in the human cerebral cortex. It is difficult to fuse two heterogeneous signals directly so that the method proposed in this paper extracts the features with strong expression and generalization capabilities, while effectively interacting and synergizing the features of the two modes. The feature extraction process is shown in Figure 1.

For facial expressions, BCN is used to extract facial expression features. Compared with traditional feature extraction methods, BCN has a stronger ability to mine potentially distributed expression features of the data [26]. As shown in Figure 1, for EEG signals, the first step is to convert the EEG signals into image sequences of three bands. This kind of visualization processing preserves the time-space characteristics of the EEG signals while unifying the two modes into images. Then, the features of EEG images are extracted based on the BCN and spatial band attention mechanism, explained as follows:

(1) Feature Extraction Process of the Facial Expression. First, use the Faster RCNN model to detect the face area in the video frame, then use BCN to extract features from the face area, and finally use the full connection layer to process features to output the final feature vector $f_{v,n}$.

(2) Feature Extraction Process of the EEG Signal. First, the original EEG signal is removed by the wavelet soft threshold algorithm to get a relatively pure signal. Then, the EEG signal is divided into each segment with a duration of T ($1/T$ corresponds to the frame rate of the facial expression) using the data processing method. At the same time, the spectrum energy information of α, β, and θ bands of EEG was extracted from the t-th band data, and the EEG images of three bands were visualized on the corresponding 32 electrodes of the

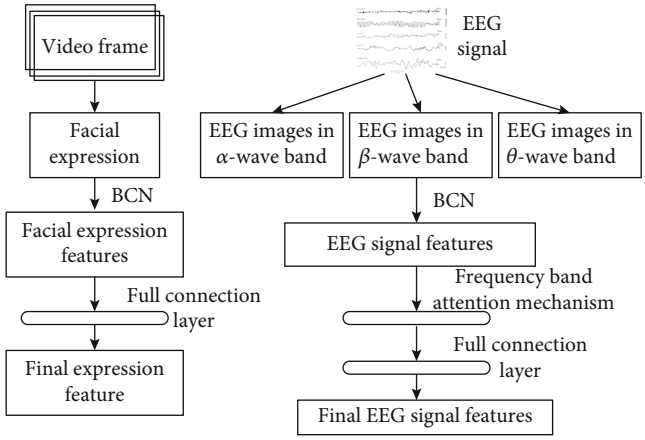


Figure 1: Feature extraction process of the facial expression and EEG signal.

electrode cap. As human emotional activation increases, β-wave forehead development increases significantly. Finally, BCN is used to extract layer features $f_{\alpha,n}$, $f_{\beta,n}$, and $f_{\theta,n}$ from three bands of EEG images; then, the importance f_n' of the three groups of features is calculated using the spatial band attention mechanism, and the f_n' output eigenvector $x_{e,n}$ is processed using the full connection layer. f_n' is calculated as

$$f_n' = f_{\alpha,n}\theta_{en,1} + f_{\beta,n}\theta_{en,2} + f_{\theta,n}\theta_{en,3}, \tag{1}$$

where $\theta_{en,1}$, $\theta_{en,2}$, and $\theta_{en,3}$ represent the importance assigned to $f_{\alpha,n}$, $f_{\beta,n}$, and $f_{\theta,n}$, respectively. The importance of each feature $\theta_{en,i}$ is as follows:

$$\theta_{en,i} = \frac{\exp\left(\omega_{h,i}h_{n-1} + b_{n,i}\right)}{\sum_{j=1}^{3}\exp\left(\omega_{h,j}h_{n-1} + b_{n,j}\right)}, \quad i = 1, 2, 3, \tag{2}$$

where $\omega_{h,i}$ is the weight matrix to be learned, $b_{n,i}$ is the deviation, and h_{n-1} is the hidden state of $n-1$ in the multilayer network.

The network structure of BCN is shown in Figure 2, which consists mainly of six network layers, namely, two projection layers, one one-dimensional convolution layer, two full connection layers, and one Softmax layer. BCN and traditional CNN have a big difference, mainly reflected in its network input as the SIFT feature matrix rather than the original picture, and we designed a new projection layer to find the optimal combination between key points and the optimal feature space and the characteristic direction of the use of the one-dimensional filter to learn the characteristics with high differentiation [27].

When constructing the input feature matrix, assuming that M key points have been located from each expression image and the N-dimensional SIFT feature vector is extracted for each key point, then the input of BCN is an $M \times N$ feature matrix. The feature matrix is then entered into the BCN for feature learning and classification [28]. In this process, the eigenmatrix first passes through a projection layer containing the multichannel left multiple projection matrix to find the optimal combination of key points. Similarly, in the second projection layer, the right multiple projection matrix will further optimize the feature space.

Suppose $H_t = \{h_{t,j}^{(l)} | j = 1,\cdots,C_l\}(t = 1,\cdots,N_l)$ is the t-th left multiple projection matrix containing C_l channels, $h_{t,j}^{(l)}$ is the j-th channel of H_t, and N_l is the number of left multiple projection matrices; then, the left multiple projection layer can be defined as

$$O_t = \sum_{j=1}^{C_l} h_{t,j}^{(l)} I_j, \quad t = 1, 2, \cdots, N_l, \tag{3}$$

where O_t is the t channel of the output feature and I_j is the j channel of the input matrix. Similarly, for the projection layer of the right multiple projection matrix, we have

$$O_t = \sum_{j=1}^{C_r} h_{t,j}^{(r)} I_j, \quad t = 1, 2, \cdots, N_r. \tag{4}$$

After the above operation, the right multiple projection layer projects each row of the input matrix from the current feature space to another feature space with a strong character and also plays the role of feature dimension reduction.

A one-dimensional convolution layer uses a set of multichannel filters to filter the characteristics entered, and the layer is often used in conjunction with the maximum pooling layer. In the process of convolution of input features, the one-dimensional filter can produce a strong response to local features with a specific structure and suppress the response value of local features different from the structure, which makes the filter capture the key local structure in the feature matrix. After the filtering operation, the convolution response is maximized to obtain the output characteristics of the convolution layer.

The output of the maximum pooling layer will pass through a nonlinear activation function to increase the flexibility of BCN before it is transferred into the full connection layer. Assuming that the output matrix of the maximum pooling layer can be represented by $Q = [Q_{c,i,j}]_{C\times I\times J}$ (where C, I, and J represent the number of channels, the number of rows, and the number of columns of the matrix, respectively), the nonlinear activation process can be expressed as follows:

$$F_{c,i,j} = \tanh\left(b_c + Q_{c,i,j}\right). \tag{5}$$

In the formula, $F = [F_{c,i,j}]_{C\times I\times J}$ is the output matrix of the nonlinear response function, and the size of the matrix is the same as that of the input matrix, and B is the element value in row D and column E on the c-th channel of the matrix.

In the formula, $F = [F_{c,i,j}]_{C\times I\times J}$ represents the output matrix of the nonlinear response function, and the size of the matrix is the same as that of the input matrix, and $F_{c,i,j}$ represents the element value on the c channel of the matrix

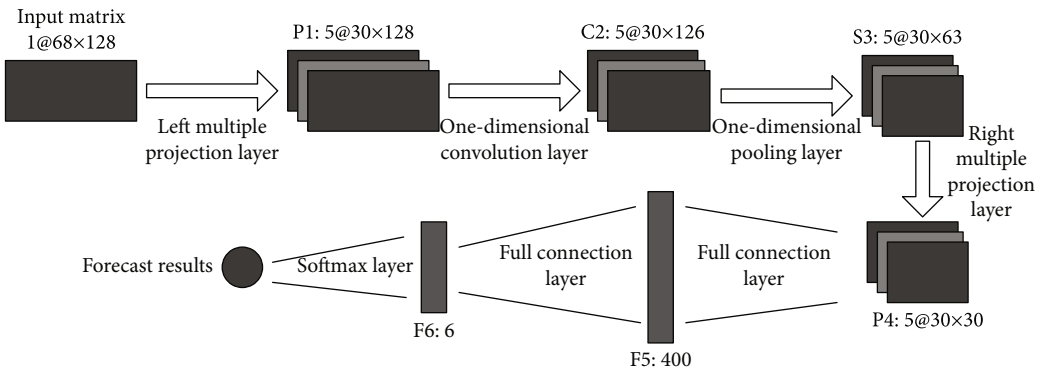


Figure 2: Network structure of BCN.

F that is in column j and row i. The nonlinear activation function is a hyperbolic tangent function, which is represented by $\tanh(\cdot)$, while b_c represents the global offset to the c channel of the input matrix during the nonlinear transformation.

The full connection layer and Softmax layer used in the BCN are the same as those in traditional CNN, where the full connection layer first stretches the nonlinear transformation-after response matrix to vectors and projects it, then finally classifies it through the Softmax layer. When training BCN, the training dataset is first divided into several batches of data and then input into the network in turn. According to the defined loss function, the difference between the prediction probability and the actual category of samples is calculated.

4. Interactive Collaborative Process Based on the Bidirectional LSTM and Attention Mechanism

4.1. Feature Fusion. Feature fusion is used to integrate different types of features to achieve redundancy and get features that are conducive to analysis and processing. Generally, there are two intuitive fusion methods in neural networks: add and concatenate. The add method is the addition of feature maps, which increases the amount of information describing the features of the image, but the dimension describing the image itself does not increase, but the amount of information in each dimension is increasing, which is obviously beneficial for the final image classification. Concatenate is a merge of the number of channels; that is, the features describing the image itself increase, while the information under each feature does not increase. Direct stitching of multilayered information in the network cannot make better use of the complementarity between features, so consider mapping features to multiple subspaces for weighted fusion and stitching them together.

Similar to the general neural network method, the number of subspaces is a superparameter. For each subspace, the corresponding central variables are defined and randomly initialized, and the mapping matrix is also initialized at the same randomness, calculating the adaptive weight based on the distance between the feature and the central variable. As with neural network model parameters, eventually, all mapping matrices and central variables are obtained by training neural networks through BP algorithms.

Assuming that layer n features are extracted, each feature is represented by $f_i (i = 1, 2, \cdots, n)$, the feature dimension is df, the mapping matrix $\mathrm{SW} = [\mathrm{SW}_1, \mathrm{SW}_2, \cdots, \mathrm{SW}_k] \in R^{df \times dk}$ is defined, the features are mapped to a subspace, weighted fusion is performed under the subspace, and the central variable $M = [M_1, M_2, \cdots, M_k] \in R^{t \times dk}$ is defined, where $df = k \times dk$ and k represent the number of subspaces set, and dk represent the feature dimension mapped to a subspace. The adaptive weights are computed, and the subspace features are stitched together. The feature fusion process is as follows:

$$\begin{aligned} \mathrm{sw}_{i,j} &= e^{-(1/df)\left\| F_i \mathrm{SW}_j - M_j \right\|_2^2}, \\ \mathrm{Sf}_j &= \sum_{i=1}^{n} \mathrm{sw}_{i,j} \left(f_i \mathrm{SW}_j \right), \\ \mathrm{SF} &= \mathrm{Concat}(\mathrm{Sf}_1, \mathrm{Sf}_2, \cdots, \mathrm{Sf}_k). \end{aligned} \tag{6}$$

In the formula, $\mathrm{sw}_{i,j}$ represents the weight corresponding to the i feature in the j subspace, Sf_j represents the feature after the j subspace is added and fused by weight, and SF indicates the final feature fusion formed by the stitching of each subspace feature fusion.

4.2. LSTM with the Attention Model. The attention model contains two kinds of attention mechanisms. The calculation process of the soft attention mechanism is subtle, can be easily embedded in a known framework, and can spread the gradient in the attention model [29]. The hard attention mechanism is a random process in which the hidden state is sampled with a certain probability. In order to transfer the gradient, Monte Carlo sampling is usually used to estimate the gradient, which makes the hard attention mechanism not well embedded as a module in known models. Therefore, in order to facilitate model training, the soft attention mechanism is introduced into the LSTM network to form the AM-LSTM (attention mechanism-LSTM) network.

The attention model weights the input vector x_t in each time step to generate a new vector a_t as follows:

$$a_t = \varphi(\omega_{xa} x_t + \omega_{xa} h_{t-1} + b_a). \tag{7}$$

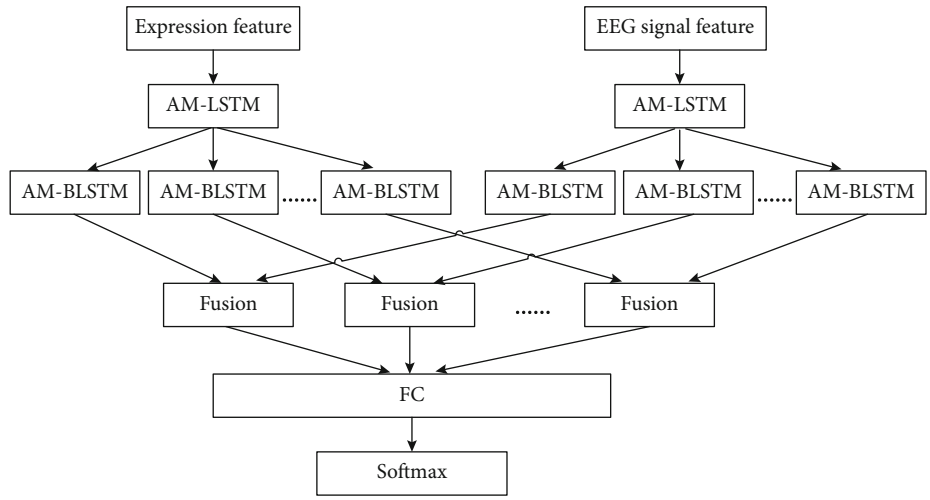


Figure 3: Structure diagram of a three-layer bidirectional LSTM network.

In the formula, φ represents the sigmoid activation function, ω represents the weight matrix, and b_a represents the offset vector. The importance of each element in the input vector x_t is determined by the current input x_t and hidden state h_{t-1}.

The attention response is multiplied by the point of the input vector to output a new input vector $\widehat{x}_t$, which is calculated as follows:

$$\widehat{x}_t = a_t \otimes x_t. \tag{8}$$

The activation function in the LSTM network module then iteratively calculates based on the new input vector element. The computational unit of the recurrent neural network for the attention module is formalized as follows:

$$\begin{aligned} i_t &= \sigma\left(\omega_{xi}\widehat{x}_t + \omega_{hi}h_{t-1} + b_i\right), \\ f_t &= \sigma\left(\omega_{xf}\widehat{x}_t + \omega_{hf}h_{t-1} + b_f\right), \\ o_t &= \sigma\left(\omega_{xo}\widehat{x}_t + \omega_{ho}h_{t-1} + b_o\right), \\ c_t &= f_t \times c_{t-1} + i_t \otimes \Phi\left(\omega_{xc}\widehat{x}_t + \omega_{hc}h_{t-1} + b_c\right), \\ h_t &= o_t \otimes \tanh\left(c_t \circ f_p\right). \end{aligned} \tag{9}$$

i, f, o, and c are the input gate, forget gate, output gates, and memory units. To prevent overfitting, the dropout algorithm is introduced in the training process. In the training process, the parameter of the weight layer is randomly sampled with a certain probability p, and f_p is the vector produced by the Bernoulli distribution with probability p. Most attention mechanisms use Softmax to calculate the weights of each element so that the sum of weights is guaranteed to be 1, but in this case, the attention weights between elements influence each other, even if they are of the same importance. So using the sigmoid activation function, the weights can be normalized between 0 and 1.

4.3. Feature Fusion Based on the Bidirectional Long-Term and Short-Term Memory Network. Feature fusion can be divided into early fusion and delayed fusion. Early fusion cannot learn the time series dynamics of each feature for tasks that require time series modeling. Delayed fusion may lose the time series information that exists in the feature representation of multiple models. To solve the above problem, the proposed method designs a three-layer bidirectional LSTM network to fuse features, as shown in Figure 3.

The first layer of the bidirectional LSTM network uses a separate bidirectional LSTM layer to model each feature in time series. The second layer combines hidden features from the RGB image model and hidden features from the depth image model using linear functions, using the sigmoid function to add nonlinear factors, and generating new feature representations at each time step. The third layer uses the bidirectional LSTM network to model the output of the second layer in time series.

Assuming that x_t^m is the feature representation extracted from RGB or depth models at time t, the three levels of the network are described as follows:

(1) Layer 1. The input of this layer is x_t^m, assuming that Ψ represents a two-way long-term and short-term memory network layer, and h_t^m is the implicit layer representation of t moment.

$$h_t^m = \Psi(x_t^m). \tag{10}$$

(2) Layer 2. This layer uses a linear function to fuse the characteristics of the hidden layer and a sigmoid function to add nonlinear factors.

$$f_t = \sigma\left(\omega\left(h_t^{\mathrm{RGB}}, h_t^{\mathrm{BCN}}\right) + b\right). \tag{11}$$

In the formula, ω is the weight matrix, h_t^{RGB} is the hidden layer of the RGB feature at time t, h_t^{BCN} is the hidden layer of the BCN feature, b is the offset, and σ is the sigmoid function.

(3) Layer 3. This layer uses a bidirectional LSTM layer to model the time series of fused feature f_t.

$$h_t^f = \Psi(f_t). \tag{12}$$

The final output is used as the input to the fully connected layer for final prediction.

4.4. Loss Function Used in Training. The BP (backpropagation) algorithm is commonly used in training neural networks, and the most intuitive function of the loss function is to update the model parameters by calculating its backpropagation gradient. The use of different loss functions tends to make the model more focused on learning some aspects of the data characteristics and can better guarantee the unique characteristics of the extracted features later, so the loss function has a guiding effect in network optimization [30].

Classification networks are generally trained with the cross-entropy loss function, which is as follows:

$$L_S = -\frac{1}{N}\sum_{i=1}^{N}\log\frac{e^{\omega_{yi}^T x_i + b_{yi}}}{\sum_{j=1}^{c} e^{\omega_j^T x_i + b_j}}, \tag{13}$$

where $\omega_j^T x_i + b$ represents the full connection layer output and *yi* represents the real category label corresponding to the input sample. The essence of the loss function decline is to increase the proportion so that the samples of this class fall more within the decision boundary of this class.

However, for facial expression recognition, the cross-entropy loss function cannot guarantee that the distance between classes of extracted features increases. It mainly considers whether the samples can be classified correctly and lacks the constraint of the distance between classes and within classes [31]. Generally, the problem of the loss function in face recognition is solved in two ways. On the one hand, it combines the measure learning method; on the other hand, it is improved on the basis of this Softmax cross-entropy loss function. Add an additional loss function after the cross-entropy loss function to strengthen the constraint on the interclass distance, making the interclass feature distance of the same sample become compact but not sufficiently constrained between classes [32, 33]. The cross-entropy loss function is improved directly by normalizing ω and x and converting them to cosine distances.

$$L = -\frac{1}{N}\sum_{i=1}^{N}\log\frac{e^{\cos\theta_{yi}}}{\sum_{j=1}^{c} e^{\cos\theta_j}}. \tag{14}$$

Then, introduce the interval constraint and use the AM-Softmax method to calculate the following:

$$L = -\frac{1}{N}\sum_{i=1}^{N}\log\frac{e^{s(\cos\theta_{yi}-m)}}{e^{s(\cos\theta_{yi}+m)} + \sum_{j=1,j\neq yi}^{c} e^{s\cos\theta_j}}. \tag{15}$$

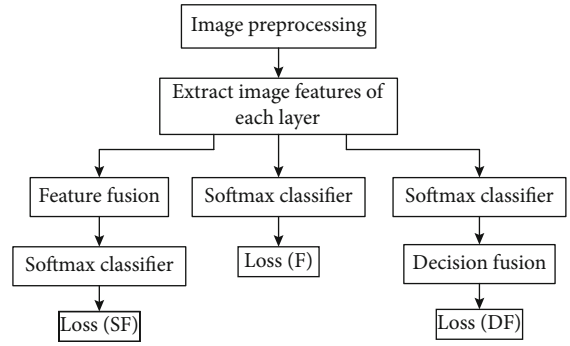


Figure 4: Calculation process of the loss function.

For all features, feed into the Softmax classifier, using the above loss function, as follows:

$$\text{loss}(F) = \sum_{i=1}^{n} L(f_i). \tag{16}$$

On this basis, the loss of the fused part is added, and the loss function calculation process is shown in Figure 4.

For the feature fusion part, the fused features are fed into the Softmax classifier, and loss(SF) is used to represent the above loss function. For the decision fusion section, loss(DF) is used to represent the above loss function [34, 35]. Add all the loss functions as the final loss function, in the following form:

$$\text{Loss} = \text{loss}(F) + \text{loss}(\text{SF}) + \text{loss}(\text{DF}). \tag{17}$$

5. Experimental Scheme and Result Discussion

To verify the validity of the proposed method, the MATLAB simulation platform experimented on the MAHNOB-HCI dataset and the DEAP dataset and used RA (recognition accuracy) and F1-score as evaluation indicators for recognition. In addition, the two-dimensional arousal-valence emotional space is used in the experiment, and its principle is shown in Figure 5.

In the figure, the horizontal valence represents the validity, indicating the positive and negative degrees of emotion, and the vertical arousal represents the activation degree, indicating the intensity and depression of emotion. By setting validity and wakefulness, complex and subtle emotions can be expressed and distinguished, such as ecstasy and cheerfulness, which describe varying degrees of pleasure and elation, which express two different kinds of joy. The two-dimensional affective space has become the main dimensional space used for dimension affective recognition because of its simple structure and rich emotional expression ability.

The MAHNOB-HCI dataset is a multimodal affective recognition and latent labeling dataset, which includes 527 sets of original face videos and audio and electroencephalogram signals collected from 27 participants who watched 20 videos. After watching each video, the participants used the activation and effect value of the calibration emotions, which were divided into 9 levels (1-9, respectively). Emotions are

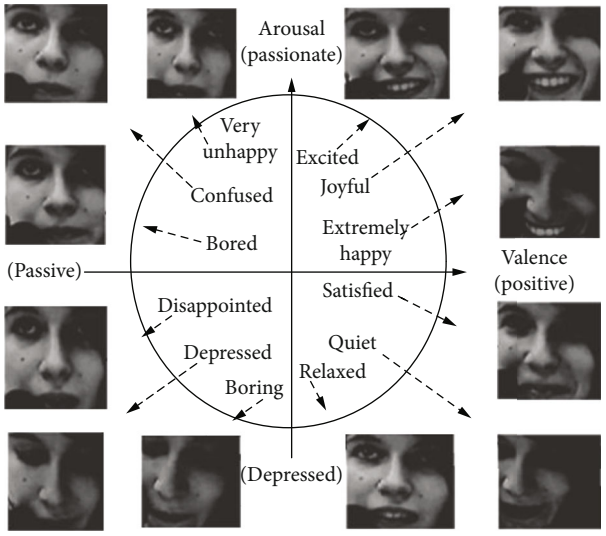


Figure 5: Schematic diagram of the two-dimensional arousal-valence emotional state space.

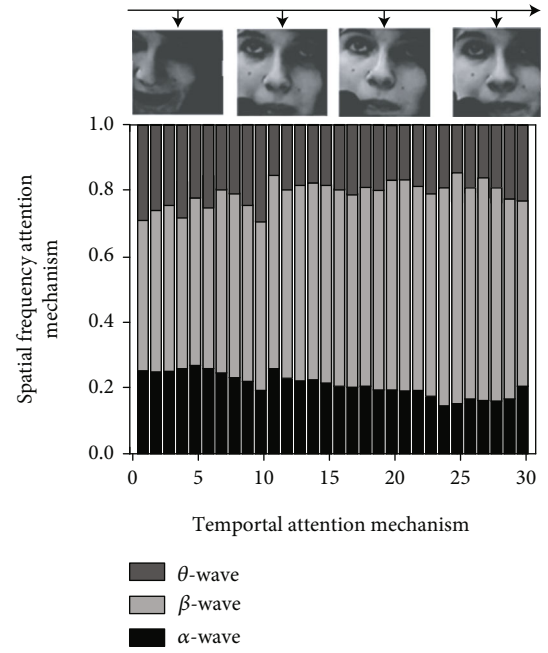


Figure 6: Visualization results of the attention mechanism in high activation data samples.

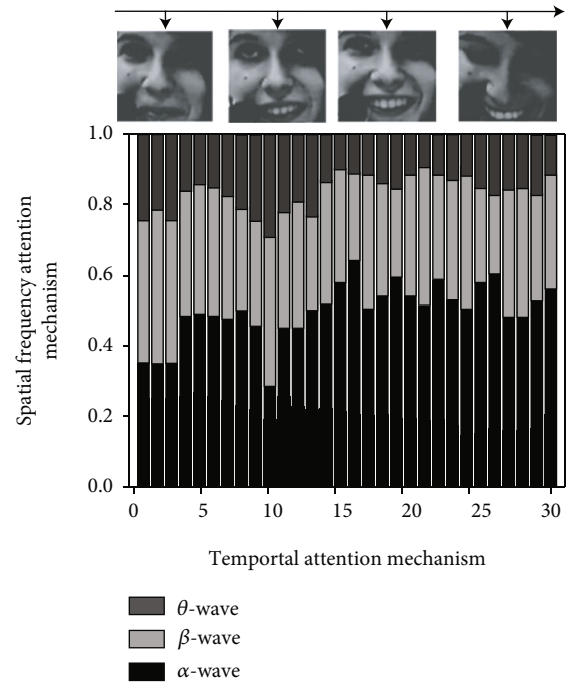


Figure 7: Visualization results of the attention mechanism in low activation data samples.

also calibrated using discrete affective tags, which categorized the activation and validity of experimenters' emotions into three categories.

DEAP is a multimodal emotion recognition dataset that includes face videos, external physiological signals, and electroencephalogram signals collected from 32 participants during 40 music videos. The data of 10 participants do not include facial expression videos. After viewing each video, the participants used the activation and validity of the calibrated emotions (values 1-9). The dataset divides emotional activation and validity into three levels according to their numerical size.

The results of model recognition are evaluated using two indicators: recognition accuracy and F1-score, where recognition accuracy RA represents the percentage of the number of samples correctly classified in the test set to the total number of samples in the test set. F1-score is a statistical measure of the accuracy of a multiclassification model. It can be viewed as a weighted average of model accuracy (precision) and recall rate (recall), which can take into account both the model accuracy and the recall rate. Recognition accuracy and F1-score are calculated as follows:

$$\begin{aligned} \mathrm{RA} &= \frac{N_{\mathrm{TP}}}{N_{\mathrm{data}}}, \\ \mathrm{F1} &= \frac{2N_{\mathrm{TP}}}{2N_{\mathrm{TP}} + N_{\mathrm{FP}} + N_{\mathrm{FN}}}. \end{aligned} \tag{18}$$

In the formula, N_{data} represents the total number of samples of emotion data in the test set; N_{TP}, N_{FP}, and N_{FN} represent the total number of positive, false, and missed tests in all the test samples, respectively.

5.1. Visual Effect Comparison of the Attention Mechanism. Select samples from the MAHNOB-HCI dataset for a single set of emotion recognition tests and visualize the emotional key information on each step to Figures 6 and 7. Only four time steps are selected to visualize the time domain attention mechanism. The top bar in the graph represents the importance of three bands of EEG signals α, β, and θ at each time from bottom to top.

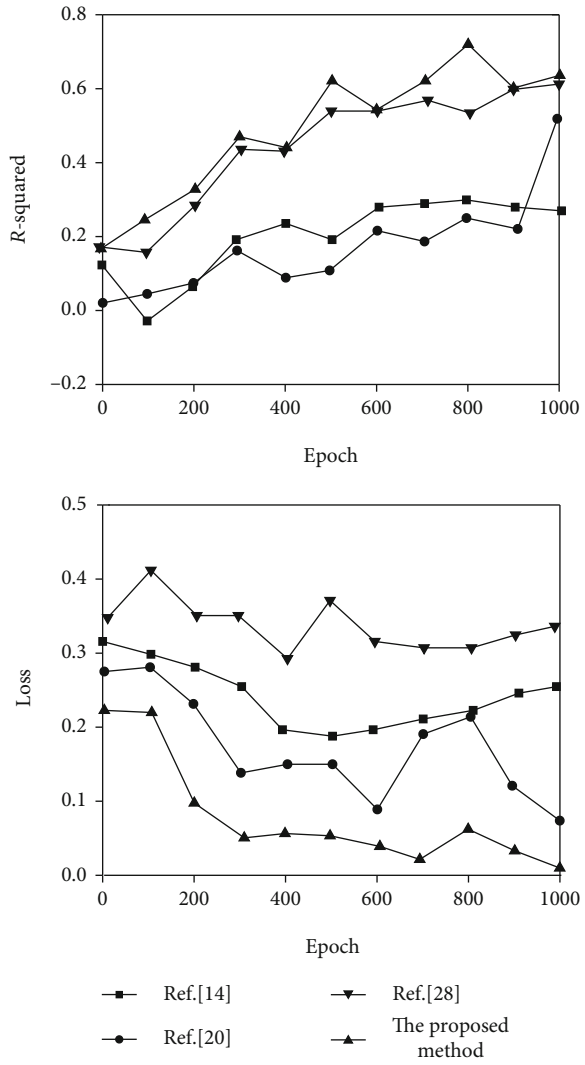


Figure 8: Arousal dimension video emotion salient feature learning results.

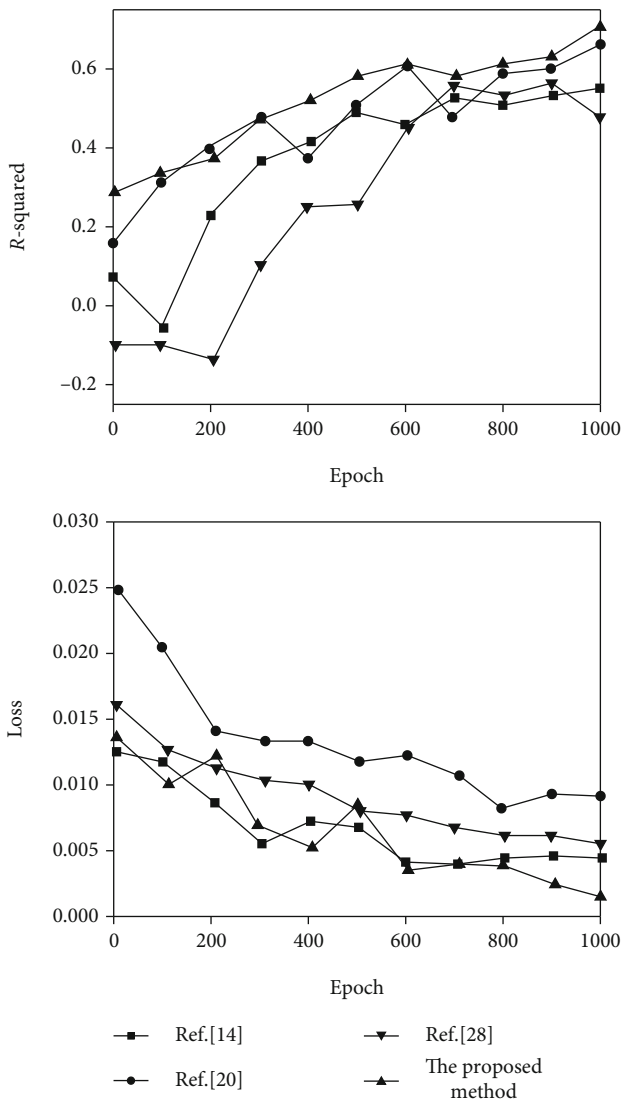


Figure 9: Valence dimension video emotion salient feature learning results.

As you can see from the above two figures, the proposed method can locate valid information quickly and accurately and analyze the data more frequently in areas of valid information to obtain more accurate identification results. As can be seen from Figure 6, as the participants become more nervous (and more emotionally active), the EEG β-wave gradually dominates, which is consistent with the conclusion that the EEG β-wave in physiology dominates when humans are in an emotional state of stress, anxiety, panic, and so on. With the enhancement of the EEG β-wave, the human body will be more and more in a tense state. Under this condition, the human body and mind energy expenditure is rapid, and it is easy to feel pressure and fatigue. As can be seen from Figure 7, α-wave dominates when the human emotional state remains stable at a low activation level. Physiological studies have shown that when the main frequency of the human brain wave is α-wave, the human being is awake and relaxed, and it is also the best state for human thinking.

5.2. Comparison of the Learning Effect on Facial Emotion Significant Features. In the learning phase of facial expression features, four methods (proposed method and References [14, 20, 28]) were used to compare the results of feature learning based on the DEAP dataset in the two dimensions of arousal and valence, respectively, as shown in Figures 8 and 9. The R-squared coefficient and loss function Loss are used as the evaluation indicators of feature learning. The R-squared coefficient of the regression task represents the degree of fit between the predicted value and the label value by calculating the change of the data. The larger the value of the R-squared coefficient, the better the fit degree and the better the feature extraction effect. The R-squared coefficient function is as follows:

$$R^2 = 1 - \frac{\sum \left(Y_{\text{act}} - Y_{\text{pre}}\right)^2}{\sum \left(Y_{\text{act}} - \bar{Y}\right)^2}. \tag{19}$$

Table 1: Comparison of dimensional emotion recognition results of different methods.

	Arousal		Valence	
	R^2	Loss	R^2	Loss
Ref. [16]	0.442	0.263	0.477	0.174
Ref. [19]	0.535	0.181	0.469	0.115
Ref. [20]	0.687	0.152	0.642	0.338
The proposed method	0.749	0.174	0.706	0.127

In the formula, Y_{act} is the emotional reality tag sequence, Y_{pre} is the emotional prediction value sequence, and $\bar{Y}$ is the average value of the emotional reality tag sequence.

From Figure 8, it can be seen that compared with other methods, the proposed method has the highest R-squared coefficient on the arousal dimension, and the loss is very low, which is close to 0.05, indicating that the feature extraction effect is the best. As can be seen from Figure 9, the squared coefficient of the proposed method is 0.63, the loss is very close to 0.002, and the network structure is simpler than several other methods. Therefore, the proposed method achieves better feature extraction results in the two-dimensional arousal-valence emotional state space.

5.3. Compared with the Dimensional Emotion Recognition Results of Other Methods. The two-dimensional emotional spatial expression was used to compare the results of dimensional affective recognition with those of References [16, 19, 20] in the MAHNOB-HCI dataset, as shown in Table 1.

As can be seen from Table 1, although the proposed method is inferior to the optimal method in terms of loss, the correlation coefficient better reflects the degree of fit between the emotional prediction value and the emotional tag value. Moreover, the recognition accuracy of the proposed method has exceeded the results of many methods, and the recognition accuracy of the hierarchical attention mechanism built on this basis performs best. Finally, the loss is optimized and the accuracy of the recognition rate is 0.749 and 0.706 in two dimensions, respectively, indicating that the proposed method can more effectively extract the emotional characteristics of the emoticon and the brain electrical signal to combine.

5.4. Comparison of Results from Multiple Emotion Recognition Methods. To demonstrate the performance of the proposed method, the results of this paper are compared with those of References [16, 19, 20], as shown in Table 2.

As can be seen from Table 2, the recognition accuracy and F1 of the proposed method are significantly improved compared with those of other methods. Because other methods directly analyze multimodal emotional signals that contain a lot of redundant information, the proposed method introduces the attention mechanism, which compresses the redundant information and improves the accuracy. On the MAHNOB-HCI dataset, the accuracy of affective activation recognition and F1 were improved by 1.1% and 0.021, respectively, and the accuracy of affective efficacy value recognition and F1 were improved by 0.4% and 0.024, respectively, compared with the methods in Reference [19], which showed better recognition results. And the recognition accuracy on the DEAP dataset is higher than that on the MAHNOB-HCI dataset because the DEAP dataset has more information, which is conducive to feature learning and classification.

Table 2: Comparison of recognition results of different methods.

	Arousal		Valence	
	RA (%)	F1	RA (%)	F1
Ref. [16]/MAHNOB-HCI	67.8	0.620	74.8	0.742
Ref. [19]/MAHNOB-HCI	73.5	0.711	74.9	0.715
Ref. [20]/MAHNOB-HCI	63.4	—	66.7	—
The proposed method/ MAHNOB-HCI	74.6	0.732	75.3	0.739
The proposed method/DEAP	86.8	—	86.2	—

In addition, it can be seen from Table 2 that the recognition of the emotional effect value is better than that of emotional activation. This is because the degree of emotional activation is used to show the degree of emotional motivation, and the value of the emotional effect is used to show whether people are good or bad in evaluating their emotional state. Compared with the value of the emotional effect, it is easier to analyze and understand intuitively. In particular, FAWT in Reference [16] decomposes EEG signals into different subband signals on the MANOB-HCI dataset, extracts features from them using the information potential, and smoothes them into random forests and SVM classifiers that classify emotions. It achieves good results in validity value recognition. Overall, the method proposed is more effective.

6. Conclusion

At present, most emotion recognition focuses on speech signals, facial expressions, ECG, EEG, and other bioelectrical signals. However, when the emotional signal of a single channel is interfered by other signals, the emotion recognition rate is often reduced. Therefore, a multimodal emotion recognition method based on the bidirectional LSTM and attention mechanism is proposed. Facial expression and EEG features are extracted based on BCN, and the attention mechanism is introduced into the LSTM network. Facial expression and EEG features are fused by a feature fusion network with a three-layer bidirectional LSTM structure to improve the accuracy of emotion recognition. The proposed method is tested on MAHNOB-HCI and DEAP datasets based on the MATLAB simulation platform. The experimental results show that the attention mechanism can enhance the visual effect of the image, and compared with other methods, the proposed method can extract emotion features in expressions and EEG signals more effectively and achieve a more accurate emotion recognition effect.

In real life, different emotions often have a certain correlation; for example, sad emotions often contain certain anger. Therefore, the emotion recognition system based on the relationship between emotions can reflect the human emotional information more effectively. So, the establishment of an effective emotion recognition model will be the focus of the next step of emotion recognition research.

Data Availability

The data included in this paper are available without any restriction.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

References

[1] K. Gabriels, "Response to uncertainty in emotion recognition," *Journal of Information Communication and Ethics in Society*, vol. 17, no. 3, pp. 295–298, 2019.

[2] L. Singh, S. Singh, and N. Aggarwal, "Improved TOPSIS method for peak frame selection in audio-video human emotion recognition," *Multimedia Tools and Applications*, vol. 78, no. 5, pp. 6277–6308, 2019.

[3] T. Ozseven, "A novel feature selection method for speech emotion recognition," *Applied Acoustics*, vol. 146, no. 3, pp. 320–326, 2019.

[4] P. Song, W. Zheng, and L. Zhao, "Joint subspace learning and feature selection method for speech emotion recognition," *Qinghua Daxue Xuebao/Journal of Tsinghua University*, vol. 58, no. 4, pp. 347–351, 2018.

[5] J. Li, S. Qiu, Y.-Y. Shen, C.-L. Liu, and H. He, "Multisource transfer learning for cross-subject EEG emotion recognition," *IEEE Transactions on Cybernetics*, vol. 50, no. 7, pp. 3281–3293, 2020.

[6] Z. T. Liu, Q. Xie, M. Wu, W. H. Cao, Y. Mei, and J. W. Mao, "Speech emotion recognition based on an improved brain emotion learning model," *Neurocomputing*, vol. 309, no. 10, pp. 145–156, 2018.

[7] Y. L. Hsu, J. S. Wang, W. C. Chiang, and C. H. Hung, "Automatic ECG-based emotion recognition in music listening," *IEEE Transactions on Affective Computing*, vol. 11, no. 1, pp. 85–99, 2020.

[8] A. Raheel, S. M. Anwar, and M. Majid, "Emotion recognition in response to traditional and tactile enhanced multimedia using electroencephalography," *Multimedia Tools and Applications*, vol. 78, no. 10, pp. 13971–13985, 2019.

[9] Y. Dong, X. Yang, X. Zhao, and J. Li, "Bidirectional convolutional recurrent sparse network (BCRSN): an efficient model for music emotion recognition," *IEEE Transactions on Multimedia*, vol. 21, no. 12, pp. 3150–3163, 2019.

[10] Y. Huang, K. Tian, A. Wu, and G. Zhang, "Feature fusion methods research based on deep belief networks for speech emotion recognition under noise condition," *Journal of Ambient Intelligence and Humanized Computing*, vol. 10, no. 5, pp. 1787–1798, 2019.

[11] Y. Li, W. Zheng, Z. Cui, Y. Zong, and S. Ge, "EEG emotion recognition based on graph regularized sparse linear regression," *Neural Processing Letters*, vol. 49, no. 2, pp. 555–571, 2019.

[12] C. Deepika, "Speech emotion recognition feature extraction and classification," *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 9, no. 2, pp. 1257–1261, 2020.

[13] P. Wei and Y. Zhao, "A novel speech emotion recognition algorithm based on wavelet kernel sparse classifier in stacked deep auto-encoder model," *Personal and Ubiquitous Computing*, vol. 23, no. 3-4, pp. 521–529, 2019.

[14] T. Zhang, W. Zheng, Z. Cui, Y. Zong, and Y. Li, "Spatial–temporal recurrent neural network for emotion recognition," *IEEE Transactions on Cybernetics*, vol. 49, no. 3, pp. 839–847, 2019.

[15] S. B. Wankhade and D. D. Doye, "Deep learning of empirical mean curve decomposition-wavelet decomposed EEG signal for emotion recognition," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 28, no. 1, pp. 153–177, 2020.

[16] V. Gupta, M. D. Chopda, and R. B. Pachori, "Cross-subject emotion recognition using flexible analytic wavelet transform from EEG signals," *IEEE Sensors Journal*, vol. 19, no. 6, pp. 2266–2274, 2019.

[17] N. Jain, S. Kumar, A. Kumar, P. Shamsolmoali, and M. Zareapoor, "Hybrid deep neural networks for face emotion recognition," *Pattern Recognition Letters*, vol. 115, no. 11, pp. 101–106, 2018.

[18] S. Hwang, M. Ki, K. Hong, and H. Byun, "Subject-independent EEG-based emotion recognition using adversarial learning," in *2020 8th International Winter Conference on Brain-Computer Interface (BCI)*, Gangwon, Korea (South), February 2020.

[19] J. Ma, Y. Sun, and X. Zhang, "Multimodal emotion recognition for the fusion of speech and EEG signals," *Xi'an Dianzi Keji Daxue Xuebao/Journal of Xidian University*, vol. 46, no. 1, pp. 143–150, 2019.

[20] F. Rahdari, E. Rashedi, and M. Eftekhari, "A multimodal emotion recognition system using facial landmark analysis," *Iranian Journal of ence and Technology, Transactions of Electrical Engineering*, vol. 43, no. S1, pp. 171–189, 2019.

[21] S. Zhang, S. Zhang, T. Huang, W. Gao, and Q. Tian, "Learning affective features with a hybrid deep model for audio–visual emotion recognition," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 28, no. 10, pp. 3030–3043, 2018.

[22] M. Moghimi, R. Stone, and P. Rotshtein, "Affective recognition in dynamic and interactive virtual environments," *IEEE Transactions on Affective Computing*, vol. 11, no. 1, pp. 45–62, 2020.

[23] C. I. Wasser, F. Evans, C. Kempnich et al., "Emotion recognition in Parkinson's disease: static and dynamic factors," *Neuropsychology*, vol. 32, no. 2, pp. 230–234, 2018.

[24] W. C. Fang, K. Y. Wang, N. Fahier, Y. L. Ho, and Y. D. Huang, "Development and validation of an EEG-based real-time emotion recognition system using edge AI computing platform with convolutional neural network system-on-chip design," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 9, no. 4, pp. 645–657, 2019.

[25] C. Wu, C. Huang, and H. Chen, "Text-independent speech emotion recognition using frequency adaptive features," *Multimedia Tools & Applications*, vol. 77, no. 18, pp. 24353–24363, 2018.

[26] C. V. Haritha and P. P. Thulasidharan, "Multimodal emotion recognition using deep neural network- a survey," *INTERNATIONAL JOURNAL OF COMPUTER ENCES AND ENGINEERING*, vol. 6, no. 6, pp. 95–98, 2018.

[27] L. Ye, P. Wang, H. F. Le Wang, T. Seppänen, and E. Alasaarela, "A combined motion-audio school bullying detection algorithm," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 32, no. 12, p. 1850046, 2018.

[28] A. Saha and S. N. Pradhan, "Facial expression recognition based on eigenspaces and principle component analysis,"

International Journal of Computational Vision & Robotics, vol. 8, no. 2, pp. 190–200, 2018.

[29] H. Feng, H. M. Golshan, and M. H. Mahoor, "A wavelet-based approach to emotion classification using EDA signals," *Expert Systems with Applications*, vol. 112, pp. 77–86, 2018.

[30] R. Pal, "Emotion recognition using facial expression," *International Journal on Computer ence & Engineering*, vol. 2, no. 5, pp. 2220–2224, 2018.

[31] H.-D. Nguyen, S. Yeom, G.-S. Lee, H.-J. Yang, I.-S. Na, and S.-H. Kim, "Facial emotion recognition using an ensemble of multi-level convolutional neural networks," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 33, no. 11, 2019.

[32] F. Athishmon, N. Narayanan, and K. Suthendran, "Recognizing spontaneous emotion from the eye region under different head poses," *International Journal of Pure and Applied Mathematics*, vol. 118, no. 8, pp. 257–262, 2018.

[33] M. Hu, H. Wang, X. Wang, J. Yang, and R. Wang, "Video facial emotion recognition based on local enhanced motion history image and CNN-CTSLSTM networks," *Journal of Visual Communication & Image Representation*, vol. 59, no. 2, pp. 176–185, 2019.

[34] H. Bo, L. Ma, Q. Liu, R. Xu, and H. Li, "Music-evoked emotion recognition based on cognitive principles inspired EEG temporal and spectral features," *International Journal of Machine Learning and Cybernetics*, vol. 10, no. 9, pp. 2439–2448, 2019.

[35] Y. Yang, Q. M. J. Wu, W. L. Zheng, and B. L. Lu, "EEG-based emotion recognition using hierarchical network with subnetwork nodes," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 10, no. 2, pp. 408–419, 2018.

Hindawi
Computational and Mathematical Methods in Medicine
Volume 2023, Article ID 9859656, 1 page
https://doi.org/10.1155/2023/9859656

Retraction

Retracted: PSSPNN: PatchShuffle Stochastic Pooling Neural Network for an Explainable Diagnosis of COVID-19 with Multiple-Way Data Augmentation

Computational and Mathematical Methods in Medicine

Received 26 September 2023; Accepted 26 September 2023; Published 27 September 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] S. Wang, Y. Zhang, X. Cheng, X. Zhang, and Y. Zhang, "PSSPNN: PatchShuffle Stochastic Pooling Neural Network for an Explainable Diagnosis of COVID-19 with Multiple-Way Data Augmentation," *Computational and Mathematical Methods in Medicine*, vol. 2021, Article ID 6633755, 18 pages, 2021.

Hindawi
Computational and Mathematical Methods in Medicine
Volume 2021, Article ID 6633755, 18 pages
https://doi.org/10.1155/2021/6633755

Research Article

PSSPNN: PatchShuffle Stochastic Pooling Neural Network for an Explainable Diagnosis of COVID-19 with Multiple-Way Data Augmentation

Shui-Hua Wang,[1,2] **Yin Zhang**,[3] **Xiaochun Cheng**,[4] **Xin Zhang**,[5] **and Yu-Dong Zhang**[6]

[1]*School of Computer Science, Henan Polytechnic University, China, Henan 454001, China*
[2]*School of Architecture Building and Civil Engineering, Loughborough University, Loughborough LE11 3TU, UK*
[3]*School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China*
[4]*School of Science & Technology, Middlesex University, London NW4 4BT, UK*
[5]*Department of Medical Imaging, The Fourth People's Hospital of Huai'an, Huai'an, Jiangsu Province 223002, China*
[6]*School of Informatics, University of Leicester, Leicester LE1 7RH, UK*

Correspondence should be addressed to Xiaochun Cheng; x.cheng@mdx.ac.uk, Xin Zhang; 973306782@qq.com, and Yu-Dong Zhang; yudongzhang@ieee.org

Received 9 November 2020; Revised 23 December 2020; Accepted 18 February 2021; Published 10 March 2021

Academic Editor: Martti Juhola

Aim. COVID-19 has caused large death tolls all over the world. Accurate diagnosis is of significant importance for early treatment. *Methods*. In this study, we proposed a novel PSSPNN model for classification between COVID-19, secondary pulmonary tuberculosis, community-captured pneumonia, and healthy subjects. PSSPNN entails five improvements: we first proposed the n-conv stochastic pooling module. Second, a novel stochastic pooling neural network was proposed. Third, PatchShuffle was introduced as a regularization term. Fourth, an improved multiple-way data augmentation was used. Fifth, Grad-CAM was utilized to interpret our AI model. *Results*. The 10 runs with random seed on the test set showed our algorithm achieved a microaveraged F1 score of 95.79%. Moreover, our method is better than nine state-of-the-art approaches. *Conclusion*. This proposed PSSPNN will help assist radiologists to make diagnosis more quickly and accurately on COVID-19 cases.

1. Introduction

COVID-19 is a type of disease caused by a new strain of coronavirus. "CO" means corona, "VI" virus, and "D" disease. Up to 22 December 2020, COVID-19 has caused more than 78.0 million confirmed cases and over 1.71 million deaths (US 326.5 k, Brazil 188.2 k, India 146.1 k, Mexico 119.4 k, Italy 69.8 k, and UK 68.3 k).

To diagnose COVID-19, two methods exist: (i) real-time reverse transcriptase PCR with nasopharyngeal swab samples to test the existence of RNA fragments and (ii) chest imaging (CI) examines the evidence of COVID-19. The first type of rRT-PCR method needs to wait for a few days to get the results, while the second type of CI approach could get quick results within minutes. The CI approaches have several advantages compared to rRT-PCR. First, the swab may be contaminated [1, 2]. Second, CI can detect the lesions of lungs where "ground-glass opacity (GGO)" will be observed to distinguish COVID-19 from healthy subjects. Third, CI can provide an immediate result as soon as imaging is complete. Fourth, reports show that chest computed tomography (CCT), one type of CI approach, can detect 97% of COVID-19 infections [3].

Currently, there are three types of CI approaches: chest X-ray (CXR) [4], chest CT (CCT), and chest ultrasound (CUS) [5]. Among all three types of approaches, CCT can provide the finest resolution than the other two CI approaches,

allowing visualization of extremely small nodules in the lung. The additional advantage of CCT is that it can provide high-quality, three-dimensional chest data where radiologists can clearly view the COVID-19 lesions, which may be obscure in the other two CI approaches.

However, manual labeling by human experts is tedious, labor-intensive, and time-consuming [6]. Besides, the labeling performances are easily affected by interexpert and intraexpert factors (e.g., emotion, tiredness, and lethargy). Moreover, the labeling throughputs of radiologists are not comparable with artificial intelligence (AI) models. For example, senior radiologists may diagnose one scanning within five minutes, but AI can analyze thousands of samples within one minute. Particularly, early lesions are small and similar to surrounding normal tissues, which make them more challenging to measure and hence can potentially be neglected by radiologists.

Traditional AI methods have successfully been applied in many medical fields. For instance, Wang et al. [7] chose wavelet packet Tsallis entropy as a feature descriptor and employed a real-coded biogeography-based optimization (RCBO) classifier. Jiang and Zhang [8] proposed a 6-level convolutional neural network (6L-CNN) for therapy and rehabilitation. Their performances were improved by replacing the traditional rectified linear unit with a leaky rectified linear unit. Fulton et al. [9] used ResNet-50 (RN-50) to classify Alzheimer's disease with and without imagery. The authors found that ResNet-50 models help identify AD patients. Guo and Du [10] utilized a ResNet-18 (RN-18) model to recognize thyroid ultrasound standard plane (TUSP), achieving a classification accuracy of 83.88%. The experiments verified the effectiveness of RN-18. The aforementioned four algorithms can be transferred to the multiclass classification task of COVID-19 diagnosis.

On the COVID-19 datasets, several recent publications reported promising results. For example, Cohen et al. [11] proposed a COVID severity score network (CSSNet), which achieved a mean absolute error (MAE) of 1.14 on geographic extent score and an MAE of 0.78 on lung opacity score. Li et al. [12] developed a fully automatic model (COVNet) to detect COVID-19 using chest CT and evaluated its performance. Wang et al. [13] proposed a 3D deep convolutional neural network to detect COVID-19 (DeCovNet). Zhang et al. [14] proposed a seven-layer convolutional neural network for COVID-19 diagnosis (7L-CCD). Their performance achieved an accuracy of 94.03 ± 0.80 for the binary classification task (COVID-19 against healthy subjects). Ko et al. [15] proposed a fast-track COVID-19 classification network (FCONet). For the sake of the page limit, the details of those methods are not described, but we shall compare our method with those state-of-the-art methods in the following sections. Wang et al. [16] presented a CCSHNet via transfer learning and discriminant correlation analysis.

Our study's inspiration is to improve recognition performances of COVID-19 infection in CCT images by developing a novel deep neural network, PSSPNN, short for PatchShuffle stochastic pooling neural network. Our contributions entail the following five angles:

(i) The "n-conv stochastic pooling module (NCSPM)" is proposed, which comprises n-times repetitions of convolution layers and batch normalization layers, followed by stochastic pooling

(ii) A novel "stochastic pooling neural network (SPNN)" is proposed, the structure of which is inspired by VGG-16

(iii) A more advanced neural network, PatchShuffle SPNN (PSSPNN), is proposed where PatchShuffle is introduced as the regularization term in the loss function of SPNN

(iv) An improved multiple-way data augmentation is utilized to help the network avoid overfitting

(v) Grad-CAM is used to show the explainable heatmap, which displays association with lung lesions

2. Dataset

This retrospective study was exempt by Institutional Review Board of local hospitals. Four types of CCT were used: (i) COVID-19-positive patients, (ii) community-acquired pneumonia (CAP), (iii) second pulmonary tuberculosis (SPT), and (iv) healthy control (HC). Three diseased classes (COVID-19, CAP, and SPT) were chosen since they are all infectious diseases of the chest regions. We intend to include the fifth category (chest tumors) in our future studies.

For each subject, $n(k)$ slices of CCT were chosen via a slice level selection (SLS) method. For the three diseased groups (COVID-19, CAP, and SPT represented as $k = \{1, 2, 3\}$), the slice displaying the largest number of lesions and size was chosen. For HC subjects ($k = 4$), any slice within the 3D image was randomly chosen. The slice-to-subject ratio $\bar{n}$ is defined as

$$\bar{n}(k) = \frac{N_S(k)}{N_P(k)}, k = 1, \cdots, 4, \tag{1}$$

where N_S stands for the number of slices via the SLS method and N_P is the number of patients.

In all, we enrolled 521 subjects and produced 1164 slice images using the SLS method. Table 1 lists the demographics of the four-category subject cohort with the values of triplets $[\bar{n}, N_P, N_S]$, where $\bar{n}$ of the total set equals to 2.23.

Three experienced radiologists (two juniors: $\mathscr{C}_1$ and $\mathscr{C}_2$ and one senior: $\mathscr{C}_3$) were convened to curate all the images. Suppose x^{CCT} means one CCT scan, Y stands for the labeling of each individual radiologist. The last labeling Y^F of the CCT scan x^{CCT} is obtained by

$$Y^F\left[x^{\mathrm{CCT}}\right] = \begin{cases} Y\left[\mathscr{C}_1, x^{\mathrm{CCT}}\right] & Y\left[\mathscr{C}_1, x^{\mathrm{CCT}}\right] == Y\left[\mathscr{C}_2, x^{\mathrm{CCT}}\right] \\ \mathrm{MAV}\left\{Y^{\mathrm{all}}\left[x^{\mathrm{CCT}}\right]\right\} & \text{otherwise} \end{cases}, \tag{2}$$

Table 1: Subjects and images of four categories.

Class index	Class name	$\bar{n}$	N_P	N_S
1	COVID-19	2.27	125	284
2	CAP	2.28	123	281
3	SPT	2.18	134	293
4	HC	2.20	139	306
Total		2.23	521	1164

where MAV denotes majority voting and Y^{all} the labeling of all radiologists, *viz.*,

$$Y^{\text{all}}\left[x^{\text{CCT}}\right] = \left[Y\left(\mathscr{C}_1, x^{\text{CCT}}\right), Y\left(\mathscr{C}_2, x^{\text{CCT}}\right), Y\left(\mathscr{C}_3, x^{\text{CCT}}\right)\right]. \tag{3}$$

The above two equations mean that in situations of disagreement between the analyses of two junior radiologists $(\mathscr{C}_1, \mathscr{C}_2)$, we consult a senior radiologist $(\mathscr{C}_3)$ to reach a MAV-type consensus. Data is available on request due to privacy/ethical restrictions.

3. Methodology

Table 2 gives the abbreviation and full meanings in this study. Section 3.1 shows the preprocessing procedure. Sections 3.2–3.5 offer four improvements. Finally, Section 3.6 gives the implementation, measure indicators, and explainable technology used in our method.

3.1. Preprocessing. The original raw dataset contained $|V|$ slice images $\{v_a(i), i = 1, 2, \cdots, |V|\}$. The size of each image was size$[v_a(i)] = 1024 \times 1024 \times 3$. Figure 1 presents the pipeline for preprocessing of this dataset.

First, the color CCT images of four classes were converted into grayscale by retaining the luminance channel and obtaining the grayscale data set V_B:

$$\begin{aligned} V_B &= O_{\text{gray}}(V_A) \\ &= \{v_b(1), v_b(2), \cdots, v_b(i), \cdots v_b(|V|)\}, \end{aligned} \tag{4}$$

where O_{gray} means the grayscale operation.

In the second step, the histogram stretching (HS) was utilized to increase the contrast of all images. Take the i-th image $v_b(i), i = 1, 2, \cdots, |V|$ as an example; its minimum and maximum grayscale values $v_b^l(i)$ and $v_b^h(i)$ were calculated as follows:

$$\begin{cases} v_b^l(i) = \min_{w=1}^{W_B} \min_{h=1}^{H_B} \min_{c=1}^{C_B} v_b(i \mid w, h, c), \\ v_b^h(i) = \max_{w=1}^{W_B} \max_{h=1}^{H_B} \max_{c=1}^{C_B} v_b(i \mid w, h, c). \end{cases} \tag{5}$$

Here, (w, h, c) means the index of width, height, and channel directions along image $v_b(i)$, respectively. (W_B, H_B, C_B) means the maximum values of width, height, and channel

Table 2: Abbreviation and full name.

Abbreviation	Full name
AP	Average pooling
CAP	Community-acquired pneumonia
CCT	Chest computed tomography
CI	Chest imaging
CXR	Chest X-ray
CUS	Chest ultrasound
DPD	Discrete probability distribution
FCL	Fully connected layer
FM	Feature map
FMS	Feature map size
HS	Hyperparameter size
L2P	l2-norm pooling
MA	Microaveraged
MAV	Majority voting
MDA	Multiple-way data augmentation
MP	Max pooling
NCSPM	n-conv stochastic pooling module
NWL	Number of weighted layers
SAPN	Salt-and-pepper noise
SC	Strided convolution
SLS	Slice level selection
SP	Stochastic pooling
SPNN	Stochastic pooling neural network
SPT	Second pulmonary tuberculosis
TCM	Test confusion matrix

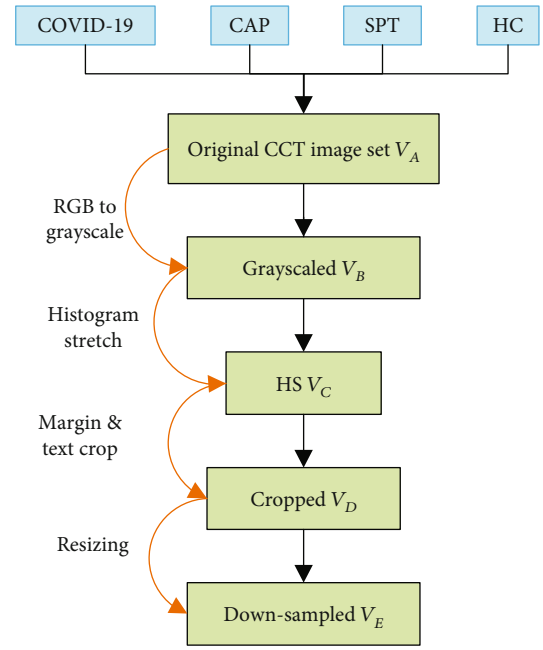
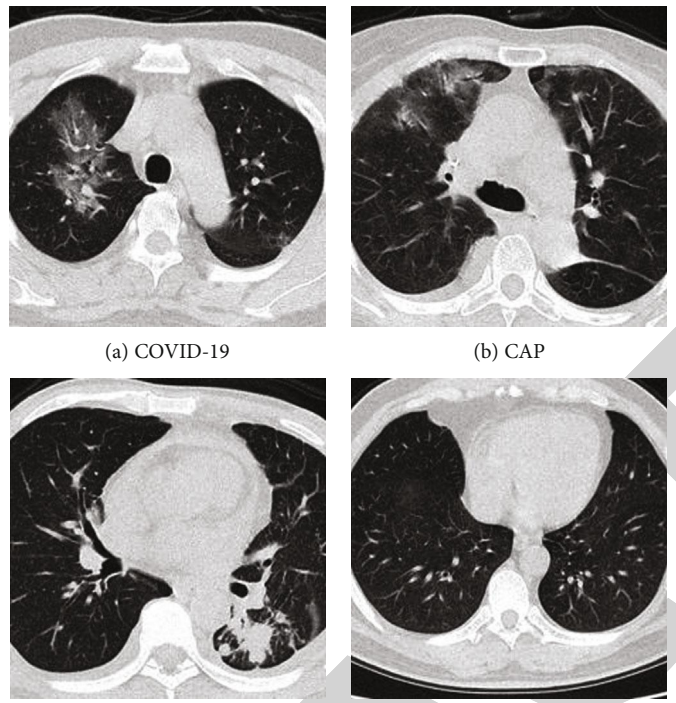


Figure 1: Illustration of preprocessing.



(a) COVID-19 (b) CAP

(c) SPT (d) HC

Figure 2: Samples of four categories (three diseased and one healthy).

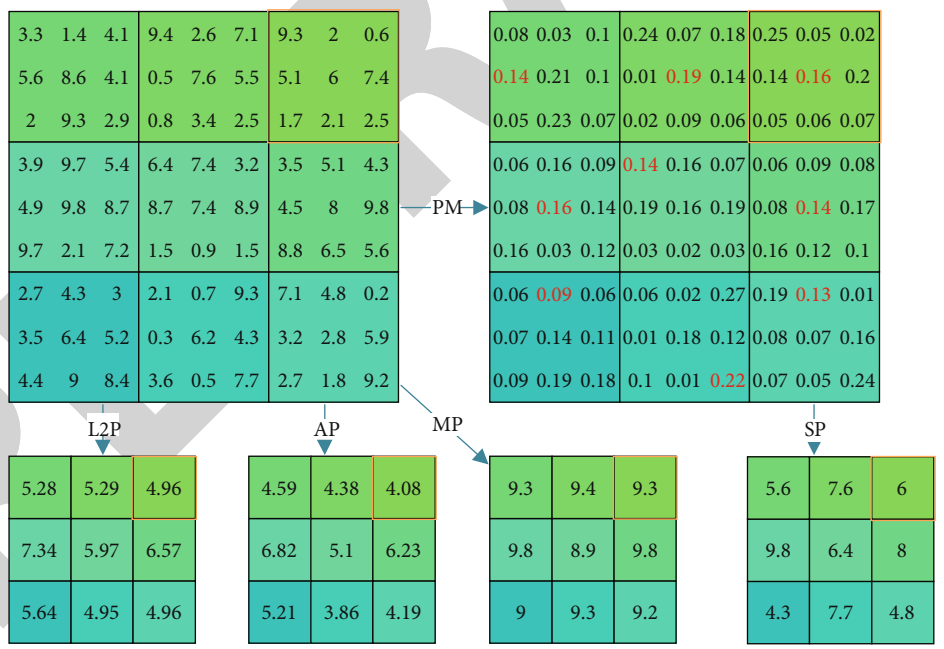


Figure 3: Comparison of four different pooling techniques (red rectangle indicates the example discussed in Section 3.2).

to the image set V_B. The new histogram stretched image set V_C was calculated as follows:

$$V_C = O_{\mathrm{HS}}(V_B) = \left\{ v_c(i) \stackrel{\mathrm{def}}{=} \frac{v_b(i) - v_b^l(i)}{v_b^h(i) - v_b^l(i)} \right\}, \tag{6}$$

where O_{HS} stands for the HS operation.

In the third step, cropping was performed to remove the checkup bed at the bottom area and eliminate the texts at the margin regions. Cropped dataset V_D is yielded as

$$\begin{aligned} V_D &= O_{\mathrm{crop}}(V_C, [a_1, a_2, a_3, a_4]) \\ &= \{v_d(1), v_d(2), \cdots, v_d(i), \cdots, v_d(|V|)\}, \end{aligned} \tag{7}$$

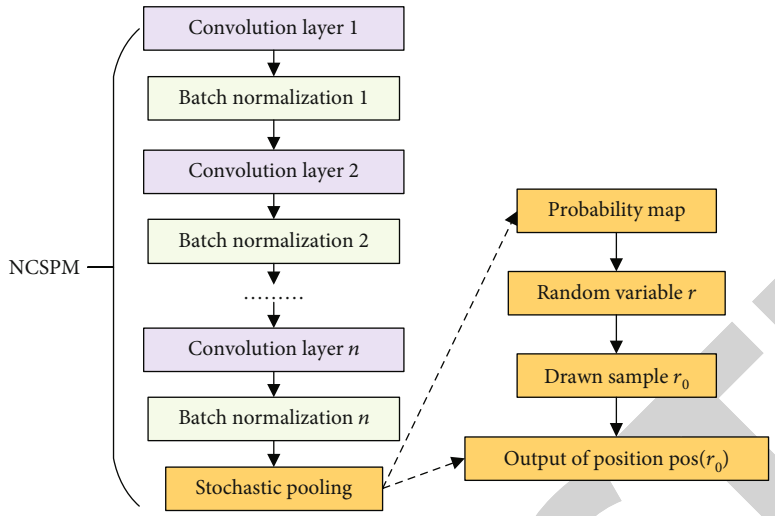


Figure 4: Schematic of proposed NCSPM.

where O_{crop} represents crop operation. Parameters (a_1, a_2, a_3, a_4) mean pixels to be cropped from four directions of the top, bottom, left, and right, respectively (unit: pixel).

In the fourth step, each image was downsampled to a size of $[W_E, H_E]$, obtaining the resized image set V_E as

$$\begin{aligned} V_E &= O_{\text{DS}}(V_D, [W_E, H_E]) \\ &= \{v_e(1), v_e(2), \cdots, v_e(i), \cdots v_e(|V|)\} \end{aligned} \tag{8}$$

where $O_{\text{DS}} : a \mapsto b$ represents the downsampling (DS) procedure, in which b stands for the downsampled image of the raw image a.

Figure 2 displays example images of the four categories, in which three are diseased, and one is healthy. The original size of each image in V_A is $1024 \times 1024 \times 3$, and the final preprocessed image is $256 \times 256 \times 1$.

3.2. Improvement I: n-conv Stochastic Pooling Module. First, stochastic pooling (SP) [17] was introduced. In the standard convolutional neural networks, pooling is an essential component after each convolution layer, which was applied to reduce the size of feature maps (FMs). SP was shown to give better performance than average pooling and max pooling in recent publications [18–21]. Recently, strided convolution (SC) is commonly used, which also can shrink the FMs [22, 23]. Nevertheless, SC could be considered a simple pooling method, which always outputs the region's fixed-position value [24].

Suppose we have a postconvolution FM $\mathbb{F} = f_{ij}(i = 1, \cdots, M \times P, j = 1, \cdots, N \times Q)$. The FM can be separated into $M \times N$ blocks, in which every block has the size of $P \times Q$. Now we focus on the block $B_{mn} = \{b(x, y), x = 1, \cdots, P, y = 1, \cdots, Q\}$ which stands for the m-th row and n-th column blocks.

Table 3: Structure of proposed 11-layer SPNN.

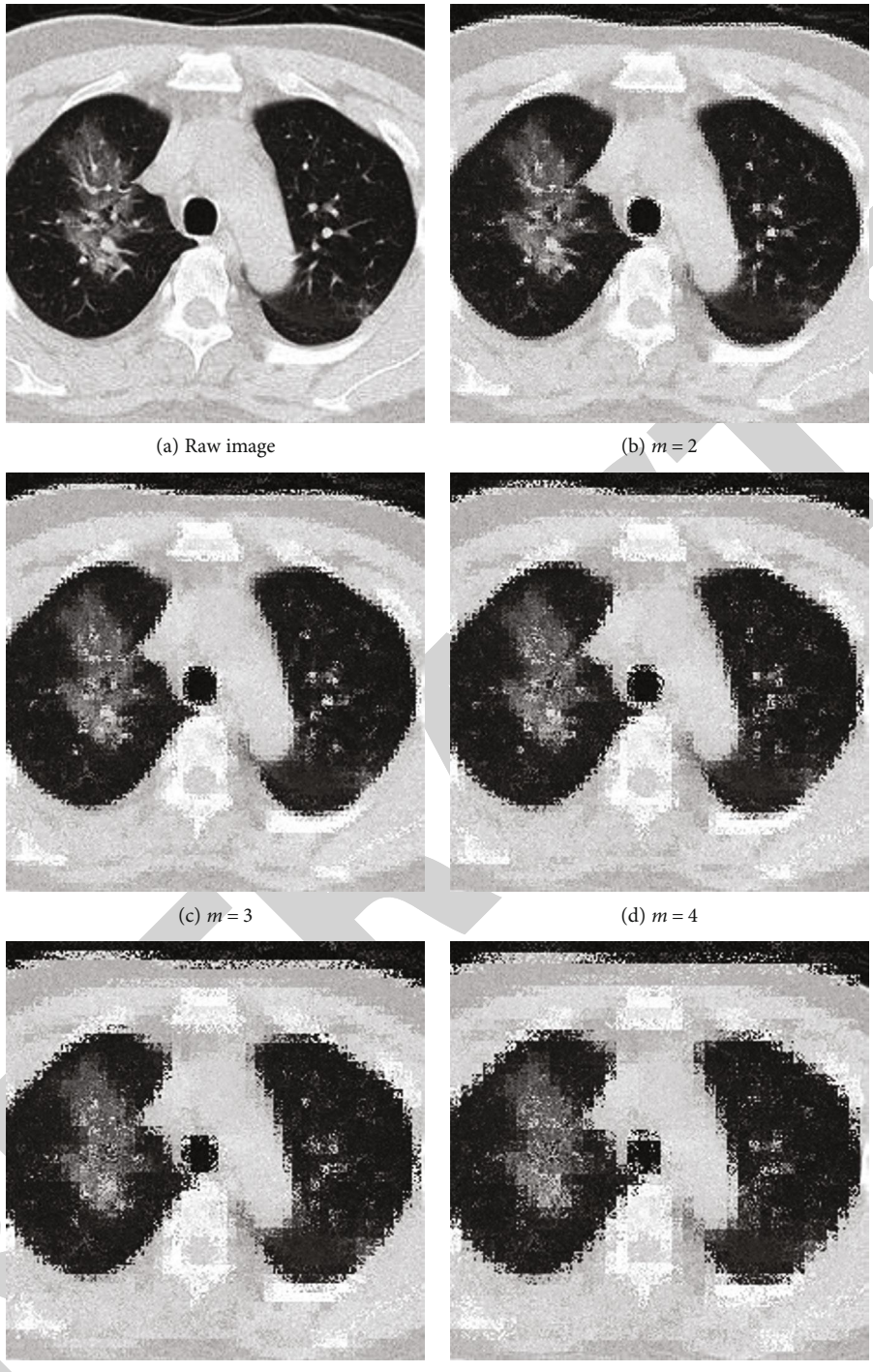
Index	Name	NWL	HS	FMS
1	Input			$256 \times 256 \times 1$
2	NCSPM-1	2	$[3 \times 3, 32] \times 2$	$128 \times 128 \times 32$
3	NCSPM-2	2	$[3 \times 3, 32] \times 2$	$64 \times 64 \times 32$
4	NCSPM-3	2	$[3 \times 3, 64] \times 2$	$32 \times 32 \times 64$
5	NCSPM-4	2	$[3 \times 3, 64] \times 2$	$16 \times 16 \times 64$
6	NCSPM-5	1	$[3 \times 3, 128] \times 1$	$8 \times 8 \times 128$
7	Flatten			$1 \times 1 \times 8192$
8	FCL-1	1	140×8192, 140×1	$1 \times 1 \times 140$
9	FCL-2	1	4×140, 4×1	$1 \times 1 \times 4$
10	Softmax			$1 \times 1 \times 4$

The strided convolution (SC) traverses the input activation map with the strides, which equals the size of the block (P, Q), so here its output is set to

$$B_{mn}^{\text{SC}} = b(1, 1). \tag{9}$$

The l2-norm pooling (L2P), average pooling (AP), and max pooling (MP) generate the l2-norm value, average value, and maximum value within the block B_{mn}, respectively.

$$\begin{cases} B_{mn}^{\text{L2P}} = \sqrt{\dfrac{\sum_{x=1}^{P}\sum_{y=1}^{Q} b^2(x, y)}{P \times Q}}, \\ B_{mn}^{\text{AP}} = \dfrac{1}{P \times Q}\displaystyle\sum_{x=1}^{P}\sum_{y=1}^{Q} b(x, y), \\ B_{mn}^{\text{MP}} = \max_{x=1}^{P} \max_{y=1}^{Q} b(x, y). \end{cases} \tag{10}$$



(a) Raw image (b) $m = 2$

(c) $m = 3$ (d) $m = 4$

(e) $m = 5$ (f) $m = 6$

Figure 5: Illustration of PatchShuffle on a grayscale COVID-19 image.

The SP provides a solution to the shortcomings of AP and MP. The AP outputs the average, so it will downscale the largest value, where the important features may sit on. On the other hand, MP reserves the maximum value but worsens the overfitting problem. SP is a three-step procedure. First, it generates the probability map (PM) for each entry in the block B_{mn}.

$$\begin{cases} p_m(x,y) = \dfrac{b(x,y)}{\sum_{x=1}^{P}\sum_{y=1}^{Q} b(x,y)}, \\ s.t. \sum_{x=1}^{P}\sum_{y=1}^{Q} p_m(x,y) = 1, \end{cases} \tag{11}$$

where $p_m(x,y)$ stands for the PM value at pixel (x,y).

(a) Raw image (b) $m = 2$

(c) $m = 3$ (d) $m = 4$

(e) $m = 5$ (f) $m = 6$

Figure 6: Illustration of PatchShuffle on a colorful house image.

In the second step, create a random variable r that takes the discrete probability distribution (DPD) as

$$\begin{cases} \Pr[r=(1,1)]=p_m(1,1) & \Pr[r=(1,2)]=p_m(1,2) & \cdots & \Pr[r=(1,Q)]=p_m(1,Q), \\ \Pr[r=(2,1)]=p_m(2,1) & \Pr[r=(2,2)]=p_m(2,2) & \cdots & \Pr[r=(2,Q)]=p_m(2,Q), \\ & \cdots \quad \cdots \quad \cdots \quad \cdots & & \\ \Pr[r=(P,1)]=p_m(P,1) & \Pr[r=(P,2)]=p_m(P,2) & \cdots & \Pr[r=(P,Q)]=p_m(P,Q), \end{cases} \tag{12}$$

where Pr represents the probability.

In the third step, a sample r_0 is drawn from the random variable r, and the corresponding position $\text{pos}(r_0) = (x_{r_0}, y_{r_0})$. Then, the output of SP is at location $\text{pos}(r_0)$, namely,

$$B_{mn}^{\text{SP}} = b\left(x_{r_0}, y_{r_0}\right). \tag{13}$$

Figure 3 presents the comparison of four different pooling techniques. The top left shows the raw FM in which the pooling will take place at a 3×3 kernel. If we take the top-right block (in an orange rectangle) as an example, the L2P outputs 4.96, while AP and MP output 4.08 and 9.3, respectively. For the SP method, it will first generate the PM and then sample a position based on the PM (see the red fonts), and thus, SP outputs the value of 6.

A new "n-conv stochastic pooling module" (NCSPM) is proposed in this study based on the SP layer discussed in previous paragraphs. The NCSPM entails n-repetitions of a conv layer and a batch normalization layer, followed by an SP layer. Figure 4 shows the schematic of the proposed NCSPM module. In this study, $n = 1 \vee 2$, since we experimented using $n = 3$, but the performance using $n = 3$ did not improve.

3.3. Improvement II: Stochastic Pooling Neural Network. The second improvement of this study is to propose a stochastic pooling neural network (SPNN), whose structure was inspired by VGG-16 [25]. In VGG-16, the network used small kernels instead of large kernels and always used 2×2 filters with a stride of 2 for pooling. In the end, VGG-16 has two fully connected layers (FCLs).

This proposed SPNN will follow the same structure design of VGG-16 but using the NCSPM module to replace the convolution block in VGG-16. The details of SPNN are shown in Table 3, where NWL means the number of weighted layers, HS is the hyperparameter size, and FMS is the feature map size.

Compared to ordinary CNN, the advantages of SPNN are two folds: (i) SPNN helps prevent overfitting; (ii) SPNN is parameter-free. (iii) SPNN can be easily combined with other advanced techniques, such as batch normalization and dropout. In total, we create this 11-layer SPNN. We have attempted to insert more NCSPMs or more FCLs, which does not show performance improvement but more computation burden. The structure of the proposed model is summarized in Table 3. The $[c_1 \times c_1, c_2] \times c_3$ related to NCSPM stands for c_3 repetitions of c_2 filters with size of $c_1 \times c_1$. For the FCL, the $c_1 \times c_2$, $c_3 \times c_4$ stands for a weight matrix is with size of $c_1 \times c_2$ and a bias matrix is with size of $c_3 \times c_4$. In the last column of Table 3, the format of $c_1 \times c_2 \times c_3$ represents the feature map's size in three dimensions: height, width, and channel. Directly using transfer learning is another alternative.

In this study, we chose to create a custom neural network by designing its structure and training the whole network using our own data. The reason is some reports have shown this "built your own network from scratch" can achieve better performance than transfer learning [26, 27].

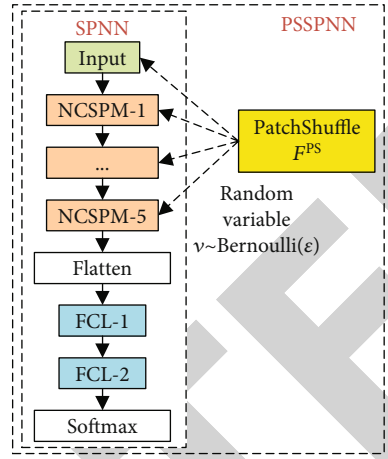


Figure 7: Build PSSPNN from SPNN.

3.4. Improvement III: PatchShuffle SPNN. Kang et al. [28] presented a new PatchShuffle method. In each minibatch, images and feature maps undergo a transformation such that pixels with that patch are shuffled. By generating fake images/feature maps with interior order-less patches, PatchShuffle creates local variations and reduces the possibility of the AI model overfitting. Therefore, PatchShuffle is a beneficial supplement to various existing training regularization methods [28].

Assume there is a matrix X of $M \times M$ entries. A random variable ν controls whether the matrix X to be PatchShuffled or not. The random variable ν obeys the Bernoulli distribution

$$\nu \sim \text{Bernoulli}(\varepsilon). \tag{14}$$

Namely, $\nu = 1$ with probability ε, and $\nu = 0$ with probability $1 - \varepsilon$. The resulted matrix $\widehat{X}$ is written as

$$\widehat{X} = (1 - \nu)X + \nu F^{\text{PS}}(X), \tag{15}$$

where F^{PS} is the PatchShuffle operation. Suppose the size of each patch is $m \times m$, we can express the matrix X as

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1,M/m} \\ x_{21} & x_{22} & \cdots & x_{2,M/m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{M/m,1} & x_{M/m,2} & \cdots & x_{M/m,M/m} \end{bmatrix}, \tag{16}$$

where x_{ij} stands for a nonoverlapping patch at i-th row and j-th column. The PatchShuffle transformation works on every patch.

$$F^{\text{PS}}(X) = \left\{ F^{\text{PS}}\left(x_{ij}\right), i = 1, \cdots, \frac{M}{m}, j = 1, \cdots, \frac{M}{m} \right\}. \tag{17}$$

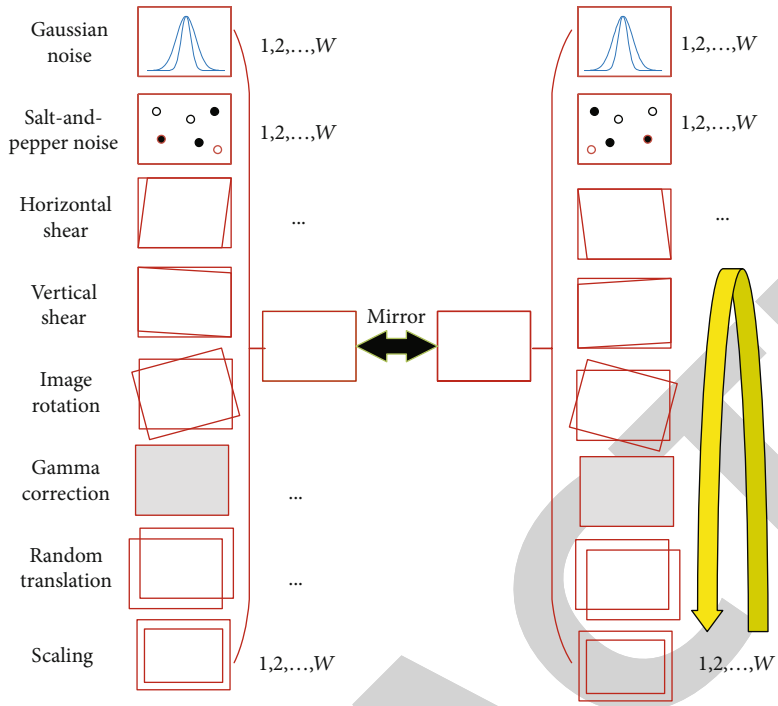


Figure 8: The proposed 16-way improved data augmentation method.

Table 4: Pseudocode of proposed 16-way improved data augmentation.

Input	Raw image $v(i)$
Step 1	Eight geometric or photometric DA transforms were utilized on raw image $v(i)$.
Step 2	A horizontal mirror image is generated.
Step 3	The raw image $v(i)$, the mirrored image, the above 8-way DA results of the raw image, and the 8-way DA results of the horizontal mirrored image are combined to form a dataset $\mathbb{D}(i)$.
Output	Enhanced dataset of raw image $\mathbb{D}(i)$

The shuffled patch is formulated as

$$F^{\mathrm{PS}}\left(x_{ij}\right) = e_{ij} \times x_{ij} \times e'_{ij}, \tag{18}$$

where e_{ij} is the row permutation matrix and e'_{ij} is the column permutation matrix [29]. In practice, a randomly shuffle operation is used to replace the row and column permutation operation. Each patch will undergo one of the $m^2!$ possible permutations.

We proposed integrating PatchShuffle into our SPNN, and this new network model is named as PatchShuffle stochastic pooling neural network (PSSPNN). The PatchShuffle operation acts on both the input image layer (see grayscale image Figure 5 and colorful image Figure 6 with different values of m) and the feature maps of all the convolutional layers (9 conv layers from NCSPM-1 to NCSPM-5).

The schematics of building PSSPNN from the SPNN are drawn in Figure 7, where either inputs or feature maps are randomly selected to undergo the PatchShuffle operation. To reach the best bias-variance trade-off, only a small percentage (ε) of the images/feature maps will undergo F^{PS} operation.

For simplicity, we consider the PatchShuffling images as an example, and the training loss function ℓ of the proposed PSSPNN is written as

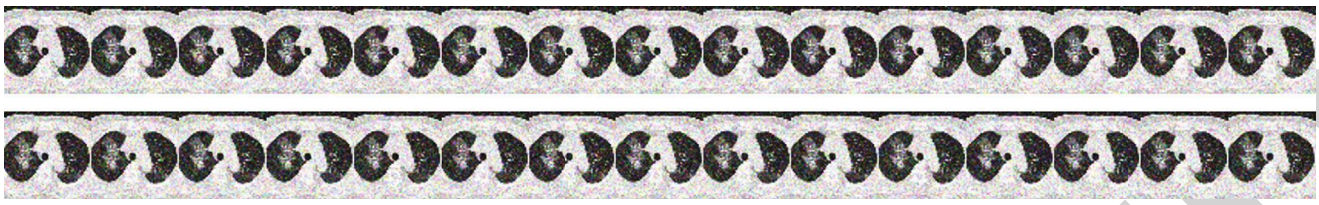
$$\ell^{\mathrm{PSSPNN}}(\mathbf{X}, y, \mathscr{W}) = (1 - v)\ell(\mathbf{X}, y, \mathscr{W}) + v\ell\left[F^{\mathrm{PS}}(\mathbf{X}), y, \mathscr{W}\right], \tag{19}$$

where ℓ stands for the ordinary loss function and ℓ^{PSSPNN} the loss function of PSSPNN. $\mathbf{X}$ represents the original images and $F^{\mathrm{PS}}(\mathbf{X})$ the PatchShuffled images. The label is symbolized as y, and the weights are symbolized as $\mathscr{W}$.

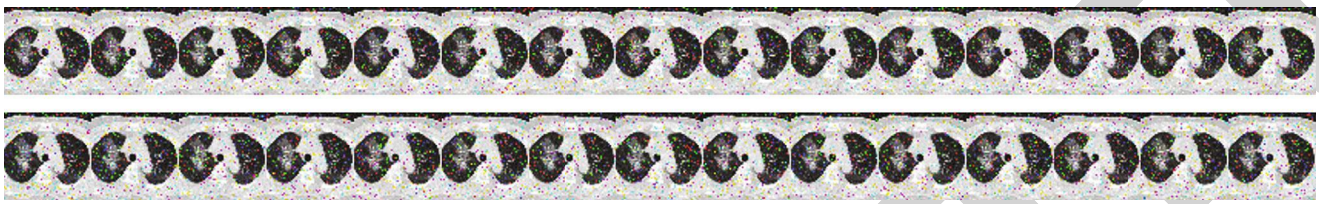
Considering the extreme situations when $v = 0 \vee 1$, we have

$$\ell^{\mathrm{PSSPNN}}(\mathbf{X}, y, \mathscr{W}) = \begin{cases} \ell(\mathbf{X}, y, \mathscr{W}) & v = 0, \\ \ell\left[F^{\mathrm{PS}}(\mathbf{X}), y, \mathscr{W}\right] & v = 1, \end{cases} \tag{20}$$

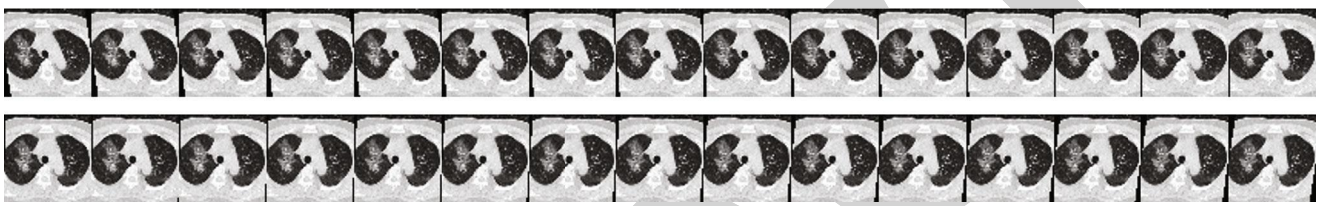
which means the loss function degrades to ordinary loss function when $v = 0$, and meanwhile, the loss function equals



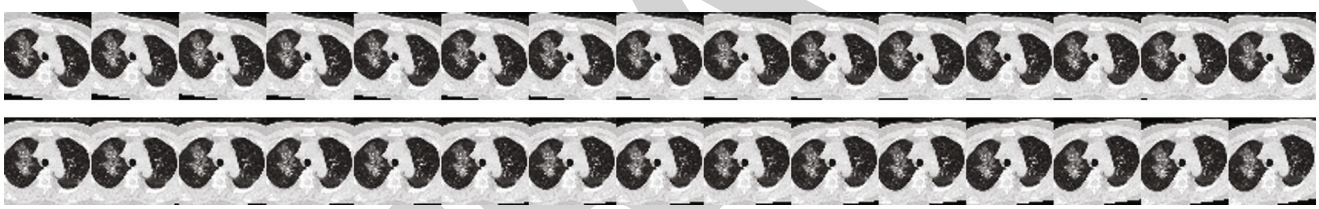
(a) Gaussian noise



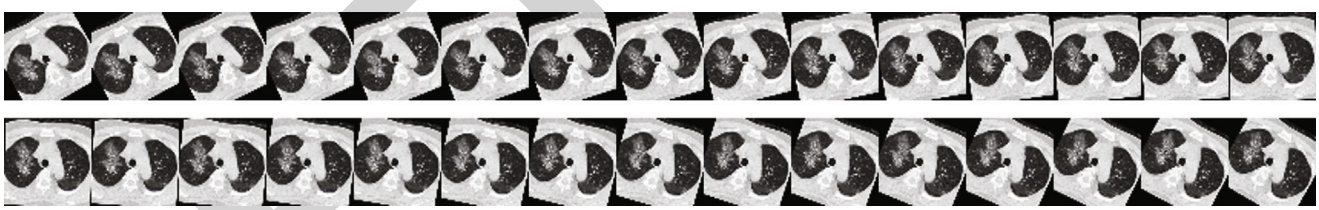
(b) Salt-and-pepper noise



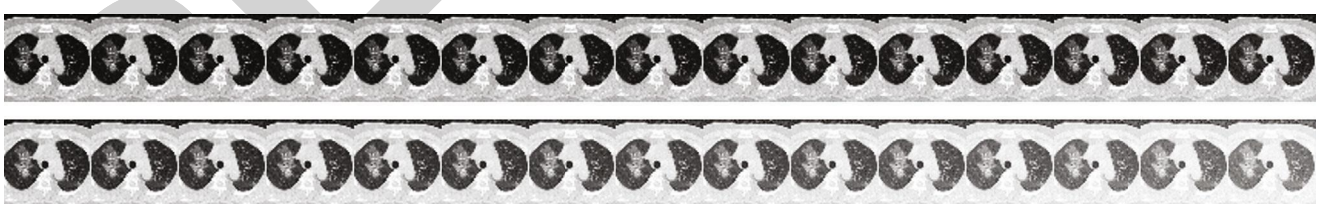
(c) Horizontal shear



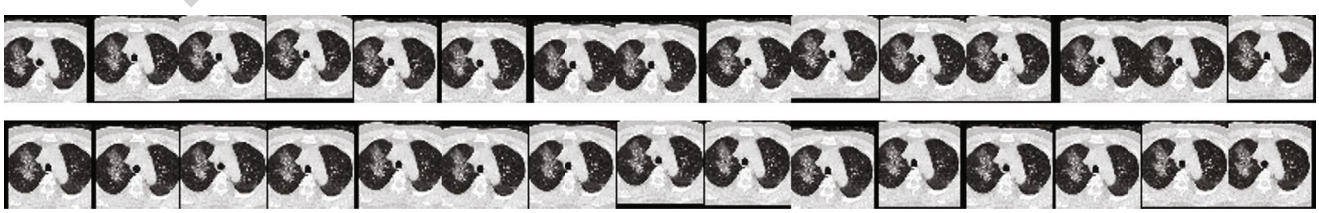
(d) Vertical shear



(e) Rotation

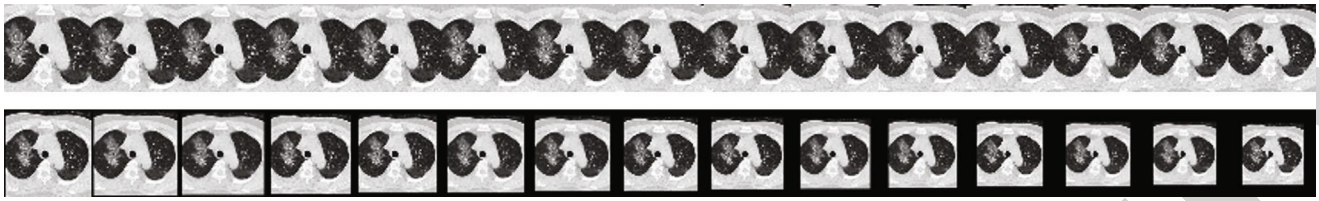


(f) Gamma correction



(g) Random translation

Figure 9: Continued.



(h) Scaling

Figure 9: Proposed 16-way data augmentation results.

Table 5: Dataset splitting.

	Nontest (10-fold cross-validation)	Test (10 runs)	Total
COVID-19	$\left\|V_1^{\text{ntest}}\right\| = 227$	$\left\|V_1^{\text{test}}\right\| = 57$	$\|V_1\| = 284$
CAP	$\left\|V_2^{\text{ntest}}\right\| = 225$	$\left\|V_2^{\text{test}}\right\| = 56$	$\|V_2\| = 281$
SPT	$\left\|V_3^{\text{ntest}}\right\| = 234$	$\left\|V_3^{\text{test}}\right\| = 59$	$\|V_3\| = 293$
HC	$\left\|V_4^{\text{ntest}}\right\| = 245$	$\left\|V_4^{\text{test}}\right\| = 61$	$\|V_4\| = 306$

to training all images PatchShuffled when $v = 1$. Taking mathematical expectation of v, equation (19) turns to

$$\frac{1}{1-\varepsilon}\mathbb{E}_v \ell^{\text{PSSPNN}}(\mathbf{X}, y, \mathscr{W}) = \ell(\mathbf{X}, y, \mathscr{W}) + \frac{\varepsilon}{1-\varepsilon}\ell\left[F^{\text{PS}}(\mathbf{X}), y, \mathscr{W}\right], \tag{21}$$

where $\varepsilon/1-\varepsilon\ell[F^{\text{PS}}(\mathbf{X}), y, \mathscr{W}]$ acts as a regularization term.

3.5. Improvement IV: Improved Multiple-Way Data Augmentation. This small four-category dataset makes our AI model prone to overfitting. In order to alleviate the overfitting and handle the low sample-size problem, the multiple-way data augmentation (MDA) [30] method was chosen and further improved. In the original 14-way MDA [30], the authors used seven different data augmentation (DA) techniques to the raw image and its horizontal image. Their seven DA techniques are as follows: noise injection, horizontal shear, vertical shear, rotation, Gamma correction, scaling, and translation.

Figure 8 shows a 16-way data augmentation method. The difference between the proposed 16-way DA with the traditional 14-way DA is that we add the salt-and-pepper noise (SAPN). Although the SAPN defies intuition as it never takes place in realistic CCT images, we found that it can increase performance. The same observation was reported by Li et al. [31], where the authors used salt and pepper noise for the identification of early esophageal cancer. Table 4 shows the pseudocode of this proposed 16-way improved data augmentation.

$$v(i) \mapsto \mathbb{D}(i) = \mathbb{S}\left\{\begin{array}{cc} v(i) & M[v(i)] \\ \underset{k=1,..,8}{\mathbb{S}} \underbrace{f_k^{\text{DA}}[v(i)]}_{W} & \underset{k=1,..,8}{\mathbb{S}} \underbrace{f_k^{\text{DA}}\{M[v(i)]\}}_{W} \end{array}\right\}, \tag{22}$$

Table 6: Comparison of SPNN with four standard CNNs.

Model	Class	Sen	Prc	F1
SC-CNN	C1	93.51	93.35	93.43
	C2	86.79	91.87	89.26
	C3	94.41	91.01	92.68
	C4	94.59	93.37	93.97
	MA			92.40
AP-CNN	C1	92.86	93.02	92.94
	C2	95.76	92.93	94.32
	C3	94.26	93.95	94.11
	C4	93.22	93.22	93.22
	MA			93.18
SPNN (ours)	C1	98.07	94.59	96.30
	C2	91.79	95.19	93.45
	C3	94.92	94.92	94.92
	C4	95.25	95.40	95.32
	MA			95.02
L2P-CNN	C1	88.77	93.36	91.01
	C2	93.21	93.72	93.46
	C3	94.41	92.68	93.53
	C4	93.61	90.63	92.10
	MA			92.53
MP-CNN	C1	95.44	94.77	95.10
	C2	94.64	91.70	93.15
	C3	94.07	93.43	93.75
	C4	91.64	95.72	93.63
	MA			93.91

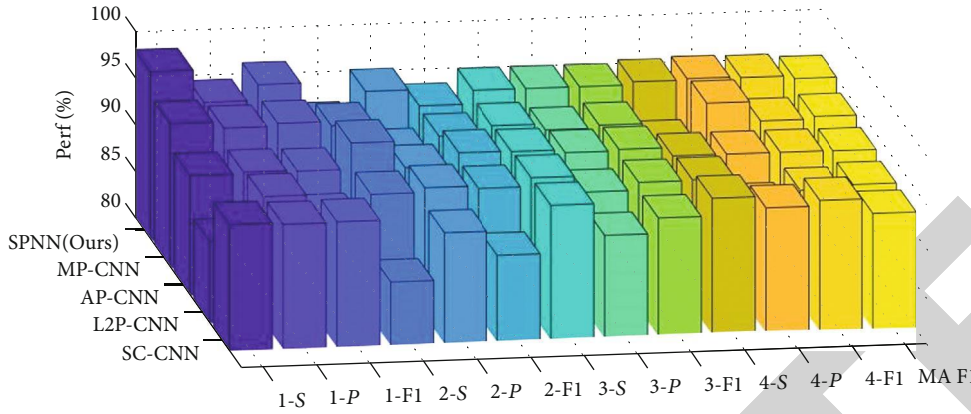


Figure 10: 3D bar plot of SPNN against other network models.

where $W = 30$ in this study. We tested a greater value of W, but it does not bring about significant improvement. Hence, one image $v(i)$ will generate $|\mathbb{D}(i)| = 16W + 2 = 482$ images (including original image), as shown in Figure 8.

Step 1. Eight geometric or photometric DA transforms were utilized on raw image $v(i)$, as shown in Figure 8. We use $f_k^{\mathrm{DA}}, k = 1, \cdots, 8$ to denote each DA operation. Note each DA operation f_k^{DA} will yield W new images. So, for a given image $v(i)$, we will produce an enhanced dataset $\mathop{\mathbb{S}}_{k=1,..,8} f_k^{\mathrm{DA}}[v(i)]$, where $\mathbb{S}$ stands for concatenation function.

Step 2. Horizontal mirror image is generated as $M[v(i)]$, where M means horizontal mirror function.

Step 3. The raw image $v(i)$, the mirrored image $M[v(i)]$, all the above 8-way DA results of raw image $\mathop{\mathbb{S}}_{k=1,..,8} f_k^{\mathrm{DA}}[v(i)]$, and 8-way DA results of horizontal mirrored image $\mathop{\mathbb{S}}_{k=1,..,8} f_k^{\mathrm{DA}}\{\mathrm{M}[v(i)]\}$ are combined. Mathematically, one training image $v(i)$ will generate to a dataset $\mathbb{D}(i)$, which contains $16W + 1$ new images.

Taking Figure 2(a) as an example raw image, Figure 9 shows the 8-way DA results, i.e., $f_k^{\mathrm{DA}}[v(i)], k = 1, \cdots, 8$. Due to the page limit, the mirror image and its corresponding 8-way DA results are not shown here.

3.6. Implementation, Measure, and Explainability. Table 5 itemizes the nontest and test sets for each category. The whole dataset V contains four nonoverlapping categories $V = \{V_1, V_2, V_3, V_4\}$. For each category, the dataset will be split into nontest set and test set $V_k \to \{V_k^{\mathrm{ntest}}, V_k^{\mathrm{test}}\}$, $k = 1, 2, 3, 4$.

$$V = \begin{bmatrix} V_1 \\ V_2 \\ V_3 \\ V_4 \end{bmatrix} = \begin{bmatrix} V^{\mathrm{ntest}} & V^{\mathrm{test}} \end{bmatrix} = \begin{bmatrix} V_1^{\mathrm{ntest}} & V_1^{\mathrm{test}} \\ V_2^{\mathrm{ntest}} & V_2^{\mathrm{test}} \\ V_3^{\mathrm{ntest}} & V_3^{\mathrm{test}} \\ V_4^{\mathrm{ntest}} & V_4^{\mathrm{test}} \end{bmatrix}. \quad (23)$$

Table 7: Hyperparameter optimization of PSSPNN in terms of microaveraged F1 score.

ε	Patch size			
	1×2	2×2	2×4	3×3
0.01	95.19	95.54	95.15	94.98
0.05	95.49	95.79	95.45	95.11
0.10	95.24	95.28	95.54	95.41
0.15	95.24	95.11	95.28	95.06
0.20	95.06	95.15	94.85	94.72

Our experiment entails two phases. At phase I, 10-fold cross-validation was used for validation on the nontest set to select the best hyperparameters and best network structure. The 16-way DA was used on the training set of 10-fold cross-validation. The hyperparameter of the proposed PSSPNN was determined over the nontest set V^{ntest}. Afterward at phase II, we train our model using the nontest set V^{ntest} 10 times with different initial seeds and attain the test results over the test set V^{test}. After combining the R^{test} runs, we attain a summation of the test confusion matrix (TCM) D^{test}. Table 5 shows the dataset splitting, where $|x|$ stands for the number of elements in the dataset x.

The ideal TCM is a diagonal matrix with the form of

$$D_{\mathrm{ideal}}^{\mathrm{test}} = R^{\mathrm{test}} \times \begin{bmatrix} \left|V_1^{\mathrm{test}}\right| & 0 & 0 & 0 \\ 0 & \left|V_2^{\mathrm{test}}\right| & 0 & 0 \\ 0 & 0 & \left|V_3^{\mathrm{test}}\right| & 0 \\ 0 & 0 & 0 & \left|V_4^{\mathrm{test}}\right| \end{bmatrix}, \quad (24)$$

in which all the off-diagonal elements are zero, meaning no prediction errors. In realistic scenarios, the AI model will make errors, and the performance is calculated per category. For each class $z = 1, \cdots, 4$, we set the label of that class as *positive*, and the labels of all the rest classes as *negative*. Three

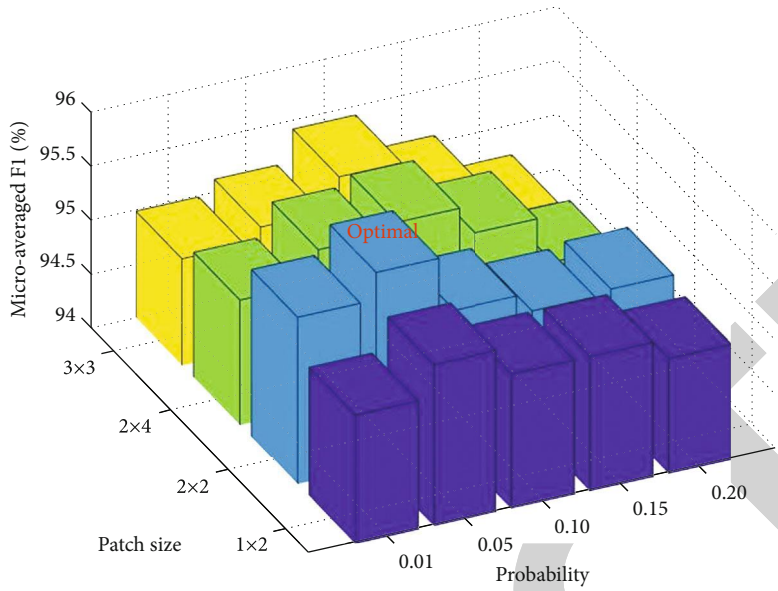


Figure 11: 3D bar chart of microaveraged F1 against patch size and probability.

performance metrics (sensitivity, precision, and F1 score) per category are defined below:

$$\begin{cases} \mathrm{Sen}(z) = \dfrac{\mathrm{TP}(z)}{\mathrm{TP}(z) + \mathrm{FN}(z)}, \\ \mathrm{Prc}(z) = \dfrac{\mathrm{TP}(z)}{\mathrm{TP}(z) + \mathrm{FP}(z)}, \\ \mathrm{F1}(z) = \dfrac{2 * \mathrm{Prc}(z) * \mathrm{Sen}(z)}{\mathrm{Prc}(z) + \mathrm{Sen}(z)}. \end{cases} \tag{25}$$

The performance can be measured over all four categories. The microaveraged (MA) F1 (symbolized as $\mathrm{F1}_\mu$) is used since our dataset is slightly unbalanced:

$$\mathrm{F1}_\mu = \frac{2 * \mathrm{Pr}\, c_\mu * \mathrm{Sen}_\mu}{\mathrm{Pr}\, c_\mu + \mathrm{Sen}_\mu}, \tag{26}$$

where

$$\begin{cases} \mathrm{Sen}_\mu = \dfrac{\sum_{z=1}^{4} \mathrm{TP}(z)}{\sum_{z=1}^{4} \mathrm{TP}(z) + \mathrm{FN}(z)}, \\ \mathrm{Pr}\, c_\mu = \dfrac{\sum_{z=1}^{4} \mathrm{TP}(z)}{\sum_{z=1}^{4} \mathrm{TP}(z) + \mathrm{FP}(z)}. \end{cases} \tag{27}$$

Finally, gradient-weighted class activation mapping (Grad-CAM) [32] was employed to provide explanations on how our model makes the decision. It exploits the gradient of the categorization score regarding the convolutional features decided by the deep model to visualize the regions of the image that are the most vital for the image classification task [33]. The output of NCSPM-5 in Table 3 was used for Grad-CAM.

Table 8: Comparison of SPNN against PSSPNN.

Model	Class	Sen	Prc	F1
SPNN (ours)	C1	98.07	94.59	96.30
	C2	91.79	95.19	93.45
	C3	94.92	94.92	94.92
	C4	95.25	95.40	95.32
	MA			95.02
PSSPN (ours)	C1	97.89	95.06	96.46
	C2	92.86	96.30	94.55
	C3	95.76	95.44	95.60
	C4	96.56	96.40	96.48
	MA			95.79

4. Experiments, Results, and Discussions

The experiment was carried out on the programming platform of Matlab 2020b. The programs ran on Windows 10 with 16GB RAM and 2.20GHz Intel Core i7-8750H CPU. The performances are reported over the test set with 10 runs.

4.1. Comparison of SPNN and Other Pooling Methods. In the first experiment, we compared the proposed SPNN with four standard CNNs with different pooling methods. The first CNN uses strided convolution in five modules to replace the stochastic pooling. The second to fourth comparison CNN models use L2P, AP, and MP, respectively. Those four baseline methods are called SC-CNN, L2P-CNN, AP-CNN, and MP-CNN, respectively. The results of 10 runs over the test set are shown in Table 6. The bar plot is displayed in Figure 10, where "k-S," "k-P," and ""k-F1" stand for the sensitivity, precision, and F1 score for category $k \in \{1, 2, 3, 4\}$.

The results in Table 6 and Figure 10 are coherent with our expectation that SPNN obtained the best results among all FM reduction approaches. The SPNN arrives at the

Table 9: Comparison with state-of-the-art approaches.

Model	Class	Sen	Prc	F1
RCBO [7]	C1	71.93	84.19	77.58
	C2	72.86	72.73	72.79
	C3	73.56	76.41	74.96
	C4	80.66	68.91	74.32
	MA			74.85
RN-50 [9]	C1	87.72	85.03	86.36
	C2	87.68	91.26	89.44
	C3	93.39	89.89	91.60
	C4	84.92	87.65	86.26
	MA			88.41
CSSNet [11]	C1	94.04	92.25	93.14
	C2	93.75	95.11	94.42
	C3	91.36	93.58	92.45
	C4	94.43	92.75	93.58
	MA			93.39
DeCovNet [13]	C1	91.05	90.58	90.81
	C2	93.75	90.99	92.35
	C3	90.51	86.97	88.70
	C4	88.69	95.58	92.01
	MA			90.94
FCONet [15]	C1	92.28	95.64	93.93
	C2	96.79	94.43	95.59
	C3	94.75	95.88	95.31
	C4	94.92	92.94	93.92
	MA			94.68
6L-CNN [8]	C1	72.46	83.94	77.78
	C2	78.93	77.82	78.37
	C3	81.86	75.00	78.28
	C4	89.84	87.54	88.67
	MA			80.94
RN-18 [10]	C1	82.81	82.66	82.73
	C2	81.07	74.43	77.61
	C3	74.24	76.98	75.58
	C4	82.13	86.38	84.20
	MA			80.04
COVNet [12]	C1	89.82	86.63	88.20
	C2	89.82	92.63	91.21
	C3	93.73	90.66	92.17
	C4	87.38	90.96	89.13
	MA			90.17
7L-CCD [14]	C1	89.47	93.58	91.48
	C2	93.93	92.44	93.18
	C3	93.73	95.18	94.45
	C4	95.08	91.34	93.17
	MA			93.09
PSSPNN (ours)	C1	97.89	95.06	96.46
	C2	92.86	96.30	94.55
	C3	95.76	95.44	95.60
	C4	96.56	96.40	96.48
	MA			95.79

sensitivities of all four categories are 98.07%, 91.79%, 94.92%, and 95.25%, respectively. The precisions of all four categories are 94.59%, 95.19%, 94.92%, and 95.40%, respectively. The F1-scores of the four categories are 96.30%, 93.45%, 94.92%, and 95.32%, respectively. The overall microaveraged F1 is 95.02%.

In terms of microaveraged F1, the second-best algorithm is MP-CNN, which obtains a microaveraged F1 score of 93.91%. The third best is AP-CNN, with a microaveraged F1 score of 93.18%. The two comparably worst algorithms are SC-CNN and L2P-CNN, with the microaveraged F1 scores of 92.40% and 92.53%, respectively.

SPNN obtains the best results because SP can prevent overfitting [17], which is the main shortcoming of max pooling. On the other hand, AP and L2P will average out the maximum activation values, which will impair the performances of convolutional neural network models. For SC-CNN, it only uses one quarter information of the input FM, and therefore may neglect those greatest values [34]. In all, this proposed SPNN can be regarded as an improved version of vanilla CNN models, where the SP is used to replace traditional MP.

4.2. PSSPNN versus SPNN. In this second experiment, we compared our two proposed network models, PSSPNN against SPNN, to validate the effectiveness of PatchShuffle. The results of 10 runs over the test set with different combinations of hyperparameters are shown in Table 7, and the 3D bar chart is shown in Figure 11. The optimal hyperparameter we found from the 10-fold cross-validation of the nontest set is $\varepsilon = 0.05$ and patch size is 2×2, which are coherent with reference [28].

In addition, the PSSPNN with optimal hyperparameter is compared with SPNN. The results are shown in Table 8. From the table, we can observe that PSSPNN provides better F1 values for all four categories and the overall microaverage, which shows the effectiveness of PatchShuffle. The reason is PatchShuffle adds regularization terms [28] in the loss function, and thus can improve the generalization ability of our SPNN model.

4.3. Comparison to State-of-the-Art Approaches. We compared our proposed PSSPNN method with 9 state-of-the-art methods: RCBO [7], 6L-CNN [8], RN-50 [9], RN-18 [10], CSSNet [11], COVNet [12], DeCovNet [13], 7L-CNN-CD [14], and FCONet [15]. All the comparison was carried on the same test set of 10 runs. The comparison results are shown in Table 9.

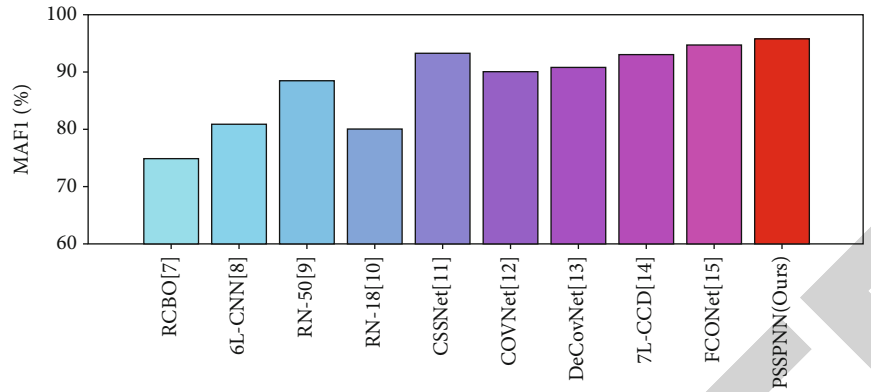
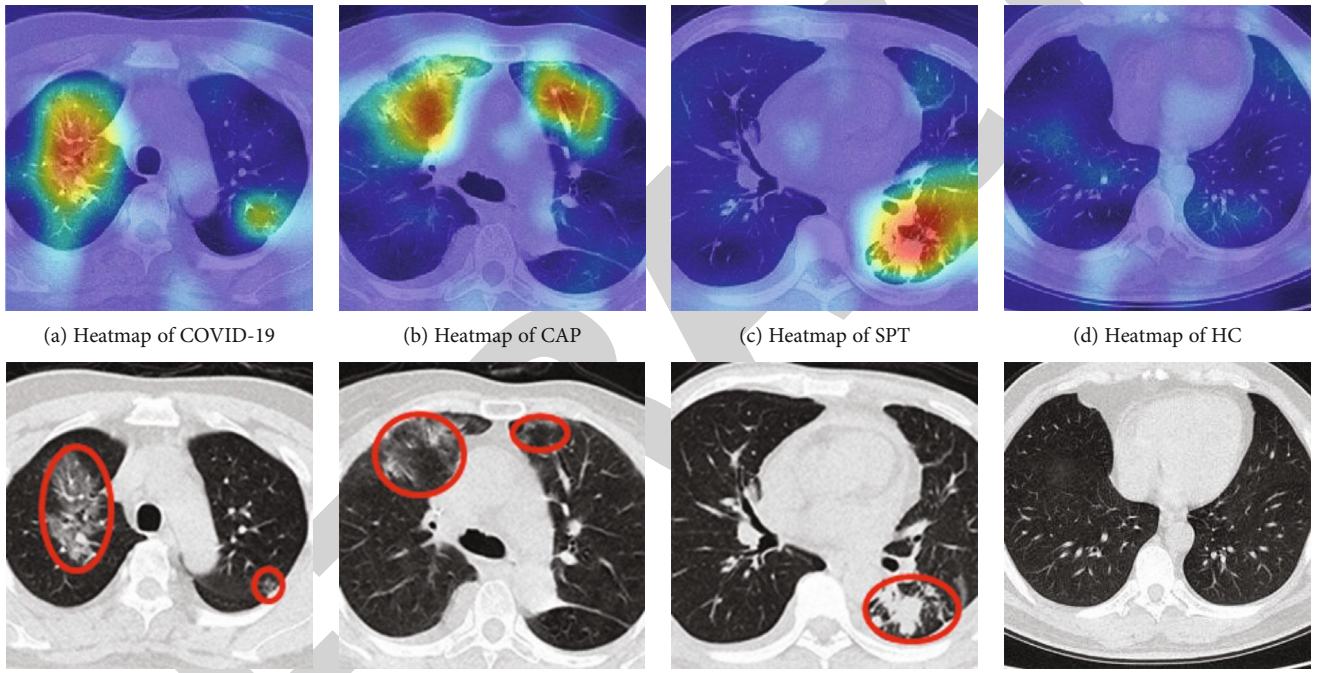


Figure 12: Algorithm comparison in terms of MA F1.



(a) Heatmap of COVID-19 (b) Heatmap of CAP (c) Heatmap of SPT (d) Heatmap of HC

(e) Lesion of COVID-19 (f) Lesion of CAP (g) Lesion of SPT (h) No lesion for HC

Figure 13: Delineation of three diseased samples.

For ease of comparison, Figure 12 only compares the microaveraged F1 (MA F1) score of all algorithms, from which we can observe this proposed PSSPNN achieves the best performance among all the algorithms. This experiment is a simulation-based comparison. In the future, we will apply our algorithm to rigorous clinical testing and verification.

4.4. Explainability of Proposed PSSPNN. We take Figure 2 images as examples; the heatmaps of those four images are shown in Figures 13(a)–13(d), and the manual delineation, shown in Figures 13(e)–13(h), delineates the lesions of the three disease samples. Note there are no lesions of healthy control (HC) image. The NCSPM-5 feature map in PSSPNN was used to generate heatmaps with the Grad-CAM approach.

We can observe from Figure 13 that the heatmaps via our PSSPNN model and Grad-CAM can capture the lesions effectively and meanwhile neglect those nonlesion regions. Traditionally, AI is regarded as a "black box," which impairs its widespread usage, e.g., the black box properties of traditional AI are problematic for the FDA. Nevertheless, with the help of explainability of modern AI techniques [35], the radiologist and patients will gain confidence in our proposed AI model, as the heatmap provides a clear and understandable interpretation of how AI predicts COVID-19 and other chest infectious disease from healthy subjects, which was also stated in reference [36]. Many new AI-based anatomic pathological systems now pass through FDA approval, such as whole slide images (WSI) [37], since the doctors know the relationships between the diagnosis and the explained answer.

In the future, the explainability of our proposed AI model can be used in patient monitoring [38] and health big data [39]. Some novel network improvement and signal

processing techniques may help our AI model in future researches, such as filters [40, 41], fuzzy [42, 43], edge computing [44], knowledge-aid [45, 46], autofocus [47], graph integration, and cross-domain knowledge exploitation [48–50].

5. Conclusion

In this paper, we proposed a PSSPNN, which entails five improvements: (i) proposed NCSPM module, (ii) usage of stochastic pooling, (iii) usage of PatchShuffle, (iv) improved multiple-way data augmentation, and (v) explainability via Grad-CAM. Those five improvements enable our AI model to deliver improved performances compared to 9 state-of-the-art approaches. The 10 runs on the test set showed our algorithm achieved a microaveraged F1 score of 95.79%.

There are three shortcomings of our method, which will be resolved in the future: (i) the dataset currently contains three chest infectious diseases. In the future, we shall try to include more classes of chest diseases, such as thoracic cancer. (ii) Some new network techniques and models are not tested, such as transfer learning, wide network module design, attention mechanism, and graph neural network. Those advanced AI technologies will be studied. (iii) Our model does not go through strict clinical validation, so we will attempt to release our software to hospitals and get feedback from radiologists and consultants.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Shui-Hua Wang and Yin Zhang contributed equally to this paper.

Acknowledgments

We appreciate Qinghua Zhou for helping us revise our English. This paper is partially supported by the Royal Society International Exchanges Cost Share Award, UK (RP202G0230); Medical Research Council Confidence in Concept Award, UK (MC_PC_17171); Hope Foundation for Cancer Research, UK (RM60G0680); British Heart Foundation Accelerator Award, UK; Open fund for Jiangsu Key Laboratory of Advanced Manufacturing Technology (HGAMTL-1703); Guangxi Key Laboratory of Trusted Software (kx201901).

References

[1] A. A. Szigiato, M. Anderson, M. Mabon, M. Germain, G. M. Durr, and A. C. Labbe, "Usefulness of prestorage corneal swab culture in the prevention of contaminated corneal tissue in corneal transplantation," *Cornea*, vol. 39, no. 7, pp. 827–833, 2020.

[2] R. Mögling, A. Meijer, N. Berginc et al., "Delayed laboratory response to covid-19 caused by molecular diagnostic contamination," *Emerging Infectious Diseases*, vol. 26, no. 8, pp. 1944–1946, 2020.

[3] T. Ai, Z. Yang, H. Hou et al., "Correlation of chest CT and RT-PCR testing for coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases," *Radiology*, vol. 296, no. 2, pp. E32–E40, 2020.

[4] R. Jain, M. Gupta, S. Taneja, and D. J. Hemanth, "Deep learning based detection and analysis of COVID-19 on chest X-ray images," *Applied Intelligence*, vol. 51, no. 3, pp. 1690–1700, 2021.

[5] H. Conway, G. Lau, and V. Zochios, "Personalizing invasive mechanical ventilation strategies in coronavirus disease 2019 (COVID-19)-associated lung injury: the utility of lung ultrasound," *Journal of Cardiothoracic and Vascular Anesthesia*, vol. 34, no. 10, pp. 2571–2574, 2020.

[6] A. K. Goel, D. DiLella, G. Dotsikas, M. Hilts, D. Kwan, and L. Paxton, "Unlocking radiology reporting data: an implementation of synoptic radiology reporting in low-dose CT cancer screening," *Journal of Digital Imaging*, vol. 32, no. 6, pp. 1044–1051, 2019.

[7] S. Wang, P. Li, P. Chen et al., "Pathological brain detection via wavelet packet Tsallis entropy and real-coded biogeography-based optimization," *Fundamenta Informaticae*, vol. 151, no. 1-4, pp. 275–291, 2017.

[8] X. Jiang and Y. D. Zhang, "Chinese sign language fingerspelling via six-layer convolutional neural network with leaky rectified linear units for therapy and rehabilitation," *Journal of Medical Imaging and Health Informatics*, vol. 9, no. 9, pp. 2031–2090, 2019.

[9] L. V. Fulton, D. Dolezel, J. Harrop, Y. Yan, and C. P. Fulton, "Classification of Alzheimer's disease with and without imagery using gradient boosted machines and ResNet-50," *Brain Sciences*, vol. 9, no. 9, p. 212, 2019.

[10] M. H. Guo and Y. Z. Du, "Classification of thyroid ultrasound standard plane images using ResNet-18 networks," in *2019 IEEE 13th International Conference on Anti-counterfeiting, Security, and Identification (ASID)*, pp. 324–328, Xiamen, China, 2019.

[11] J. P. Cohen, L. Dao, K. Roth et al., "Predicting COVID-19 pneumonia severity on chest X-ray with deep learning," *Cureus*, vol. 12, article e9448, 2020.

[12] L. Li, L. Qin, Z. Xu et al., "Using artificial intelligence to detect COVID-19 and community-acquired pneumonia based on pulmonary CT: evaluation of the diagnostic accuracy," *Radiology*, vol. 296, no. 2, pp. E65–E71, 2020.

[13] X. G. Wang, X. B. Deng, Q. Fu et al., "A weakly-supervised framework for COVID-19 classification and lesion localization from chest CT," *IEEE Transactions on Medical Imaging*, vol. 39, no. 8, pp. 2615–2625, 2020.

[14] Y. D. Zhang, S. C. Satapathy, L. Y. Zhu, J. M. Gorriz, and S. H. Wang, "A seven-layer convolutional neural network for chest CT based COVID-19 diagnosis using stochastic pooling," *IEEE Sensors Journal*, 2020.

[15] H. Ko, H. Chung, W. S. Kang et al., "COVID-19 pneumonia diagnosis using a simple 2D deep learning framework with a single chest CT image: model development and validation," *Journal of Medical Internet Research*, vol. 22, article e19569, p. 13, 2020.

[16] S.-H. Wang, D. R. Nayak, D. S. Guttery, X. Zhang, and Y. D. Zhang, "COVID-19 classification by CCSHNet with deep fusion using transfer learning and discriminant correlation analysis," *Information Fusion*, vol. 68, pp. 131–148, 2021.

[17] M. D. Zeiler and R. Fergus, "Stochastic pooling for regularization of deep convolutional neural networks," 2013, https://arxiv.org/abs/1301.3557.

[18] Y.-D. Zhang, D. R. Nayak, X. Zhang, and S.-H. Wang, "Diagnosis of secondary pulmonary tuberculosis by an eight-layer improved convolutional neural network with stochastic pooling and hyperparameter optimization," *Journal of Ambient Intelligence and Humanized Computing*, 2020.

[19] X. W. Jiang, M. Lu, and S. H. Wang, "An eight-layer convolutional neural network with stochastic pooling, batch normalization and dropout for fingerspelling recognition of Chinese sign language," *Multimedia Tools and Applications*, vol. 79, no. 21-22, pp. 15697–15715, 2020.

[20] A. Jahanbakhshi, M. Momeny, M. Mahmoudi, and Y. D. Zhang, "Classification of sour lemons based on apparent defects using stochastic pooling mechanism in deep convolutional neural networks," *Scientia Horticulturae*, vol. 263, article 109133, p. 10, 2020.

[21] B. C. Zhou, X. L. Wang, and Q. Q. Qi, "Optimal weights decoding of M-ary suprathreshold stochastic resonance in stochastic pooling network," *Chinese Journal of Physics*, vol. 56, no. 4, pp. 1718–1726, 2018.

[22] H. Vu, H. C. Kim, M. Jung, and J. H. Lee, "fMRI volume classification using a 3D convolutional neural network robust to shifted and scaled neuronal activations," *Neuroimage*, vol. 223, article 117328, 2020.

[23] J. Yang, Z. X. Ji, S. J. Niu, Q. Chen, S. T. Yuan, and W. Fan, "RMPPNet: residual multiple pyramid pooling network for subretinal fluid segmentation in SD-OCT images," *OSA Continuum*, vol. 3, no. 7, pp. 1751–1769, 2020.

[24] B. Brzycki, A. P. Siemion, S. Croft et al., "Narrow-band signal localization for SETI on noisy synthetic spectrogram data," *Publications of the Astronomical Society of the Pacific*, vol. 132, article 114501, 2020.

[25] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *International Conference on Learning Representations (ICLR)*, pp. 1–14, San Diego, CA, USA, 2015.

[26] A. Hazra, P. Choudhary, S. Inunganbi, and M. Adhikari, "Bangla-Meitei Mayek scripts handwritten character recognition using convolutional neural network," *Applied Intelligence*, 2020.

[27] Y. H. Sabry, W. Z. W. Hasan, A. H. Sabry, M. Z. A. A. Kadir, M. A. M. Radzi, and S. Shafie, "Measurement-based modeling of a semitransparent CdTe thin-film PV module based on a custom neural network," *IEEE Access*, vol. 6, pp. 34934–34947, 2018.

[28] G. Kang, X. Dong, L. Zheng, and Y. Yang, "Patchshuffle regularization," 2017, https://arxiv.org/abs/1707.07103.

[29] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *Journal of Big Data*, vol. 6, no. 1, p. 60, 2019.

[30] S.-H. Wang, V. V. Govindaraj, J. M. Górriz, X. Zhang, and Y. D. Zhang, "Covid-19 classification by FGCNet with deep feature fusion from graph convolutional network and convolutional neural network," *Information Fusion*, vol. 67, pp. 208–229, 2021.

[31] X. Li, Y. Chai, W. Chen, and F. Ao, "Identification of early esophageal cancer based on data augmentation," in *2020 39th Chinese Control Conference (CCC)*, pp. 6307–6312, Shenyang, China, 2020.

[32] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: visual explanations from deep networks via gradient-based localization," *International Journal of Computer Vision*, vol. 128, no. 2, pp. 336–359, 2020.

[33] J. Kim and J. M. Kim, "Bearing fault diagnosis using Grad-CAM and acoustic emission signals," *Applied Sciences*, vol. 10, no. 6, article 2050, 2020.

[34] T. Akilan and Q. M. J. Wu, "sEnDec: an improved image to image CNN for foreground localization," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 10, pp. 4435–4443, 2020.

[35] L. Guiga and A. W. Roscoe, "Neural network security: hiding CNN parameters with guided Grad-CAM," in *Proceedings of the 6th International Conference on Information Systems Security and Privacy (ICISSP)*, pp. 611–618, Valletta, Malta, 2020.

[36] M. S. Hossain, G. Muhammad, and N. Guizani, "Explainable AI and mass surveillance system-based healthcare framework to combat COVID-I9 like pandemics," *IEEE Network*, vol. 34, no. 4, pp. 126–132, 2020.

[37] A. B. Tosun, F. Pullara, M. J. Becich, D. L. Taylor, J. L. Fine, and S. C. Chennubhotla, "Explainable AI (xAI) for anatomic pathology," *Advances in Anatomic Pathology*, vol. 27, no. 4, pp. 241–250, 2020.

[38] M. S. Hossain, "Cloud-supported cyber-physical localization framework for patients monitoring," *IEEE Systems Journal*, vol. 11, no. 1, pp. 118–127, 2017.

[39] M. S. Hossain and G. Muhammad, "Emotion-aware connected healthcare big data towards 5G," *IEEE Internet of Things Journal*, vol. 5, no. 4, pp. 2399–2406, 2018.

[40] S. Liu, D. Y. Liu, G. Srivastava, D. Połap, and M. Woźniak, "Overview and methods of correlation filter algorithms in object tracking," *Complex & Intelligent Systems*, vol. 23, 2020.

[41] S. Liu, W. L. Bai, G. Srivastava, and J. A. T. Machado, "Property of self-similarity between baseband and modulated signals," *Mobile Networks and Applications*, vol. 25, no. 4, pp. 1537–1547, 2020.

[42] S. Liu, X. Y. Liu, S. Wang, and K. Muhammad, "Fuzzy-aided solution for out-of-view challenge in visual tracking under IoT-assisted complex environment," *Neural Computing & Applications*, vol. 33, no. 4, pp. 1055–1065, 2021.

[43] S. Liu, S. Wang, X. Liu, C. Lin, and Z. Lv, "Fuzzy detection aided real-time and robust visual tracking under complex environments," *IEEE Transactions on Fuzzy Systems*, vol. 29, no. 1, 2020.

[44] S. Liu, C. L. Guo, F. al-Turjman, K. Muhammad, and V. H. C. de Albuquerque, "Reliability of response region: a novel mechanism in visual tracking by edge computing for IIoT environments," *Mechanical Systems and Signal Processing*, vol. 138, article 106537, 2020.

[45] X. Mao, L. Ding, Y. Zhang, R. Zhan, and S. Li, "Knowledge-aided 2-D autofocus for spotlight SAR filtered backprojection imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 11, pp. 9041–9058, 2019.

[46] X. H. Mao, X. L. He, and D. Q. Li, "Knowledge-aided 2-D autofocus for spotlight SAR range migration algorithm

Hindawi
Computational and Mathematical Methods in Medicine
Volume 2021, Article ID 8853787, 9 pages
https://doi.org/10.1155/2021/8853787

Research Article

Decision Model for Allocation of Intensive Care Unit Beds for Suspected COVID-19 Patients under Scarce Resources

Eduarda Asfora Frej, Lucia Reis Peixoto Roselli, Rodrigo José Pires Ferreira, Alexandre Ramalho Alberti, and Adiel Teixeira de Almeida

Universidade Federal de Pernambuco, Av. Acadêmico Hélio Ramos, s/n-Cidade Universitária, Recife, PE CEP 50740-530, Brazil

Correspondence should be addressed to Eduarda Asfora Frej; eafrej@cdsid.org.br

Received 22 September 2020; Revised 14 December 2020; Accepted 8 January 2021; Published 28 January 2021

Academic Editor: Alireza Jolfaei

This paper puts forward a decision model for allocation of intensive care unit (ICU) beds under scarce resources in healthcare systems during the COVID-19 pandemic. The model is built upon a portfolio selection approach under the concepts of the Utility Theory. A binary integer optimization model is developed in order to find the best allocation for ICU beds, considering candidate patients with suspected/confirmed COVID-19. Experts' subjective knowledge and prior probabilities are considered to estimate the input data for the proposed model, considering the particular aspects of the decision problem. Since the chances of survival of patients in several scenarios may not be precisely defined due to the inherent subjectivity of such kinds of information, the proposed model works based on imprecise information provided by users. A Monte-Carlo simulation is performed to build a recommendation, and a robustness index is computed for each alternative according to its performance as evidenced by the results of the simulation.

1. Introduction

The coronavirus disease 2019 (COVID-19) has been affecting the whole world and changing the routines of society in many cities. The COVID-19 disease, caused by coronavirus (SARS-CoV-2), is very contagious and has spread rapidly in many cities around the world. This disease is exhibited in different forms in the human organism and, in severe cases, results in an acute respiratory syndrome, which requires treatment in hospital Intensive Care Units (ICUs), with the support of specific equipment, such as mechanical ventilation [1–3]. The rapid spread of COVID-19 disease leads to a chaotic scenario, whereby an imbalance arises between the number of severe cases requiring treatment and the limited number of resources available to treat these patients.

In this context, the COVID-19 pandemic brings critical medical decisions related to the restriction of resources available in healthcare systems around the world, and doctors take responsibility for such decisions when deciding how to best allocate those scarce resources. Hence, the proposition of structured decision-making models based on well-founded elements of decision theory may be of great interest for health systems managers during the COVID-19 pandemic.

Therefore, this study is aimed at presenting a decision model to support an important medical decision faced daily by doctors in health system routines during the COVID-19 pandemic, considering that scarce hospital resources are available: the ICU beds allocation problem. Doctors usually make decisions based on their personal judgments and previous experience; therefore, the proposed approach is aimed at getting this knowledge and previous experiences as input for a decision model, which seeks to provide a structured and rational framework for the decision-making process with regard to critical decision problems that deal with human lives.

Decision analysis techniques have been widely used to support how to structure and resolve medical decision-making problems [4–8]. For example, Jiang et al. [6] and Xu et al. [8] use the decision tree technique to support how

best to structure the screening problem. In the same context, Janssen et al. [9] perform a literature review on methods which were used to structure the screening problem. Roselli et al. [10] tackled the screening problem for patients with suspected/confirmed COVID-19, considering a multiattribute model based on the utility theory.

Therefore, in this study, decision analysis concepts [11–14] are used to construct a utility-based model for handling the ICU bed allocation problem. A risk scenario is considered, which is related to the prognosis as to whether the patient is likely to survive or die. Also, subjected probabilities are considered [14], estimated by doctors in the form of imprecise information, the doctors being the decision-makers in these decision-problems. Due to the inherent imprecision of the information about chances of survival in different scenarios, the proposed model works based on the Monte-Carlo simulation in order to calculate a robustness index for each alternative. Therefore, the model provides doctors with recommendations regarding the medical decision problem investigated. The proposed approach is operationalized by a decision system which is freely available for use of doctors in health systems all over the world, and a database was designed to store the data of every occurrence, which will be analyzed and investigated in future research.

The main contribution of the present work relies on giving focus to the scarce resources scenario faced by health systems all over the world during the COVID-19 pandemic, which is overcharging health units and forcing doctors to make hard decisions on how to allocate those scarce resources. A specific modeling approach based on portfolio selection tools specifically directed for the COVID-19 context is developed in this paper, different from previous generic approaches present in the literature (e.g., Almeida et al. [15]). Moreover, a web-based decision information system was developed as the main product of this work, which can be easily used by health professionals all over the world.

This paper is organized as follows. "The ICU Allocation Problem" presents a brief review on the Intensive Care Units allocation problem and its relevance, highlighting its important aspects. "Portfolio-Based Approach for the ICU Allocation Problem" presents the mathematical model proposed to address the ICU allocation problem, which is based on a portfolio selection approach. "Allocation of ICU Beds under Scarce Resources: Practical Application" describes how the proposed model is operationalized, as well as a practical application of the proposed approach. Finally, "Final Remarks" presents the conclusions and final remarks of this work.

2. The ICU Allocation Problem

The ICU allocation problem has been investigated in the literature for a long time, since the resources for critical care are limited and cost-intensive. It is not uncommon to experience a situation where the number of ICU beds available is less than the number required to attend to patients who require them: the availability of this scarce resource is highly impacted by stochastic patient demands and stochastic service times, in a way that makes managing such a resource a complex problem [16, 17].

In this context, a model that aims at maximizing the expected number of lives of patients in a Pediatric ICU has been preliminary proposed by Almeida et al. [15]. Giannini and Consonni [18] investigated the perceptions and attitudes of health professionals regarding inappropriate admissions and resource allocation in the intensive care setting. They conducted a survey by applying a questionnaire to ICU doctors in Milan, Italy, and observed that 86% of them recognized that there were inappropriate admissions to ICU beds. This was due to several reasons, such as clinical doubt, limited time to make decision, errors in assessment, and pressure from superiors. Also, 5% of the respondents reported refusing appropriate admission due to financial issues. Finally, 67% reported that they frequently received requests to admit patients to ICU installations when no beds were available.

Concerned about the risk of occupational stress, Coomber et al. [19] performed a study to investigate the occupational stress of ICU doctors in the United Kingdom: they conducted a postal survey and observed that 29% of the respondents could be considered distressed, while 12% of them could be considered depressed. Thus, supporting doctors to deal with the ICU allocation problem by conducting this decision process in an easy, structured and rational way, may be very opportune, since this decision process is not trivial and is undertaken in a stressful situation.

Therefore, some studies presented in the literature propose techniques or models to support the ICU allocation problem. He et al. [20] presented a systematic literature review of research design and modeling techniques to support inpatient bed management. The authors recognized the complexity of this problem, which is affected by several factors, such as uncertainties about the patients' length of stay, fluctuations in demands, and unexpected admissions. They verified that simulation has been the main tool used in studies in this area. Reiz et al. [21] discussed the use of big data and machine learning to improve the way the ICU allocation problem is handled.

Shmueli and Sprung [22] investigated the survival benefits of allocating a patient to an ICU. The authors pointed out that in a situation of resource limitation, the policies for ICU admission should distinguish between the probability of survival and the benefits of survival. In their field study, which was undertaken in the ICU of the Hadassah-Hebrew University Medical Center during a seven-month period, the benefit was computed by using a model which considers admission and survival variables combined with the observable characteristics of patients. As a result of this study, the authors concluded that the ICU admission policy practiced until that moment in that hospital does not maximize the potential survival benefits.

On the other hand, Edenharter et al. [17] pointed out that most of the literature about the ICU allocation problem deals with the admission problem, but few studies tackle supporting the discharge decision problem. They investigated the ICU discharge problem: a univariate logistic regression model was proposed in order to assess the impact of the length of stay in the ICU, using data from two surgical ICUs

of a large academic medical center. They observed that the absence of appropriate beds in the regular ward is the main cause of the delay in ICU discharge. They emphasized that this problem is of economic and ethical relevance, since the resources of the ICU are scarce. Azcarate et al. [23] also focused on the ICU discharge problem: they present a review of the literature on patient discharge decisions and propose a simulation framework that enables the real-world processes for discharging patients to be modeled in a more realistic way.

Some studies have addressed the ethical issues inherent to the ICU allocation problem: Oerlemans et al. [24] conducted interviews with health professionals concerning ethical problems, such as how full ICU occupancy and treatment decisions are reached in terms of choosing what patients should benefit from them. Health professionals' attitudes were collected in order to provide insights to improve the management of intensive care resources. As a conclusion, the authors suggested that the collective responsibility and effort by health professionals (ICU professionals and different professionals in the wards) have to be reinforced in a hospital routine in order to alleviate moral distress caused by the ethical dilemmas faced, since these two factors are mutually dependent on each other. Consequently, health professionals have to work together for an optimal transfer of patients between hospital departments. McGuire and McConnell [25] also discuss fairness and ethics in the ICU allocation problem, suggesting that an alliance of ethical and moral principles has to be applied in order to obtain a moral, ethical, and common-sense approach to deal with this complex problem.

The ICU allocation problem acquires a special dimension when there are public health emergencies, which can be caused by several factors, such as natural disasters and major outbreaks of infectious diseases [26, 27]. Christian et al. [28], concerned about outbreaks of avian influenza (H5N1), highlighted the importance of preparing a plan for allocating resources, such as mechanical ventilators, which can become scarce during a pandemic. The authors proposed a triage protocol for allocating resources for critical care during an influenza pandemic: the protocol uses the SOFA score and has four main components: inclusion criteria, exclusion criteria, minimum qualifications for survival, and a prioritization tool. The prioritization tool they proposed determines that the highest priority for accessing ICU beds be given to patients who meet the inclusion criteria and whose probabilities of survival are greatest.

Cao and Huang [27] created a discrete event simulation model to evaluate the performance of four principles that have been often proposed as alternatives to guide the allocation of scarce resources during a public health emergency. The four principles are as follows: First Come-First Served (FCFS), which recommends the allocation of the resources to the earliest arrivals; Random Selection (RAN), which recommends the random allocation of the resources; Most Serious First (MSF), which recommends the allocation of the resources to the most seriously ill patients; and Least Serious First (LSF), which recommends the allocation of the resources to the least seriously ill patients. The authors observed that the MSF principle is intuitively favored by many authors. However, according to the results that they obtained, among the four principles evaluated, this principle performs poorest, resulting in a greater death toll. On the other hand, the LSF principle presented the best performance based on the death toll in different scarcity scenarios. However, as the authors pointed out, this principle may be problematic from an ethical perspective.

The ethical issues involved in allocating resources during a public health emergency were analyzed by Ghanbari et al. [29] in a systematic literature review. The authors observed that several clinical and nonclinical factors have been considered in protocols to prioritize patients. However, there is no clear definition about the most appropriate principles that should underpin such a prioritization. Despite this lack, the authors highlighted the importance of maintaining clear and explicit guidelines for prioritizing limited resources, in order to improve how the general public perceives the basis for such prioritization.

This review demonstrates that the ICU allocation problem is not a trivial decision problem, first, because the scarcity of resources for intensive care cannot be overcome quickly because the cost of doing so is very high and there are shortages of appropriately qualified and experienced personnel and, secondly, because the nonallocation of a place in an ICU in some cases is likely to increase the probability that the patient will die. This problem involves ethical and financial issues for which there are no instantly applicable solutions, and this dilemma becomes all the more acute during public health emergencies in which mortality rates are expected to be much higher than at other times. The current example of this is the COVID-19 pandemic which is being experienced in several regions worldwide including in countries which simply do not have the budgetary and human resources needed for intensive care. This situation leads to health professionals involved in patient care working under very stressful conditions, which can hinder rational decision-making in line with policies established by public health authorities. This paper presents important contributions for this context: a decision model based on the utilitarian principle [29] is presented for the ICU allocation problem. The approach that we propose seeks to save the largest possible number of lives, by maximizing the expected number of lives saved in all groups analyzed by defining guidelines on to which patients ICU beds should be allocated. In order to support the use of the proposed model, a system that is available online for free is presented as a tool that aids operationalizing the proposed methodology. In summary, Table 1 presents an overview to support the understanding of previous approaches related to ICU allocation issues.

3. Portfolio-Based Approach for the ICU Allocation Problem

The ICU allocation model addresses the following situation: there are n candidate patients for occupying w available ICU beds (in which $w < n$), and the doctor should decide which of these patients are going to be allocated into ICU

Table 1: Previous studies about ICU allocation problem.

Theme	Authors	Contribution
Perceptions and attitudes of health professionals	Giannini and Consonni [18]	Perceptions about inappropriate admissions and resource allocation
	Coomber et al. [19]	Occupational stress of ICU doctors
Techniques or models to support the ICU allocation problem (or similar issues)	Azcarate et al. [23]	Proposition of a framework
	He et al. [20]	Systematic literature review of research design and modeling techniques
	Reiz et al. [21]	Use of big data and machine learning
	Edenharter et al.[17]	Use of logistic regression model
	Cao and Huang [27]	Use of discrete event simulation model
	Shmueli and Sprung [22]	Application about survival benefits of allocating a patient to an ICU
	Almeida et al. [15]	Model to maximize the expected number of lives of patients in a Pediatric ICU
Ethical issues	McGuire and McConnell [25]	Discuss fairness and ethics in the ICU allocation problem
	Ghanbari et al. [29]	Systematic literature review about ethical issues involved in allocating resources
	Oerlemans et al. [24]	Interviews with health professionals concerning ethical problems
	White et al. [26]	Used ethical principles to improve allocation decisions

beds, considering their chances of survival in the ICU and outside the ICU.

Before introducing the mathematical model, an important issue that should be highlighted here is how the input data is given by the users. The input given by the doctors concerns the chances of survival of the patient in different scenarios: the user estimates the chances of survival of every candidate patient in the ICU and outside the ICU.

It is not trivial, however, for the user to provide these probability estimations. It may be hard for a doctor, even after analyzing the patient's clinical state and symptoms, to establish probabilities of survival in these different scenarios. This information is extremely subjective and may be imprecise. This type of medical decisions inherently involve uncertainty into the model. Such uncertainties sometimes derive from a random pattern of the variable being analyzed, as well as lack of knowledge and/or lack of understanding about a future condition. De Almeida et al. [14] list some factors from which uncertainties may arise, such as inaccuracy of measurement techniques, lack of details, and lack of data, among others. For the medical decision problem treated in this paper, there is uncertainty related to what will happen to a patient's life, depending on the treatment conduction adopted with him/her. Therefore, probabilities of survival and death of a patient in certain treatment conditions should be estimated.

In decision theory, a key element of many decision problems is the prior probability of the state of nature (θ), and the so-called prior probability distribution ($\pi(\theta)$) is shown as a convenient manner to quantify this information [30]. In the context of our decision-making problem, two states of nature are possible: patient survives or patient does not survive. Considering that two possible alternatives are available for such patient (allocate or not allocate an ICU bed for that patient), the prior probabilities for this problem can be represented as in Table 2.

The role of experts' knowledge in this process is crucial, since their experience about the variables of the decision problems can be used to estimate those prior probabilities [31]. According to Garcez et al. [32], a purely frequentist notion of probability cannot be applied in some cases, because some events are very rare, and therefore, their repetition is difficult to be predicted, especially when historical data is insufficient. Hence, it becomes impractical to estimate frequentist probabilities in such cases. According to Berger [30], subjective probabilities are not correct or accurate probabilities, but a measure of the degree of beliefs of the experts about the chance of occurrence of a particular event.

Clemen and Winkler [33] state that factors that have influence on the probabilities should be correlated with technical characteristics of the analyzed system. In order to do so, all the experience acquired by experts should by applied, considering their knowledge and expertise regarding the system. In this way, experts are able to provide valuable and insightful information for the decision problem being treated.

Experts' prior knowledge should, therefore, be elicited in order to be useful for the decision problem. Kadane and Wolfson [34] claim that the main purpose of the elicitation is to gather the main characteristics of the opinion of these experts and therefore to integrate their academic knowledge and previous experiences. Frequentist inference allows the interpretation of probabilities, while the Bayesian approach for statistics is completely based on subjective or personal interpretations of probabilities [35].

Table 2: Representation of prior probabilities.

	θ_1 = *Patient Survives*	θ_2 = *Patient does* not *survive*
Patient *i* allocated to ICU	$\pi_i(S_{\text{IN}})$	$\pi_i(D_{\text{IN}})$
Patient *i* not allocated to ICU	$\pi_i(S_{\text{OUT}})$	$\pi_i(D_{\text{OUT}})$

Therefore, the proposed approach considers probabilities of survival of a patient inside and outside ICU as prior probabilities, which are estimated by doctors, who act as experts in this case, considering their prior knowledge and experience about the situation. Hence, our model assumes that the doctor will be able to specify a measure the probability of survival of each patient i in the ICU ($\pi_i(S_{\text{IN}})$) and outside the ICU ($\pi_i(S_{\text{OUT}})$).

However, we recognize that such information inherently involves imprecision on its estimation. Therefore, our approach considers ranges of probabilities, instead of exact values of probabilities. Before designing the form of these input data, three doctors acted as specialists for this research and gave their opinion regarding the way in which they feel more comfortable and self-confident about providing such information. All of them stated that expressing these chances in a verbal scale makes them much more secure and comfortable than providing numbers does.

The combination of multiple experts' knowledge has advantages that were listed by Winkler et al. [36]. First, combined probability distributions leads to a better result than a single probability distribution ("two heads are better than one," according to the psychological point of view). Second, the final probability distribution may be considered a way of agreement between different experts' knowledge. Finally, the analysis becomes more complete when several opinions are considered.

In this sense, a 5-point Likert scale (*very low*, *low*, *medium*, *high*, and *very high*) was built for users to estimate chances of survival in each specific scenario, as a consensus reached by those three doctors. Each level of this scale is then converted into a range of probabilities of survival. The option *very low* means the chance of survival varies from 0 to 20%; the option *low* covers from 20% to 40%; a *medium* chance is from 40% to 60%; a *high* chance of survival covers from 60% to 80%; and *very high* means that the patient survives with 80% to 100% probability. Within those ranges, a Monte-Carlo simulation is conducted for generating a recommendation for the user, based on a robustness index of each alternative, which is detailed later on in this paper.

As previously mentioned, the input data of this model are the probability of survival of each patient i in the ICU ($\pi_i(S_{\text{IN}})$) and outside the ICU ($\pi_i(S_{\text{OUT}})$). The users estimate these chances of survival using a verbal scale, and probabilities are given in ranges of 20% (quintiles) that are derived from levels of a 5-point Likert scale. The probabilities of death in and outside the ICU can be obtained by one minus the respective probability of survival.

The utility of survival in the ICU ($U(S_{\text{IN}})$) and outside the ICU ($U(S_{\text{OUT}})$) are parameters of the model which are considered the same for every patient, since the lives of all of them have the same value for the doctor/user. The utilities of death in the ICU $U(D_{\text{IN}})$ and outside the ICU $U(D_{\text{OUT}})$ are also parameters considered the same for every patient.

Let $X_i (i = 1, \cdots, n)$ be a binary decision variable, which indicates whether patient i goes to the ICU ($X_i = 1$) or patient i does not go to the ICU ($X_i = 0$). Then, the expected utility of patient i when he/she goes to ICU ($U_i(X_i = 1)$) can be calculated as per Equation (1), and the expected utility of patient I when he/she does not go to ICU $U_i(X_i = 0)$ can be calculated using Equation (2).

$$U_i(X_i = 1) = \pi_i(S_{\text{IN}})\, x\, U(S_{\text{IN}}) + \pi_i(D_{\text{IN}})\, x\, U(D_{\text{IN}}), \tag{1}$$

$$U_i(X_i = 0) = \pi_i(S_{\text{OUT}})\, x\, U(S_{\text{OUT}}) + \pi_i(D_{\text{OUT}})\, x\, U(D_{\text{OUT}}). \tag{2}$$

In (1) and (2), the probabilities of death in ICU and outside ICU are calculated according to (3) and (4), respectively.

$$\pi_i(D_{\text{IN}}) = 1 - \pi_i(S_{\text{IN}}), \tag{3}$$

$$\pi_i(D_{\text{OUT}}) = 1 - \pi_i(S_{\text{OUT}}). \tag{4}$$

Given these utilities of staying inside and outside the ICU for each patient, the overall utility of the ICU allocation ($U_{\text{ICU}}(X_1, X_2, \cdots, X_n)$) can be calculated based on Equation (5), by the sum of the expected utility of each patient.

$$U_{\text{ICU}}(X_1, X_2, \cdots, X_n) = \left[\sum_{i=1}^{n} U_i(X_i = 1)\, x\, X_i\right] + \left[\sum_{i=1}^{n} U_i(X_i = 0)\, x\, (1 - X_i)\right]. \tag{5}$$

In order to find the combination of patients that maximize the overall utility of the ICU allocation ($U_{\text{ICU}}(X_1, X_2, \cdots, X_n)$), a binary integer linear optimization model is run ((6)–(8)). The constraints of the optimization model are Equations (7) and (8). Equation (7) guarantees that the sum of patients that go to the ICU does not exceed the number of ICU beds (w), since Equation (8) imposes that the decision variables are binary.

$$Max\, U_{\text{ICU}}(X_1, X_2, \cdots, X_n), \tag{6}$$

s.t.

$$\sum_{i=1}^{n} X_i = w, \tag{7}$$

$$X_i \in \{0, 1\} \forall i = 1, \cdots, n. \tag{8}$$

This optimization model recalls a portfolio selection model, and the output of this is the optimal combination of patients that should go to the ICU.

The design rationale for our bed allocation approach relies mainly on maximizing the number of lives saved,

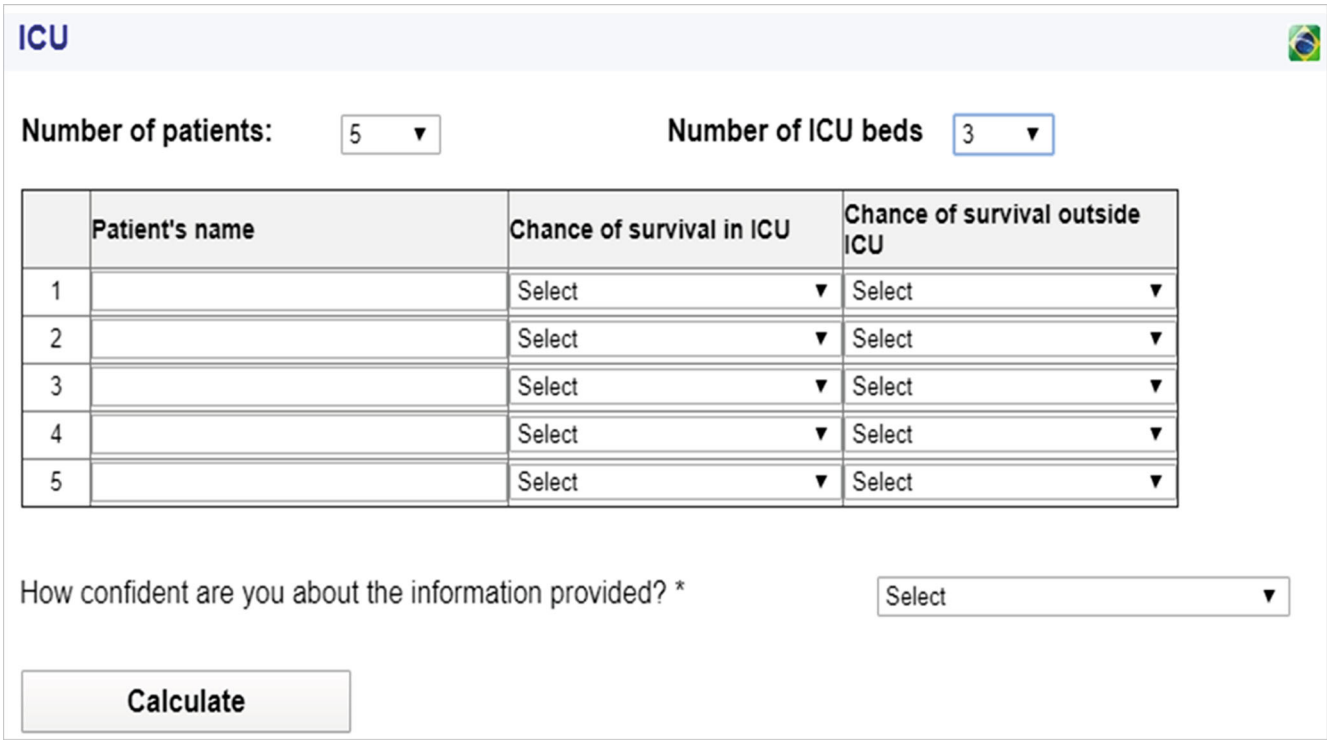


Figure 1: Input data for the ICU allocation problem.

considering a scenario in which the health system is overcharged due to the effects of the COVID-19 pandemic. In order to do so, we consider a portfolio selection approach that takes into account expected utility concepts. The objective is to maximize the overall utility of the system, considering subjective probabilities of survival of each patient inside ICU and outside ICU. In order to model such situation, the survival scenario is considered the best situation, and therefore, the utility of survival in the ICU and outside the ICU is considered to have the maximum value of utility within the considered scale. In an analogous manner, the utility of death inside the ICU and outside the ICU is considered to have the lowest utility value within the considered scale. Our approach is based on the utilitarian principle that treats the lives of all patients with equal importance, without distinction between them. The following section describes how the proposed model is operationalized to be applied in practical cases.

4. Allocation of ICU Beds under Scarce Resources: Practical Application

The model presented in "Portfolio-Based Approach for the ICU Allocation Problem" is operated by means of a Decision Information System, which is freely available for users at http://insid.org.br/sidtriagem/app/. The software was developed in a web-based environment, with a user-friendly interface that allows doctors to interact with the platform. The software has mainly two operation modules: the computational module and the interactive module. The computation module of the software works according the calculations detailed in "Portfolio-Based Approach for the ICU Allocation Problem," based on a Monte-Carlo simulation model. The interaction module works as explained in the following paragraphs. The system also has connection with a database, and the data of all occurrences performed are stored on it, in order to allow data analysis for future research to be performed.

It should be highlighted here that the role of the system is to act as a support tool for the users, but the final decision is always the responsibility of the user, who can choose whether or not to follow the recommendation given by the system.

Figure 1 shows the initial interface of the system. First, the user should enter the total number of candidate patients to go to the ICU (n) and the number of ICU beds available (w). Optionally, the user may enter the name of the patients. Then, the user should enter, for each patient, the chances of survival in the ICU and outside the ICU. The options given for the user are based on a 5-point Likert scale (*very low*, *low*, *medium*, *high*, and *very high*), which is converted into probability ranges of 20% (quintiles), as previously explained. Optionally, the user can also register how confident he/she feels in giving such information (*very unconfident*, *unconfident*, *neutral*, *unconfident*, *very unconfident*, or even *not applicable* (*N*/*A*)).

As previously mentioned, chances of survival in the ICU and outside the ICU are given by the user considering a 5-point Likert verbal scale, and the levels of the scale are converted into probability ranges (quintiles), as previously explained in "The ICU Allocation Problem." Therefore, a Monte-Carlo simulation is performed in order to obtain a recommendation for the user. At each simulation instance, random values for probabilities are generated according to a uniform distribution within the respective range given by

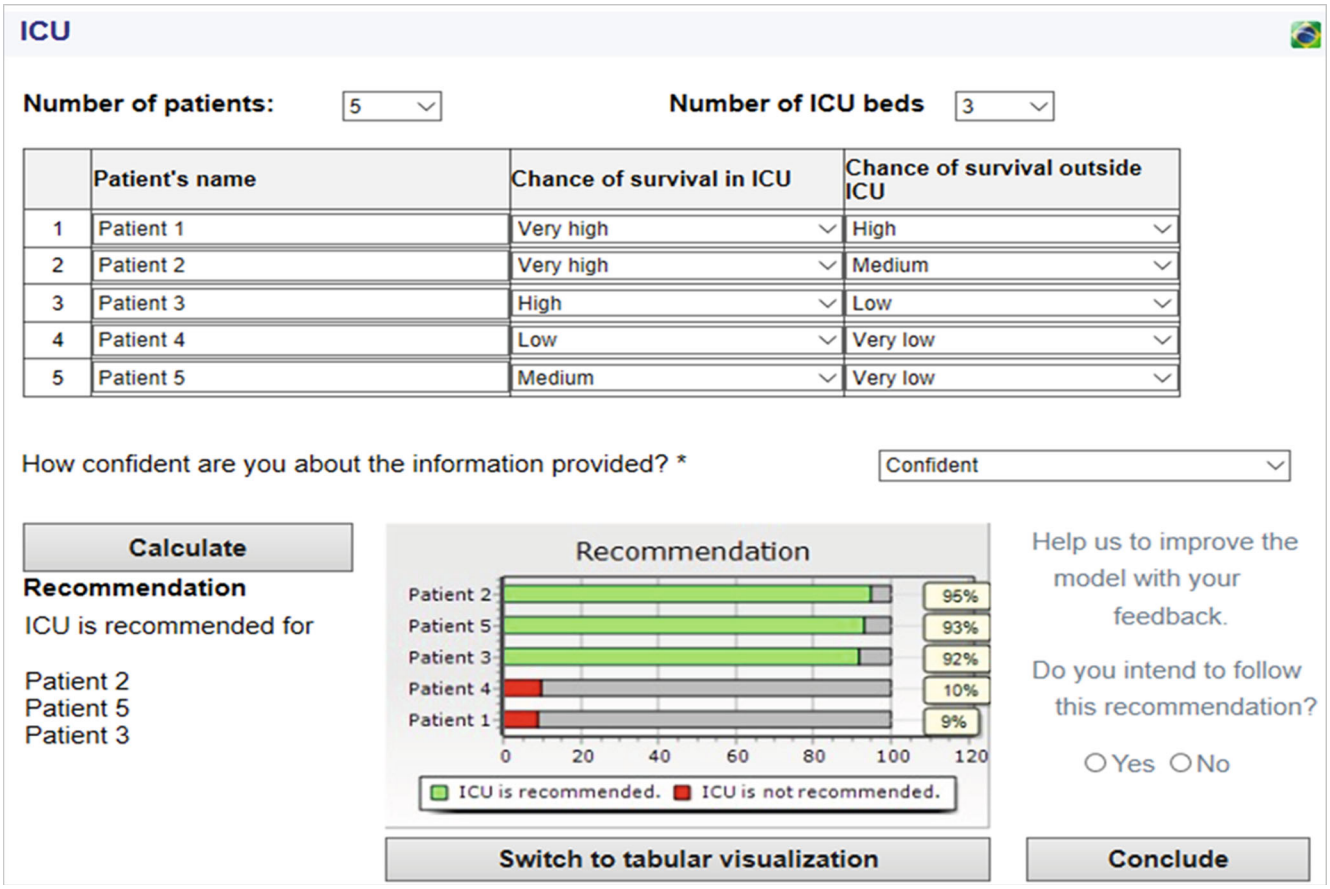


Figure 2: Practical example of the ICU allocation problem.

the input provided by the user. Then, the expected utilities are calculated for each patient and the optimization model in ((6)–(8)) is run to search for the optimal allocation of patients. At the end of the simulations, a robustness index is calculated for each patient, based on the number of simulation scenarios he/she appears in the optimal portfolio. The next topic shows a practical example to illustrate how the ICU module works.

In order to illustrate the applicability of the proposed model, let us consider a hypothetical example in which there are 5 candidate patients for only 3 available ICU beds. The user should first enter the chances of survival in the ICU and outside the ICU for each of these patients, based on the symptoms, exams, and clinical assessment of each of them. Hypothetically, let us assume that "Patient 1" has a *very high* (80%-100%) chance of survival in the ICU and a *high* (60%-80%) chance of survival outside the ICU. "Patient 2" has a *very high* (80%-100%) chance of survival in the ICU and a *medium* (40%-60%) chance of survival outside the ICU. "Patient 3" has a *high* (60%-80%) chance of survival in the ICU and a *low* (20%-40%) chance of survival outside the ICU. "Patient 4" has a *low* (20%-40%) chance of survival in the ICU and a *very low* (0-20%) chance of survival outside ICU. And "Patient 5" has a *medium* (40%-60%) chance of survival in the ICU and a *very low* (0-20%) chance of survival outside the ICU. As to the level of confidence about the information provided, let us assume the user felt "*confident*" about this. Figure 2 shows the results obtained by the model considering these input values.

The results in Figure 2 suggest that Patients 2, 5, and 3 should go to the ICU since the robustness index for them is 95%, 93%, and 92%, respectively. This means that in 92% of the simulation instances, Patient 2 was part of the optimal allocation; in 93% of the simulations instances, Patient 5 was in the optimal allocation; and in 92% of the simulation instances, Patient 3 was in the optimal portfolio. The results from the robustness index for Patient 4 and Patient 1 were 10% and 9%, respectively, which means they were part of the optimal portfolio in only 10 and 9% of the simulation instances. In the graphic of Figure 2, green bars mean the patient should go to the ICU, and red bars mean the patient should not go to the ICU. A total of 100,000 scenarios was simulated, which means that 100,000 simulation instances were performed, i.e., in each of these instances, the portfolio selection problem given by equations (6), (7), and (8) was run, and the robustness index of each patient shown in Figure 2 represents the percentage of cases in which that patient belongs to the optimal portfolio, considering all the instances performed.

The values considered for the parameters to obtain the results in Figure 2 were $U(S_{\text{IN}}) = U(S_{\text{OUT}}) = 1$, and $U(D_{\text{IN}}) = U(D_{\text{OUT}}) = 0$, but these values may be changed

whenever necessary. Parameters should be calibrated in conjunction with analysts and/or experts and can be changed in the system.

The results presented here for the ICU allocation problem are for recommendation and decision support purposes, but the final decision is always up to the doctor. The system also asks the user whether or not he/she intends to follow the recommendation provided; but the user is not obliged to answer it. A feedback space is also available for the user when clicking on the "conclude" button.

5. Final Remarks

This paper presented a utility-based portfolio selection approach to support doctors in the allocation of intensive care unit beds decision problem, in the light of the complex situation of the absence and lack of resources caused by the COVID-19 pandemic. This problem is very common in health system routines and has a special role in the complex situation brought by the pandemic. The proposed approach has potential impact related to the possibility of tackling this critical decision situation in a rational and structured way, based on proposing and implementing well-founded decision theory techniques for aiding health systems managers. Thus, this model can directly influence in the strategy to save the maximum number of patients, since the rational conduct of these medical decision processes is fundamental to enabling doctors to tackle problems arising from the COVID-19 pandemic scenario.

Moreover, giving attention to these common decision problems prompts consideration also being given to a wide variety of other applications for the proposed approach in this study. The proposed model can continue to be used after the pandemic has been brought under control, this being an important feature constructed in the pandemic period that can help doctors in their routines. It is worth mentioning, however, that the recommendations provided by this system are not normative. In other words, the proposed model is a supplement to support doctors, but it is for doctors to decide whether or not to follow the recommendations provided.

For future research, studies can be performed in order to investigate other mathematical models constructed to deal with this pandemic scenario, including the possibility of considering partial information about preferences [37]. Future works can also investigate the subjective probabilities provided by the doctors in order to perform behavioral studies.

Data Availability

The database generated from the occurrences registered in the Decision System used to support the findings of this study is restricted by the Ethical Committee in Research of the Federal University of Pernambuco with CAAE ("Certificado de Apresentação e Apreciação Ética"-Certificate of Presentation and Ethical Appreciation) number 31065820.5.0000.5208 in order to protect patient privacy.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This study was partially sponsored by the Coordination Unit for improving the Academic Qualifications of Higher Education Personnel, Brazil (CAPES), the Brazilian Research Council (CNPq), and the Foundation of Support in Science and Technology of the State of Pernambuco (FACEPE), for which the authors are most grateful.

References

[1] W.-j. Guan, Z.-y. Ni, Y. Hu et al., "Clinical characteristics of coronavirus disease 2019 in China," *New England Journal of Medicine*, vol. 382, no. 18, pp. 1708–1720, 2020.

[2] F. Zhou, T. Yu, R. Du et al., "Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study," *The Lancet*, vol. 395, no. 10229, pp. 1054–1062, 2020.

[3] M. Haghani, M. C. J. Bliemer, F. Goerlandt, and J. Li, "The scientific literature on coronaviruses, COVID-19 and its associated safety-related research dimensions: a scientometric analysis and scoping review," *Safety Science*, vol. 129, article 104806, 2020.

[4] B. A. Kiberd and K. Forward, "Screening for West Nile virus in organ transplantation: a medical decision analysis," *American Journal of Transplantation*, vol. 4, no. 8, pp. 1296–1301, 2004.

[5] K. L. Cleary, E. Paré, D. Stamilio, and G. A. Macones, "Type-specific screening for asymptomatic herpes infection in pregnancy: a decision analysis," *BJOG: An International Journal of Obstetrics & Gynaecology*, vol. 112, no. 6, pp. 731–736, 2005.

[6] B. Jiang, V. P. Ho, J. Ginsberg et al., "Decision analysis supports the use of drain amylase-based enhanced recovery method after esophagectomy," *Diseases of the Esophagus*, vol. 31, no. 10, article doy041, 2018.

[7] V. J. Ribas Ripoll, A. Wojdel, E. Romero, P. Ramos, and J. Brugada, "ECG assessment based on neural networks with pretraining," *Applied Soft Computing*, vol. 49, pp. 399–406, 2016.

[8] Y. Xu, N. S. Parikh, B. Jiao, J. Z. Willey, A. K. Boehme, and M. S. V. Elkind, "Decision analysis model for prehospital triage of patients with acute stroke," *Stroke*, vol. 50, no. 4, pp. 970–977, 2019.

[9] I. M. Janssen, A. Gerhardus, M. A. Schröer-Günther, and F. Scheibler, "A descriptive review on methods to prioritize outcomes in a health care context," *Health Expectations*, vol. 18, no. 6, pp. 1873–1893, 2015.

[10] L. R. P. Roselli, E. A. Frej, R. J. P. Ferreira, A. R. Alberti, and A. T. De Almeida, "Utility-based multicriteria model for screening patients under the COVID-19 pandemic," *Computational and Mathematical Methods in Medicine*, vol. 2020, Article ID 9391251, 8 pages, 2020.

[11] J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior*, Princeton University Press, Princeton, NJ, USA, 1944.

[12] R. L. Keeney and H. Raiffa, *Decision Analysis with Multiple Conflicting Objectives*, Wiley & Sons, New York, NY, USA, 1976.

[13] V. Belton and T. Stewart, *Multiple Criteria Decision Analysis: An Integrated Approach*, Springer Science & Business Media, Berlin, 2002.

[14] A. T. de Almeida, C. A. V. Cavalcante, M. H. Alencar, R. J. P. Ferreira, A. T. de Almeida-Filho, and T. V. Garcez, "Multicriteria and multi-objective models for risk, reliability and maintenance decision analysis," in *International Series in Operations Research & Management Science*, Springer, New York, NY, USA, 2015.

[15] L. M. A. B. Almeida, C. M. G. Miranda, and A. T. de Almeida, "Escolha de Pacientes Para UTI Pediátrica Baseada em Métodos de Apoio a Decisão (choice of patients for paediatric ICU based on decision analysis)," *Revista Brasileira de Terapia Intensiva*, vol. 13, 1, Supplement, p. 122, 2002.

[16] K. Seung-Chul, H. Ira, K. Young Karl, and A. Buckley Thomas, "Flexible bed allocation and performance in the intensive care unit," *Journal of Operations Management*, vol. 18, no. 4, pp. 427–443, 2000.

[17] G. Edenharter, D. Gartner, M. Heim et al., "Delay of transfer from the intensive care unit: a prospective observational analysis on economic effects of delayed in-house transfer," *European Journal of Medical Research*, vol. 24, no. 1, pp. 1–5, 2019.

[18] A. Giannini and D. Consonni, "Physicians' perceptions and attitudes regarding inappropriate admissions and resource allocation in the intensive care setting†," *British Journal of Anaesthesia*, vol. 96, no. 1, pp. 57–62, 2006.

[19] S. Coomber, C. Todd, G. Park, P. Baxter, J. Firth-Cozens, and S. Shore, "Stress in UK intensive care unit doctors," *British Journal of Anaesthesia*, vol. 89, no. 6, pp. 873–881, 2002.

[20] L. He, S. C. Madathil, A. Oberoi, G. Servis, and M. T. Khasawneh, "A systematic review of research design and modeling techniques in inpatient bed management," *Computers & Industrial Engineering*, vol. 127, pp. 451–466, 2019.

[21] A. N. Reiz and Organizing Committee of the Madrid 2017 Critical Care Datathon, "Big data and machine learning in critical care: opportunities for collaborative research," *Medicina Intensiva*, vol. 43, no. 1, pp. 52–57, 2019.

[22] A. Shmueli and C. L. Sprung, "Assessing the in-hospital survival benefits of intensive care," *International Journal of Technology Assessment in Health Care*, vol. 21, no. 1, pp. 66–72, 2005.

[23] C. Azcarate, L. Esparza, and F. Mallor, "The problem of the last bed: contextualization and a new simulation framework for analyzing physician decisions," *Omega*, vol. 96, article 102120, 2020.

[24] A. J. M. Oerlemans, N. van Sluisveld, E. S. J. van Leeuwen, H. Wollersheim, W. J. M. Dekkers, and M. Zegers, "Ethical problems in intensive care unit admission and discharge decisions: a qualitative study among physicians and nurses in the Netherlands," *BMC Medical Ethics*, vol. 16, no. 1, p. 9, 2015.

[25] A. McGuire and P. C. McConnell, "Resource allocation in ICU," *Current Opinion in Anaesthesiology*, vol. 32, no. 2, pp. 190–194, 2019.

[26] D. B. White, M. H. Katz, J. M. Luce, and B. Lo, "Who should receive life support during a public health emergency? Using ethical principles to improve allocation decisions," *Annals of Internal Medicine*, vol. 150, no. 2, pp. 132–138, 2009.

[27] H. CAO and S. HUANG, "Principles of scarce medical resource allocation in natural disaster relief: a simulation approach," *Medical Decision Making*, vol. 32, no. 3, pp. 470–476, 2012.

[28] M. D. Christian, L. Hawryluck, R. S. Wax et al., "Development of a triage protocol for critical care during an influenza pandemic," *Canadian Medical Association Journal*, vol. 175, no. 11, pp. 1377–1381, 2006.

[29] V. Ghanbari, A. Ardalan, A. Zareiyan, A. NEJATI, D. Hanfling, and A. Bagheri, "Ethical prioritization of patients during disaster triage: a systematic review of current evidence," *International Emergency Nursing*, vol. 43, pp. 126–132, 2019.

[30] J. O. Berger, *Statistical Decision Theory and Bayesian Analysis*, Springer Science & Business Media, New York, NY, USA, 1985.

[31] H. F. Martz and R. A. Waller, *Bayesian Reliability Analysis*, John Wiley & Sons, New York, NY, USA, 1982.

[32] T. V. Garcez, A. T. Almeida-Filho, A. T. de Almeida, and M. H. Alencar, "Experts' Elicitation of Prior Knowledge on Accidental Releases in a Natural Gas Pipeline," in *European Safety and Reliability Conference, Prague, Reliability, Risk, and Safety: Theory and Applications*, R. Bris, C. G. Soares, and S. Martorell, Eds., p. 2480, Taylor and Francis, London, UK, 2010.

[33] R. T. Clemen and R. L. Winkler, "Combining probability distributions from experts in risk analysis," *Risk Analysis*, vol. 19, no. 2, pp. 187–203, 1999.

[34] J. B. Kadane and L. J. Wolfson, "Experiences in elicitation," *Journal of the Royal Statistical Society: Series D (The Statistician)*, vol. 47, no. 1, pp. 3–19, 1998.

[35] A. O'Hagan and J. E. Oakley, "Probability is perfect, but we can't elicit it perfectly," *Reliability Engineering & System Safety*, vol. 85, no. 1-3, pp. 239–248, 2004.

[36] R. L. Winkler, S. C. Hora, and R. G. Baca, *The Quality of Expert Judgment Elicitations*, Center for Nuclear Waste Regulatory Analyses, San Antonio, TX, USA, 1992.

[37] A. T. de Almeida, E. A. Frej, and L. R. P. Roselli, "Combining holistic and decomposition paradigms in preference modeling with the flexibility of FITradeoff," *Central European Journal of Operations Research*, 2021.

Hindawi
Computational and Mathematical Methods in Medicine
Volume 2020, Article ID 8845459, 13 pages
https://doi.org/10.1155/2020/8845459

Research Article

How to Determine the Early Warning Threshold Value of Meteorological Factors on Influenza through Big Data Analysis and Machine Learning

Hui Ge,[1] Debao Fan,[2] Ming Wan,[1] Lizhu Jin,[1] Xiaofeng Wang,[1] Xuejie Du,[1] and Xu Yang[2]

[1]*Chinese Center for Disease Control and Prevention, 102206 Beijing, China*
[2]*School of Computer Science and Technology, Beijing Institute of Technology, 100081 Beijing, China*

Correspondence should be addressed to Xu Yang; yangxu@tsinghua.edu.cn

Received 11 September 2020; Revised 27 October 2020; Accepted 23 November 2020; Published 2 December 2020

Academic Editor: Kaijian Xia

Infectious diseases are a major health challenge for the worldwide population. Since their rapid spread can cause great distress to the real world, in addition to taking appropriate measures to curb the spread of infectious diseases in the event of an outbreak, proper prediction and early warning before the outbreak of the threat of infectious diseases can provide an important basis for early and reasonable response by the government health sector, reduce morbidity and mortality, and greatly reduce national losses. However, if only traditional medical data is involved, it may be too late or too difficult to implement prediction and early warning of an infectious outbreak. Recently, medical big data has become a research hotspot and has played an increasingly important role in public health, precision medicine, and disease prediction. In this paper, we focus on exploring a prediction and early warning method for influenza with the help of medical big data. It is well known that meteorological conditions have an influence on influenza outbreaks. So, we try to find a way to determine the early warning threshold value of influenza outbreaks through big data analysis concerning meteorological factors. Results show that, based on analysis of meteorological conditions combined with influenza outbreak history data, the early warning threshold of influenza outbreaks could be established with reasonable high accuracy.

1. Introduction

Infectious diseases are a major health challenge for the worldwide population. Since their rapid spread can cause great distress to the real world, in addition to taking appropriate measures to curb the spread of infectious diseases in the event of an outbreak, proper prediction and early warning before the outbreak of the threat of infectious diseases can provide an important basis for early and reasonable response by the government health sector, reduce morbidity and mortality, and greatly reduce national losses. However, if only traditional medical data is involved, it may be too late or too difficult to implement prediction and early warning of an infectious outbreak.

Influenza, commonly known as *the flu*, is an acute respiratory illness caused by influenza viruses A and B, which is a typical infectious disease [1]. It occurs all over the world and causes considerable morbidity and mortality each year. With high transmission speed, frequent pathogen variation, and a wide range of influence, rapid response and prevention of influenza remain a serious global challenge [2, 3]. WHO estimates that influenza affects 5% to 10% of adults and more than 20% of children worldwide each year [4]. About 250,000 to 500,000 people are killed each year by influenza. If we could find a way to scientifically monitor, predict, and provide early warning of influenza, governments can be prepared to prevent the outbreak and spread of influenza as early as possible. Thus, influenza early warning has received great attention from relevant departments.

Recently, medical big data has become a research hotspot and has played an increasingly important role in public health, precision medicine, and disease prediction [5–8]. In

this paper, we focus on exploring a prediction and early warning method for influenza with the help of medical big data.

It is well known that influenza disease and transmission are closely related to seasons, regions, weather and environment, demographic factors, and human behavior, among which meteorological factors are the key factors affecting the onset of influenza in a certain area. Integrating meteorological data and influenza incidence data in a certain area through machine learning and data analysis to mine the influence and effect of meteorological factors on influenza incidence is the main focus of this paper.

Traditional epidemiological surveillance systems are likely to have delayed reporting of confirmed cases. Therefore, in this paper, we will study the relationship between the number of influenza cases in the current period and the meteorological factors before a certain period of time, so as to consider the influence of lag and delay of the epidemic monitoring system, thus exploring a more accurate correlation between meteorological factors and influenza outbreaks.

Previous research mainly used machine learning or deep learning algorithms, through correlation analysis and feature selection work; first screened out important predictors such as temperature, rainfall, and relative humidity; and then made a prediction of the number of influenza incidence, but less research has been done on the establishment of influenza early warning thresholds based on meteorological factors [9, 10].

This paper studies how to determine the early warning threshold value of meteorological factors on influenza, thus providing a way to establish prediction and early warning of an influenza outbreak.

Our contributions are as follows:

(1) Establishing a preprocessing process to integrate meteorological data and influenza incidence data

(2) Selecting important meteorological indicator features for prediction and early warning of influenza outbreaks through correlation analysis and feature construction

(3) Building a prediction and early warning method for influenza outbreaks using machine learning and constructing an early warning threshold of meteorological data for influenza outbreaks through data visualization

The following is organized as follows. Related works are presented in Section 2. Our method is discussed in detail in Section 3. Experiments and results are presented in Section 4. The conclusion is given in Section 5.

2. Related Works

With the continuous development of the medical and health industry and the strengthening of the importance of public health, more and more attention has been paid to the monitoring, prediction, and early warning of infectious diseases such as influenza in the world, and the methods and technologies used have been continuously improved.

Choi and Thacker used the ARIMA model (Autoregressive Integrated Moving Average Model) in 1981 to estimate pneumonia and influenza mortality, one of the earliest studies on time series [11].

The percentage of deaths associated with pneumonia and influenza was used as an evaluation index to quantify the impact of influenza on mortality. The experimental results showed that the ARIMA model was more specific than the rule based on the regression model. The model can predict the expected mortality of pneumonia and influenza more accurately, but the factors considered in this study are far too less.

The study of Ugarte et al. [9] in 2010 and the study of Paul and Held [10] in 2011 all adopted the method of applying statistical methods to multivariate time series of infectious disease counts. The latter also introduces specific regions and possibly space-related random effects to explain different levels of incidence or changes in the spread of pathogens across regions.

Conesa et al. used a Bayesian hierarchical Poisson model with hidden Markov structures in 2015 to detect influenza epidemics [12]. By automatically monitoring influenza-related data, they detect epidemics immediately at the outbreak and predict trends in influenza epidemics and outbreaks to generate sensitive, specific, and timely warning alerts.

Marquez and Barron have created an intelligent system to support the diagnosis of influenza using the relevant factors based on historical data of the Mexican population [13]. They proposed to support the first clinical diagnosis with machine learning methods.

Some researchers have also adopted more novel techniques or included other influencing factors to analyze such problems.

Since there are many uncertain factors affecting avian influenza outbreaks, [14] has used the classification model (OOC) to solve the task of avian influenza outbreak prediction.

Dai et al. [15] presented an unsupervised word embedding-based clustering method. They try to use Twitter data to perform surveillance of influenza.

[16] combines CDC statistics, Google Trends web search data, and King Net national medical diagnosis and consultation records to propose a linear prediction framework that demonstrates that a large amount of online social behavior information can be used to indirectly monitor influenza activity.

However, due to the limitations of the linear model itself, the prediction effect is relatively general. There are also many studies on the effect of meteorological factors on influenza-like cases.

[17] compared the model error and sample fitting accuracy of the common regression model and backpropagation neural network based on the genetic algorithm and modeled the high and low flu seasons, respectively.

[18] used artificial neural networks to predict seasonal influenza epidemics in Tehran. The dataset used contains climatic characteristics such as temperature, humidity, precipitation, wind speed, sea level pressure, and the number of

Table 1: Comparison of different influenza-related works.

Reference	Methods	Data	Goal
[11]	ARIMA	Influenza data	Predict trend
[9]	Statistical methods	Influenza data	Predict trend
[10]	Statistical methods	Influenza data	Predict trend
[12]	Bayesian	Influenza data	Predict trend
[13]	Machine learning methods	Influenza data	Support diagnosis
[14]	OOC	Influenza data	Predict outbreak
[15]	Clustering	Social media data	Monitor influenza
[19]	Linear prediction	Medical data and search data	Monitor influenza
[20]	Genetic algorithm	Influenza data	Predict trend
[21]	ANN	Climatic data and influenza data	Predict trend
[16]	LSTM	Geographical data and climatic data	Predict trend
[17]	Nonlinear regression	Meteorological data	Monitor influenza
[18]	MLP	Meteorological data	Predict trend

patients (total number of referrals and number of patients with flu-like diseases). Different loss functions are defined. The results show that the model provides a satisfactory prediction possibility.

Venna et al. proposed to use long short-term memory-(LSTM-) based multistage forecasting for influenza forecasting [16]. They try to use the LSTM method to capture the temporal dynamics of seasonal flu. And they proposed a technique to capture the influence of external variables that include geographical proximity and climatic variables such as humidity, temperature, precipitation, and sun exposure.

Based on the theory of the Generalized Additive Model (GAM) and the mathematical model based on nonlinear regression, the influence of meteorological factors on the change of influenza-like cases in Urumqi is analyzed in [17]. The results of the single-factor model showed that the difference of all influencing factors was statistically significant, and the monthly sunshine hours, monthly average relative humidity, and monthly average temperature were the risk factors that caused the change of influenza-like cases. The results of the multifactor model show that only the monthly mean relative humidity and the monthly mean temperature are statistically significant.

Jhuo et al. [18] have used the meteorological and pollution parameters and acute upper respiratory infection (AVRI) outpatient number as input to a multilayer perceptron (MLP) to predict the patient number of influenza and the associated pneumonia in the following week. The meteorological parameters they used are temperature and relative humidity, and air pollution parameters are Particulate Matter 2.5 (PM 2.5) and Carbon Monoxide (CO).

We have summarized all those works in Table 1.

3. Method

3.1. Overview. In this work, we combine the influenza incidence data and meteorological data of a province in China in the past four years, to explore an effective early warning method based on machine learning and big data algorithms, thus providing useful information for influenza prevention in other regions of China.

The whole framework consists of three main parts:

(1) *Data Preprocessing*. Including the collection of meteorological data from the internet; cleaning and integrating influenza incidence data and meteorological data; normalization and exploratory analysis of data; data tagging.

(2) *Correlation Analysis and Feature Selection*. More complex features are constructed according to domain knowledge, the importance of feature calculation is calculated by the single-factor analysis method, and the feature selection is carried out by the Filter and Embedding combination algorithm.

(3) *Model Construction*. Feed data into the decision tree model, adjust the model parameters, construct the prediction model, and optimize the prediction model.

3.2. Data Preprocessing

3.2.1. Data Collection and Data Cleaning. The meteorological data is collected from the internet. This work uses Python Requests library and crawler framework Scrapy to collect meteorological data from the National Greenhouse Data System. When crawling meteorological data, first determine the crawling area, then use the *urlencode* function to send a HTTP request to get the corresponding *station_id* of the weather station in this area, and then use this *id* as the parameter of the *getWeatherData* request, plus the required date, to send a HTTP request.

Features of collected meteorological data are shown in Table 2:

3.2.2. Data Tagging. In order to train the data models, we need to annotate the original data. The data of influenza incidence and the local meteorological data collected were integrated before tagged.

Table 2: Features of meteorological data.

Name	Meaning	Data type	Data unit
t_avg	Daily average temperature	Continuous	℃
t_max	Daily highest temperature	Continuous	℃
t_min	Daily lowest temperature	Continuous	℃
precip	Cumulative precipitation	Continuous	mm
winds_avg	Average wind speed	Continuous	m/s
winds_max	Maximum wind speed	Continuous	m/s
rh_avg	Average relative humidity	Continuous	%
rh_min	Minimum relative humidity	Continuous	%
QNE_hPa	Average air pressure	Continuous	hPa
radiation	Cumulative daily radiation	Continuous	MJ/m^2

There are two basic ideas for studying the early warning threshold of meteorological factors. One is to take the daily number of influenza incidence as the explanatory variable, that is, dependent variable, to treat and solve the problem as a regression problem in machine learning, to train the model and to predict the number of future influenza incidence, and to issue an early warning when the number of predicted cases is greater than a certain threshold.

The second is to transform the continuous number of influenza cases into discrete labels of 0 and 1 by means of specific data tagging methods. The data tagging method used here is to define the threshold of influenza outbreaks and to measure whether the current incidence represents an influenza outbreak. After the tagging is completed, the problem can be solved as a classification problem in machine learning, while when the data is predicted, the dates predicted as 1 are regarded as the dates that need to be issued an early warning.

Because influenza has typical seasonal characteristics, it is not so reasonable to compare the predicted continuous values with a specific threshold according to the first idea. And if we divide the data by season and train multiple models, it complicates the problem. Compared with the first idea, the second one is more understandable and easier to implement. After comprehensive consideration and comparison, it is decided to choose the second one as the way to solve the problem in this paper.

According to the specific problem of influenza outbreaks, this paper proposes three methods of data tagging:

(i) *Moving Percentile Method*. The moving percentile method compares the number of cases in the local current observation cycle with its corresponding historical baseline data in real time. If the number of cases occurring during the current observation cycle reaches or exceeds the warning threshold, an influenza outbreak is considered; that is, the data label is defined as 1. For example, if the number of years of retrospective history is 3 years, the calculation period is 7 days, moving by day, and the historical period rocking back and forth is two reference periods. Suppose we set an early warning threshold for influenza outbreaks to P80; set the label to 1 only if the number of cases within the current observation period (7 days) is greater than or equal to 80% of the historical baseline data; otherwise, set to 0.

(ii) *Monthly Upquartile Marking*. The monthly upquartile marking, by definition, defines the label of the data corresponding to those dates in which the number of cases per month exceeds the monthly upquartile as 1.

(iii) *Dual Cycle Daily Marking*. Through the exploratory analysis of influenza incidence data, it can be found that one year can be divided into two different cycles according to the number of cases per month. The first cycle is from November to April, which is the most frequent period of influenza; the second cycle is from May to October, which is the low stage of influenza, with an average of about 1/3 of the first cycle. Because of the large gap of influenza incidence data in two cycles, it is a reasonable way to define different data tagging methods for different cycles. The specific definitions are as follows: in the first cycle, there are more influenza cases, with days as the basic unit, and the number of cases per day greater than the upper four quartiles of this cycle is marked as 1; that is, an early warning is required; in the second cycle, the number of influenza cases is less, with days as the basic unit. When the number of cases per day is greater than the 90th percentile of the cycle, mark the data as 1.

3.3. Correlation Analysis and Feature Selection. As shown in Table 2, we have collected 10 basic features of meteorological data. In order to achieve the goal of this study, we need to use the feature construction method to process the collected basic meteorological data features to construct more complex data features, in order to explore the relationship between meteorological data and influenza outbreaks from a more comprehensive perspective. Based on the obtained basic meteorological data, we constructed 48 new meteorological data features, mainly considering the delayed effect of meteorological factors on the onset time of influenza.

The purpose of feature selection is to select relevant features that are beneficial for learning algorithms from all features while sifting out irrelevant and redundant features to prevent dimensional disaster problems. Moreover, feature selection can also reduce the difficulty of learning tasks and improve the efficiency of the model.

This work uses a combination of Filter and Embedding for feature selection. We first use Filter for feature selection, calculate the correlation between each feature and output value, remove the obviously irrelevant features, reduce the feature dimension, and then use Embedding to fuse the process of feature selection with the process of classifier learning to select features in the process of learning.

After the feature selection phase, we have selected 26 constructed features and 10 basic features to train the models.

3.4. Model Construction

3.4.1. Basic Model Construction. In this work, we want to build a model that could generate early warning of influenza

Table 3: Critical parameters for CART.

Name	Meaning	Data type	Default value
max_depth	The maximum tree depth	None	
min_impurity_decrease	The minimum impurity for node splitting	0	
min_weight_fraction_leaf	The minimum weight of a leaf node	0	
class_weight	The weight of a class	None	

outbreaks based on a combination of meteorological data and influenza incidence data through machine learning and data visualization.

The decision tree algorithm could be used for classification or regression. When the relationship between independent variables and dependent variables is nonlinear or there is an interaction between variables, the effect of the linear model will be poor, and the nonlinear model should be considered. One of the important characteristics of a decision tree algorithm compared with the SVM and BP neural network is interpretability, because the process of constructing a decision tree is equivalent to forming an if-then rule set. According to the data visualization results of the decision tree model, the threshold of meteorological conditions for influenza warning is obtained. Therefore, in this work, a decision tree is used to build the basic model.

The CART decision tree algorithm uses the Gini coefficient as the evaluation standard and replaces the logarithmic operation with the quadratic operation. The smaller the Gini coefficient, the smaller the impurity representing the characteristics, and the decision tree will preferentially select the characteristics with the smallest Gini coefficient when splitting. Compared with the entropy model-based algorithm, the computational complexity of the CART algorithm is much lower. CART only produces two branches on each node, so a binary tree is formed, and each feature can be reused. And the CART algorithm can be used to deal with continuity variables.

And as discussed before, we treat the work of generating early warning of influenza outbreaks as a classification issue; in this work, we choose the CART classification tree algorithm to build the basic model.

The CART classification tree algorithm uses the Gini coefficient to perform feature selection, as described by the following equation:

$$\text{Gini}(p) = \sum_{k=1}^{K} pk(1 - pk) = 1 - \sum_{k=1}^{K} p^2 k, \tag{1}$$

where K is the number of classes in the sample, while p_k is the probability that a sample belongs to the kth class.

Since we treat the work of generating early warning of influenza outbreaks as a 0-1 two-classification issue, Equation (1) could be further simplified as follows:

$$\text{Gini}(p) = 2p(1 - p), \tag{2}$$

where p is the probability that a sample belongs to class 0.

For the given dataset D, assume the number of classes in D as K. Define$C_k$as the number of samples that belongs to class$k$. Then, the Gini coefficient of dataset D could be calculated as follows:

$$\text{Gini}(p) = 1 - \sum_{k=1}^{K} \left(\frac{|C_k|}{D} \right)^2. \tag{3}$$

For dataset D, when the CART tree splits according to feature A, D would be divided into D_1 and D_2. Under this situation, the Gini coefficient of D would be as follows:

$$\text{Gini}(D, A) = \frac{|D_1|}{D} \text{Gini}(D_1) + \frac{|D_2|}{D} \text{Gini}(D_2). \tag{4}$$

The calculation of the Gini coefficient is much simpler than that of entropy, especially for the two-classification problem, and the loss of accuracy is also smaller. Furthermore, the decision tree generated by the CART classification tree algorithm is a binary tree. Compared with the multitree formed by other decision tree algorithms, the efficiency is undoubtedly further improved.

The flow of the CART decision tree algorithm consists of two phases: decision tree generation and decision tree pruning. We have used the Cost Complexity Pruning (CCP) strategy to direct the decision tree pruning phase in our work.

3.4.2. Model Optimization. In order to fully exploit the potential of the CART algorithm, several parameters of CART need to be optimized (as illustrated by Table 3).

The *max_depth* specifies the maximum depth of the tree; limiting this parameter can ensure that the scale of the early warning model is not too complex. The *min_impurity_decrease* represents the minimum impurity of the node splitting (i.e., Gini coefficient). Since the impurity of a node decreases when the node splits, the node stops splitting immediately when the value of the impurity is less than the value of this threshold. These parameters, especially the *max_depth* parameter of the tree, are very important to limit the size of the decision tree after splitting and reduce overfitting to improve the generalization performance of the model.

Because there are far fewer days of influenza outbreaks and early warning each year than there are no early warning days, the dataset itself has an uneven sample ratio. The sample of label = 0 occupies the majority, and the sample of label = 1 is only a few. For the machine learning model, the uneven proportion of positive and negative samples will lead to the deviation of the results; that is, the effect of the model cannot reach the best, and the accuracy of prediction is not

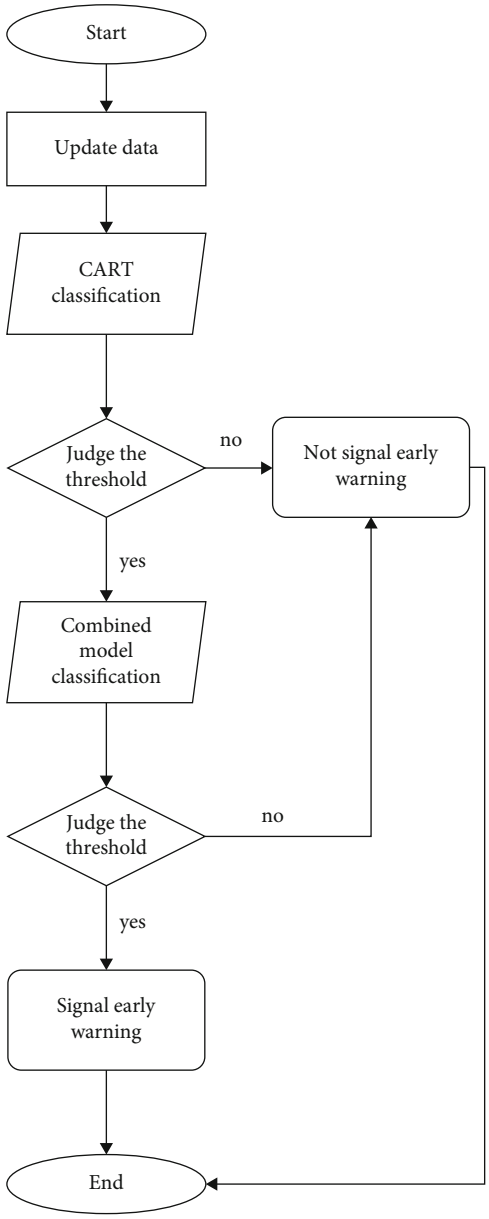


Figure 1: Flow of our method.

Table 4: Evaluation of max_depth for CART.

max_depth	ACC	f1-score	AUC
2	0.8361	0.6562	0.8019
3	0.8126	0.6793	0.7798
4	0.8135	0.7087	0.7943
5	0.7621	0.6315	0.7109
6	0.7709	0.6107	0.6954
7	0.7891	0.6051	0.6598

Table 5: Evaluation of min_impurity_decrease for CART.

min_impurity_decrease	ACC	f1-score	AUC
0	0.8135	0.7087	0.7943
0.005	0.8135	0.7087	0.7943
0.01	0.8143	0.7165	0.8029
0.02	0.8177	0.7254	0.8087
0.05	0.8268	0.7301	0.8109
0.08	0.7521	0.6342	0.7651
0.1	0.7196	0.6072	0.7535

Table 6: Evaluation of min_weight_fraction_leaf for CART.

min_weight_fraction_leaf	ACC	f1-score	AUC
0	0.8291	0.7370	0.8153
0.01	0.8043	0.6909	0.7733
0.02	0.8105	0.7144	0.7992
0.05	0.8358	0.7451	0.8208
0.1	0.8470	0.6369	0.7384
0.2	0.8578	0.6882	0.7572
0.3	0.7329	0.6153	0.7023

Table 7: Evaluation of different data tagging methods.

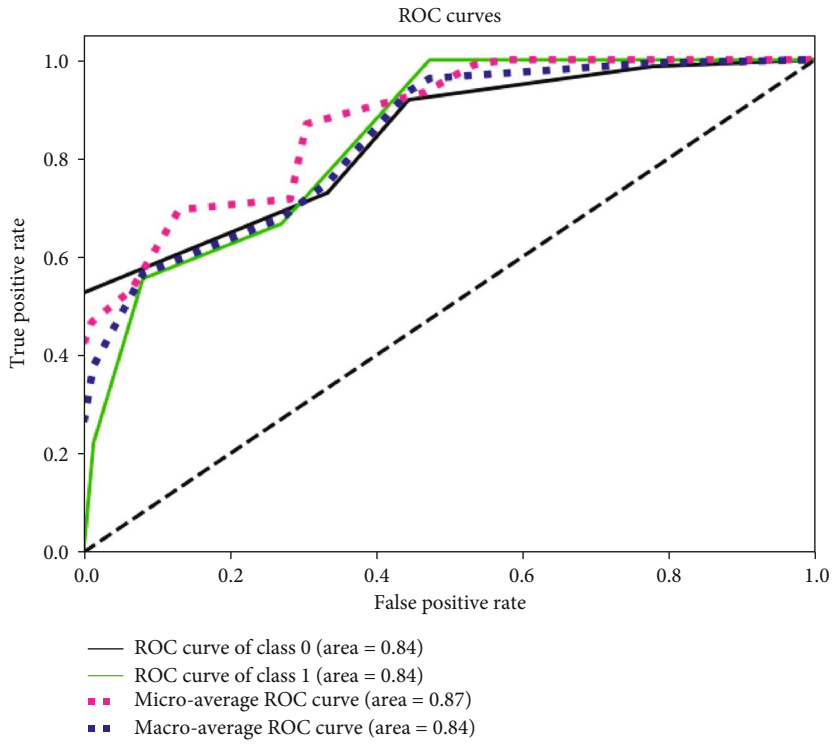
Data tagging method	ACC	f1-score	AUC
Moving percentile method	0.8586	0.7610	0.8429
Monthly upquartile marking	0.8317	0.6963	0.7967
Dual cycle daily marking	0.8391	0.7129	0.7508

good. For this purpose, it needs to be adjusted by the parameter *class_weight*. Our algorithm would calculate and give the appropriate weight to all samples in a class according to the proportion of each class in the whole sample. The *min_weight_fraction_leaf* parameter also plays an important role, since a different class has a different weight. Because upsampling will introduce a large amount of redundant data, downsampling will lose most of the information, so the most common practice is to assign different weights.

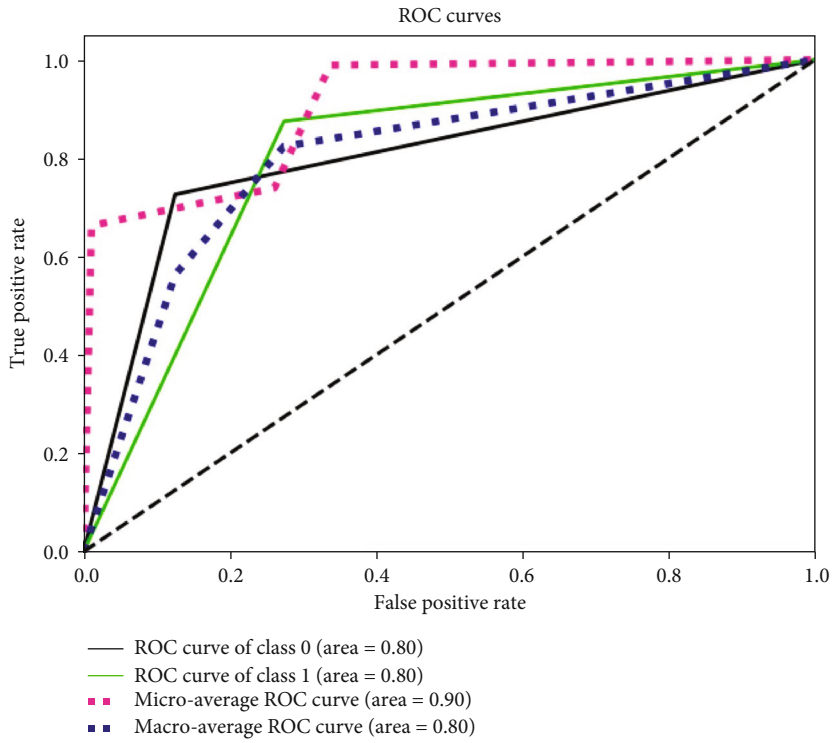
The setting of *max_depth*, *min_impurity_decrease*, and *min_weight_fraction_leaf* for the CART algorithm would be decided through experiments, which would be discussed later.

The ensemble learning method combines several simple models to form a more complex and comprehensive model. CART could be optimized through ensemble learning to enhance the stability; however, after ensemble learning optimization, it is not feasible to use a visualization method to interpret the relationship between the early warning threshold of influenza outbreaks and the certain features of meteorological data and influenza incidence data. So, we propose a method to enhance the prediction accuracy and achieve visualized interpretation of the decision of the early warning threshold of influenza outbreaks simultaneously.

Based on the idea of ensemble learning, we provided an optimized model to generate a more accurate prediction of influenza outbreaks based on meteorological data and influenza incidence data by combining CART, XGBoost, and LightGBM. XGBoost (eXtreme Gradient Boosting) is proposed by Tianqi Chen et al. in 2015, which is an optimization

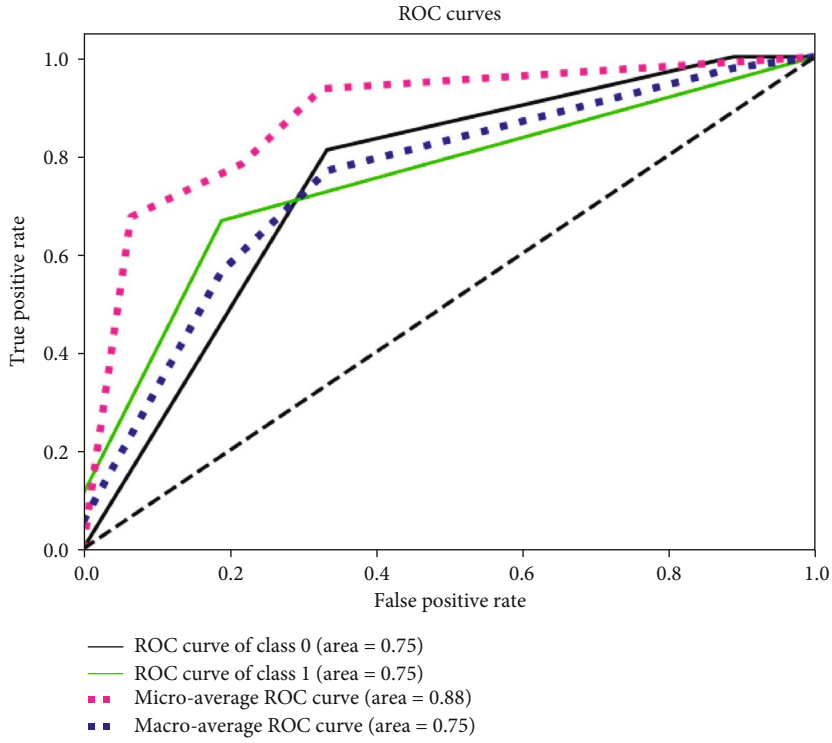


(a) ROC of moving percentile method



(b) ROC of monthly upquartile marking

Figure 2: Continued.



(c) ROC of dual cycle daily marking

Figure 2: ROC for different data tagging methods.

on GBDT. LightGBM is another optimization of GBDT, which mainly considers how to reduce the usage of memory and how to reduce the cost of multimachine communication.

The flow of our method is demonstrated in Figure 1.

The CART basic model is used to decide the early warning threshold of influenza outbreaks through data visualization. And if the CART basic model predicts that the early warning threshold is reached according to meteorological data and influenza incidence data, then the combination model formed is used to predict, and if the combination model decides that indeed the early warning threshold is reached, then our model will signal the early warning.

This method reduces the probability of prediction errors in the CART model used alone, but when the meteorological and influenza incidence big data meet the warning conditions of the CART model, they would be sent into the combination model for prediction.

The operation efficiency of the model can be greatly improved. The complexity of this algorithm is comparable with the basic CART algorithm, which is $O(\log N)$, where N represents the number of samples in the training set.

4. Experimental Results

4.1. Experimental Framework. We built our experimental framework using Python 3.5.5. The Hold-Out method divides the dataset D into two mutually exclusive subdatasets D_1 and D_2, trains the model on D_1, and tests the effect of the model on D_2. The Hold-Out method is a common method to verify model parameters and evaluate the model effect. Generally speaking, the sample size included in D_1 should account for at least 2/3 of the D of the entire dataset. In practice, there is a widely used Hold-Out method [22]: when the data has obvious time series factors, the time of online data is after the offline dataset. In this case, the training set and test set should be divided according to time.

In this work, we comply with the method. Since we have the data for a total of five years from 2012 to 2016, we divided the data from 2012 to 2015 as the training set and the data in 2016 as the test set. Dividing the dataset by year does not destroy the characteristics of the original data, preserves the characteristics of the data distribution to the greatest extent, and avoids the introduction of noise in the segmentation of the data.

The measurement metrics we used in this paper are as follows:

(1) *ACC.* Accuracy represents the ratio of the number of samples with the same predicted value as the actual value to the total sample. When the accuracy of the model is higher, it shows that the model prediction results perform better.

(2) *f1-score.* f1-score is a more combined metric, which could be calculated as $\text{f1-score} = 2 \times (\text{precision} \times \text{recall})/(\text{precision} + \text{recall})$. While recall is calculated as $\text{recall} = \text{TP}/(\text{TP} + \text{FN})$, where TP represents the number of True Positive samples, while FN is the number of False Negative samples. And precision is calculated as $\text{precision} = \text{TP}/(\text{TP} + \text{FP})$, where FP is the number of False Positive samples.

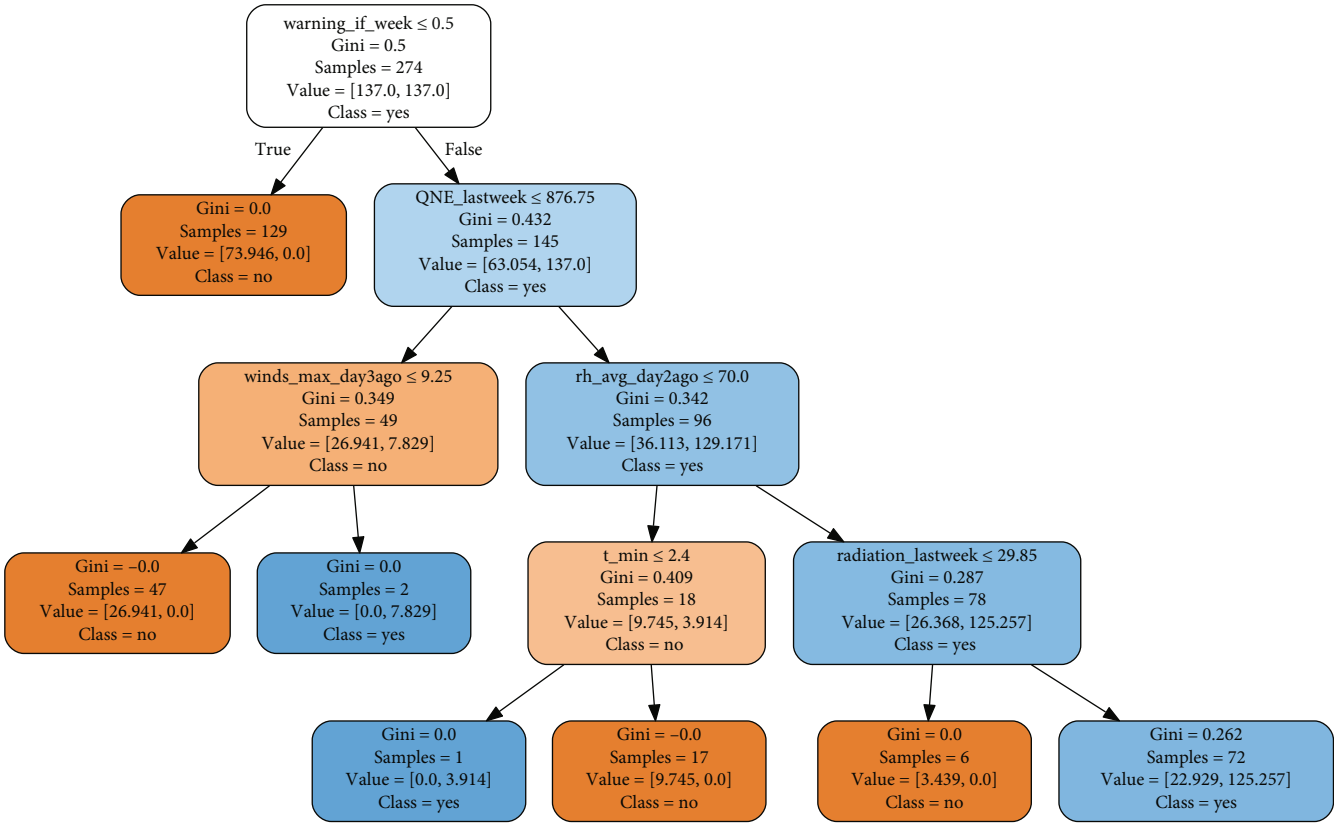


Figure 3: Data visualization of the CART model.

(3) *AUC (Area Under Curve).* AUC is often used to evaluate a two-classification model. AUC reflects a probability value that can intuitively quantify the performance of this classifier. The larger the AUC value, the better the performance of the classifier, and the maximum value is not more than 1. AUC is relatively stable and can better measure the performance of the classifier, that is, the early warning model.

4.2. Decision of CART Parameters. Experiments are conducted to decide several most important parameters for CART, as shown in Table 3. The moving percentile method is used to perform data tagging.

4.2.1. Decision of max_depth. Experimental results for different *max_depth* are shown in Table 4.

It could be seen from the results that when the $max\_depth \leq 4$, the ACC is higher. When the $max\_depth > 4$, the ACC reduces. f1-score reaches the maximum number when the $max\_depth = 4$.

And the AUC is also relatively high when the $max\_depth = 4$. f1-score and AUC reduce as *max_depth* becomes larger than 4. We could conclude that if the *max_depth* is larger than 4, the model might become overfitting. Thus, we decide that the setting of the $max\_depth = 4$.

4.2.2. Decision of min_impurity_decrease. Experimental results for different *min_impurity_decrease* are shown in Table 5.

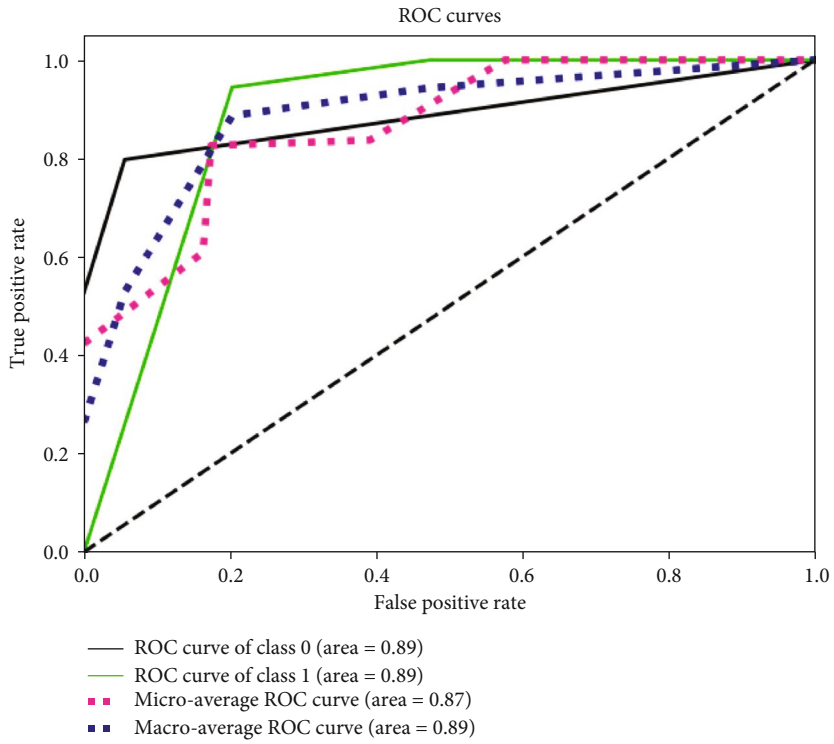
Table 8: Comparison between our model and baseline models.

Method	ACC	f1-score	AUC
Optimized model	0.8721	0.7381	0.8709
CART	0.8586	0.7610	0.8429
XGBoost	0.8804	0.6998	0.8561
LightGBM	0.8735	0.7321	0.8224

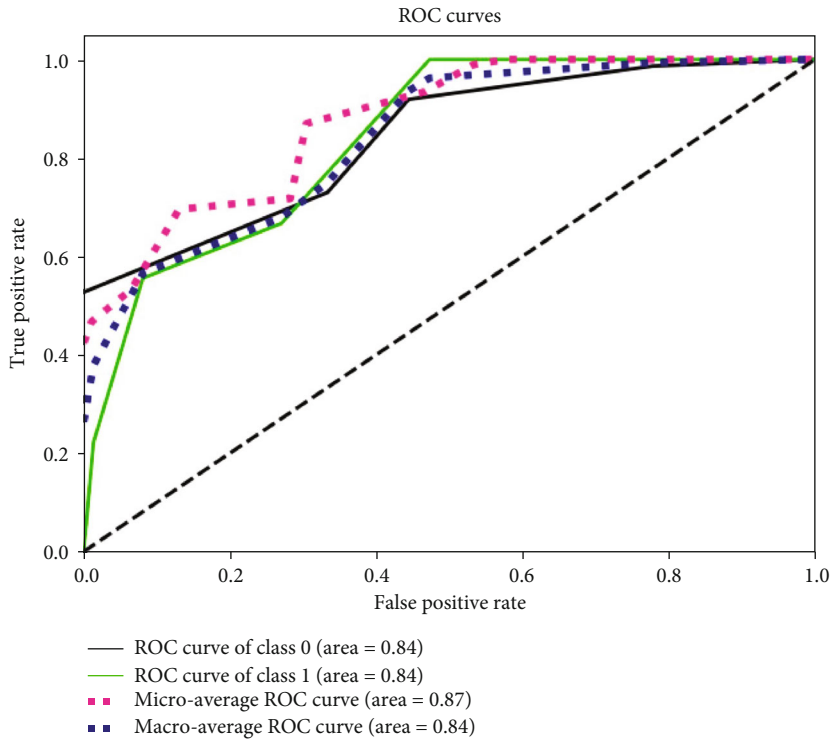
It could be seen that with the increase of *min_impurity_decrease*, ACC, f1-score, and AUC show the trend of first increasing and then decreasing. After the value of *min_impurity_decrease* is greater than 0.08, ACC, f1-score, and AUC all have a large decline. Therefore, it can be judged that the model has the best effect when the parameter is in the range of 0.02 to 0.08. After further evaluation, finally, we set $min\_impurity\_decrease = 0.04$.

4.2.3. Decision of min_weight_fraction_leaf. Experimental results for different *min_weight_fraction_leaf* are shown in Table 6.

It could be seen that when the $min\_weight\_fraction\_leaf = 0.05$, f1-score and AUC all reach the maximum value, while ACC is relatively high. Although ACC increases as min_weight_fraction_leaf increases when *min_weight_fraction_leaf* is larger than 0.05, both f1-score and AUC decline largely. Thus, we could conclude that the model gets the best effect when *min_weight_fraction_leaf* is around 0.05. After further evaluation, we set $min\_weight\_fraction\_leaf = 0.062$.

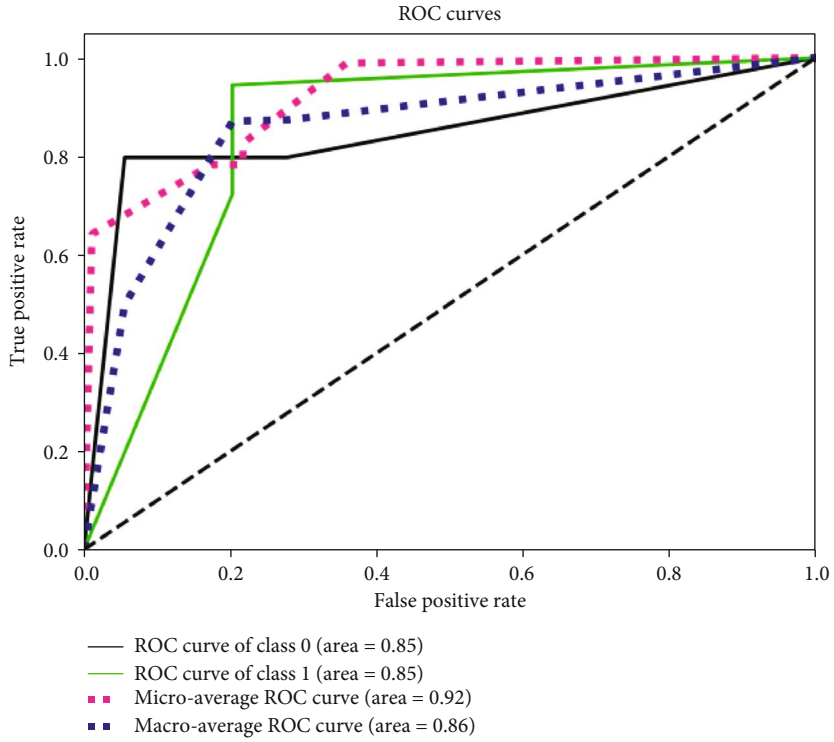


(a) ROC of our optimized model

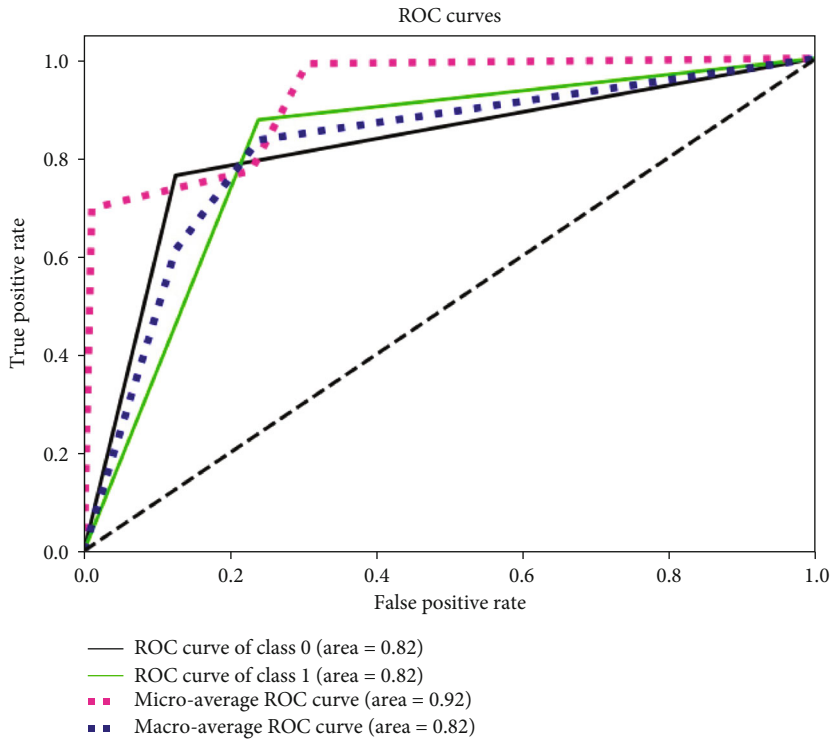


(b) ROC of the CART basic model

Figure 4: Continued.



(c) ROC of the XGBoost basic model



(d) ROC of the LightGBM basic model

Figure 4: ROC for different basic models.

4.3. Evaluation of Data Tagging Methods. An experiment is conducted to evaluate the best data tagging method for our model. Results are shown in Table 7 and Figure 2.

Through the comparison, we could decide that the moving percentile method is more suitable for our model.

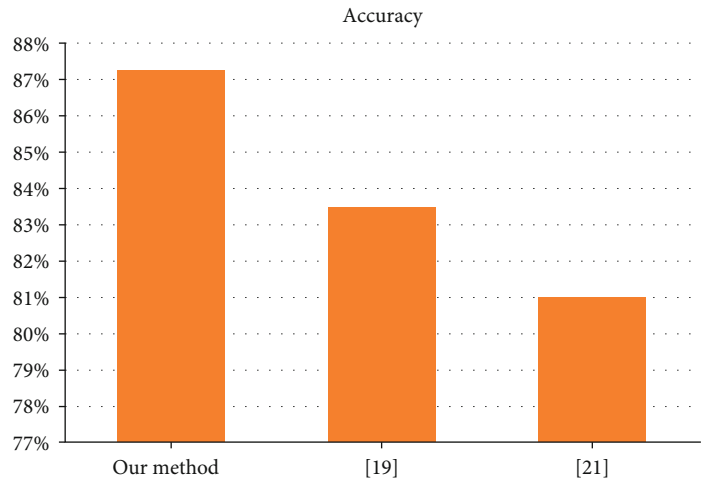


Figure 5: Comparison with other algorithms.

4.4. Definition of the Early Warning Boundary Value of Meteorological Factors on Influenza. The visualization results of the CART basic model are shown in Figure 3.

As we said before, according to the construction process of the decision tree model, the classification rules can be seen intuitively from the tree structure diagram, and then the meteorological conditions need to be issued when an early warning is given.

Thus, we could generate the early warning boundary value of meteorological factors on influenza based on using the moving percentile tagging method with the CART basic model from Figure 3 as follows: (1) (warning_if_week = 0) and (QNE_lastweek ≤ 876.75 hPa) and (winds_max_day3 ago > 9.25 m/s); (2) (warning_if_week = 0) and (QNE_lastweek > 876.75 hPa) and (rh_avg_day2ago ≤ 70%) and (t_min ≤ 2.4°C); and (3) (warning_if_week = 0) and (QNE_lastweek > 876.75 hPa) and (rh_avg_day2ago > 70%) and (radiation_lastweek > 29.85 MJ/m).

It is easy to see that the three conditions are mutually exclusive and only one of them will be satisfied at most. When one condition is satisfied, an early warning is issued.

4.5. Evaluation of the Optimized Model. We use the moving percentile method as the data tagging method. And the comparison between our optimized model and the baseline models is shown in Table 8 and Figure 4.

It could be seen that ACC and AUC of the optimized model are better than those of the CART basic model. But f1-score of the optimized model is smaller than that of the CART basic model. According to our analysis, the mechanism of the optimized model makes the number of samples predicted as 1 become less; thus, the recall rate becomes lower and the f1-score becomes lower. The ACC of CART is relatively low, but the f1-score and AUC are relatively high. The XGBoost model performs well in accuracy and AUC, but the f1-score is relatively low. The LightGBM model is slightly poor in AUC, and the ACC and f1-score tend to be intermediate.

We have also shown a comparison of the accuracy between our method and some state-of-the-art methods in Figure 5.

5. Conclusion

In this paper, we try to combine meteorological data and influenza incidence data to build a big data model to determine the early warning boundary value of meteorological factors on influenza. We exploit the data visualization method on the CART basic model to provide a way to generate an early warning threshold for influenza outbreaks based on data analysis of meteorological data. We proposed an optimized model to generate a more accurate early warning signal.

Our approach comes at the expense of slightly reducing the recall rate to improve ACC and AUC and also making full use of the results of the CART model via data visualization. Only when the CART basic model indicates that maybe an early warning should be signaled, then the more complex combination model of XGBoost and LightGBM would be needed. Overall, it is a reasonable scheme according to the evaluation.

Another strategy might be to take the "OR" operation for the construction of the combination optimized model. When at least one model is predicted to be 1, the final prediction result is 1; that is, an early warning is needed. However, the early warning threshold could not be generated through data visualization, thus without interpretability. Also, the computation effort is more. Under realistic conditions, different model combination strategies can be selected according to different needs.

Actually, in this work, we have only introduced key meteorological factors, while the influenza outbreak is also closely related to human flow, intercity migration index, vaccination, emergencies, and other factors. In future work, we would try to establish a more comprehensive way to establish the early warning system for influenza outbreaks.

Data Availability

Requests for data (6/12 months after publication of this article) will be considered by the corresponding author.

Disclosure

The funders had no role in the design of the study; in the collection, analysis, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Conflicts of Interest

The authors declare no conflict of interest.

Authors' Contributions

The contributions of the authors involved in this study are as follows: conceptualization, H.G. and D.F.; methodology, M.W.; software, L. J.; validation, X.W., D.F., and M.W.; formal analysis, H.G.; investigation, L.J.; resources, X.Y.; data curation, D.F. and X. D.; writing–original draft preparation, H.G.; visualization, X.W.; supervision, X.Y.; project administration, X.Y.; and funding acquisition, H.G. and X.Y. All authors have read and agreed to the published version of the manuscript.

Acknowledgments

This work is supported by the National Science and Technology Major Project of China under Grant No. 2018ZX10201-002 and the National Natural Science Foundation of China under Grant No. 91846303.

References

[1] WHO, "WHO report on global surveillance of epidemic-prone infectious diseases - influenza," 2020, https://www.who.int/csr/resources/publications/influenza/CSR_ISR_2000_1/en/.

[2] M. Robinson, Y. Drossinos, and N. I. Stilianakis, "Indirect transmission and the effect of seasonal pathogen inactivation on infectious disease periodicity," *Epidemics*, vol. 5, no. 2, pp. 111–121, 2013.

[3] L. Willem, F. Verelst, J. Bilcke, N. Hens, and P. Beutels, "Lessons from a decade of individual-based models for infectious disease transmission: a systematic review (2006-2015)," *BMC Infectious Diseases*, vol. 17, no. 1, p. 612, 2017.

[4] S. C. Schoenbaum, "The great influenza: the epic story of the deadliest plague in history," *Journal of Public Health Policy*, vol. 25, no. 3-4, pp. 435–443, 2004.

[5] Q. Yao, Y. Tian, P.-F. Li, L.-L. Tian, Y.-M. Qian, and J.-S. Li, "Design and development of a medical big data processing system based on Hadoop," *Journal of Medical Systems*, vol. 39, no. 3, pp. 1–11, 2015.

[6] W. Lin, W. Dou, Z. Zhou, and C. Liu, "A cloud-based framework for home-diagnosis service over big medical data," *Journal of Systems & Software*, vol. 102, pp. 192–206, 2015.

[7] C. H. Lee and H.-J. Yoon, "Medical big data: promise and challenges," *Kidney Research and Clinical Practice*, vol. 36, no. 1, pp. 3–11, 2017.

[8] D. Windridge and M. Bober, "A kernel-based framework for medical big-data analytics," in Springer Berlin Heidelberg, 2014.

[9] M. D. Ugarte, T. Goicoa, and A. F. Militino, "Spatio-temporal modeling of mortality risks using penalized splines," *Environmetrics*, vol. 21, no. 3-4, pp. 270–289, 2010.

[10] M. Paul and L. Held, "Predictive assessment of a non-linear random effects model for multivariate time series of infectious disease counts," *Statistics in Medicine*, vol. 30, no. 10, pp. 1118–1136, 2011.

[11] K. Choi and S. B. Thacker, "An evaluation of influenza mortality surveillance, 1962-1979. II. Percentage of pneumonia and influenza deaths as an indicator of influenza activity," *American Journal of Epidemiology*, vol. 113, no. 3, pp. 227–235, 1981.

[12] D. Conesa, M. A. Martínez-Beneito, R. Amorós, and A. López-Quílez, "Bayesian hierarchical Poisson models with a hidden Markov structure for the detection of influenza epidemic outbreaks," *Statistical Methods in Medical Research*, vol. 24, no. 2, pp. 206–223, 2011.

[13] E. Marquez and V. Barron, "Artificial intelligence system to support the clinical decision for influenza," in *2019 IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC)*, pp. 1–5, Ixtapa, Mexico, November 2019.

[14] J. Zhang, J. Lu, and G. Zhang, "Combining one class classification model for avian influenza outbreaks," in *2011 IEEE Symposium on Computational Intelligence in Multicriteria Decision-Making (MDCM)*, pp. 190–196, Paris, France, April 2011.

[15] X. Dai, M. Bikdash, and B. Meyer, "From social media to public health surveillance: word embedding based clustering method for Twitter classification," *SoutheastCon 2017*, 2017, pp. 1–7, IEE, Charlotte, NC, USA, 2017.

[16] S. R. Venna, A. Tavanaei, R. N. Gottumukkala, V. V. Raghavan, A. S. Maida, and S. Nichols, "A novel data-driven model for real-time influenza forecasting," *IEEE Access*, vol. 7, pp. 7691–7701, 2019.

[17] G. Feng-Yun, F. Xu-Cheng, Y. Jian-Dong, and W. Kai, "Influence of meteorological factors on influenza-like cases in Urumqi," *Occupation and Health*, 2019.

[18] S. Jhuo, M. Hsieh, T. Weng, M. Chen, C. Yang, and C. Yeh, "Trend prediction of influenza and the associated pneumonia in Taiwan using machine learning," in *2019 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS)*, pp. 1-2, Taipei, Taiwan, December 2019.

[19] T.-H. Chen, Y.-C. Chen, J.-L. Chen, and F.-C. Chang, "Flu trend prediction based on massive data analysis," in *2018 IEEE 3rd International Conference on Cloud Computing and Big Data Analysis (ICCCBDA)*, pp. 304–308, Chengdu, China, April 2018.

[20] H. Xue, Y. Bai, H. Hu, and H. Liang, "Influenza activity surveillance based on multiple regression model and artificial neural network," *IEEE Access*, vol. 6, pp. 563–575, 2018.

[21] F. Saberian, A. Zamani, M. M. Gooya, P. Hemmati, M. A. Shoorehdeli, and M. Teshnehlab, "Prediction of seasonal influenza epidemics in Tehran using artificial neural networks," in *2014 22nd Iranian Conference on Electrical Engineering (ICEE)*, pp. 1921–1923, Tehran, Iran, May 2014.

[22] A. Mihalik, F. S. Ferreira, M. Moutoussis et al., "Multiple holdouts with stability: improving the generalizability of machine learning analyses of brain-behavior relationships," *Biological Psychiatry*, vol. 87, no. 4, pp. 368–376, 2020.

Hindawi
Computational and Mathematical Methods in Medicine
Volume 2020, Article ID 8823861, 12 pages
https://doi.org/10.1155/2020/8823861

Research Article

A Multifeature Extraction Method Using Deep Residual Network for MR Image Denoising

Li Yao

School of Computing, Hubei Polytechnic University, Huangshi, Hubei 435003, China

Correspondence should be addressed to Li Yao; yaoli@hbpu.edu.cn

Received 30 September 2020; Revised 21 October 2020; Accepted 26 October 2020; Published 6 November 2020

Academic Editor: Kaijian Xia

In order to improve the resolution of magnetic resonance (MR) image and reduce the interference of noise, a multifeature extraction denoising algorithm based on a deep residual network is proposed. First, the feature extraction layer is constructed by combining three different sizes of convolution kernels, which are used to obtain multiple shallow features for fusion and increase the network's multiscale perception ability. Then, it combines batch normalization and residual learning technology to accelerate and optimize the deep network, while solving the problem of internal covariate transfer in deep learning. Finally, the joint loss function is defined by combining the perceptual loss and the traditional mean square error loss. When the network is trained, it can not only be compared at the pixel level but also be learned at a higher level of semantic features to generate a clearer target image. Based on the MATLAB simulation platform, the TCGA-GBM and CH-GBM datasets are used to experimentally demonstrate the proposed algorithm. The results show that when the image size is set to 190×215 and the optimization algorithm is Adam, the performance of the proposed algorithm is the best, and its denoising effect is significantly better than other comparison algorithms. Especially under high-intensity noise levels, the denoising advantage is more prominent.

1. Introduction

With the rapid development of information technology and computer vision, different digital imaging technologies emerge endlessly. An image has become one of the most common information transmission carriers in modern life. Up to now, the research on medical imaging has made some progress and achievements [1]. For example, methods such as positron emission tomography (PET), computer tomography (CT), magnetic resonance imaging (MRI), and functional magnetic resonance imaging (FMRI) have been used. They have become the main technical means for obtaining medical images. They have been successfully applied in clinical medical diagnosis [2]. These imaging technologies can enrich the anatomical images of clinical symptoms and can observe various lesions of the human body more clearly. It is conducive to image postprocessing and provides an important reference for clinical development. Image digital imaging technology is based on computers; collects, displays, stores, and transmits images to digitize image information; and optimizes each part individually. Electronic information, computer technology [3], and digital image processing methods are the basis of the above imaging technology. Using imaging technology to effectively complement the image information can achieve the effect of easy access to medical image information. It is helpful for clinicians to formulate treatment plans through image diagnosis acquisition, so as to accelerate the rapid development of medical research and new medical technology [4].

Among many medical imaging technologies, MRI is a high-resolution medical imaging technology for human tissues and organs. It can image various parts of the human body at multiple angles and in all directions and can obtain relatively complete medical image information [5]. For a complete image information processing system, it mainly includes modules such as information acquisition, information processing, information transmission, and information reception. However, in each of these links, the image may be polluted by the random noise of the Rician distribution. These noises reduce the signal-to-noise ratio of the image, which makes it difficult for doctors to distinguish the details of the lesion from the background [6]. In addition, noisy

images will not only affect the visual effect but also adversely affect subsequent image analysis, such as image segmentation, target recognition, and edge detection. Therefore, it is very necessary to remove the noise in an image. Due to the lack of relevant research in the field of denoising, the post-processing of an image is seriously affected, and the accuracy and effectiveness of imaging are reduced. Therefore, image denoising processing is particularly important [7].

An image denoising algorithm can effectively filter out image noise and at the same time enhance the useful information of the image [8]. That is, while eliminating image noise, the details of the image are preserved as much as possible. Image detail information is selectively strengthened and suppressed, and the edge of the image is highlighted, thereby improving the visual effect of the image [9]. When traditional image denoising methods filter medical MR image noise, it is easy to lose the edge information of the image and it is difficult to save the detailed information of the image, which is far from meeting the needs of medical diagnosis [10]. Therefore, there is an urgent need for new research methods and means to solve this problem.

2. Related Research

Image denoising is a difficult problem in the field of image processing. In order to get a clear image and effectively remove the noise introduced in the process of image generation, many domestic and foreign researchers have done a lot of research and put forward many very clever denoising algorithms. Some of the traditional image denoising methods have been proposed for a long time and have been widely used. However, most algorithms will inevitably damage the original useful information while denoising. And many algorithms need to be improved in denoising performance and model complexity [11]. How to strip out the noise and achieve a good denoising effect even under high-intensity noise has brought new challenges to the research of existing denoising algorithms. For medical MRI denoising algorithms, the more common ones are traditional image denoising algorithms, dual-domain filtering image denoising algorithms, and deep learning image denoising algorithms [12].

According to the denoising of traditional image denoising algorithms in different domains, these methods are divided into spatial domain and transform domain image denoising algorithms. In the image denoising algorithm in the spatial domain, the commonly used spatial domain filtering methods mainly include mean filtering, median filtering, and Wiener filtering [13]. The mean filter is a linear smoothing filter, which can effectively deal with the additive white Gaussian noise in the image. In Reference [14], an effective image denoising method is proposed for images damaged by salt and pepper noise. A modified mean filter (MMF) is used to restore the image by processing the value. For serious noise damage, the noise pixel value is replaced by the weighted average value of MMF sum. The change of damaged pixels can be minimized by convex optimization, and two fuzzy systems are used to determine the weight of the average value to achieve denoising. The mean filter is equivalent to a low-pass filter. Although the operation is simple and the calculation speed is fast, the mean filter will lose the details in the denoising process and make the image blurry [15]. Median filtering was originally a nonlinear processing technique used to analyze time series and was later used to remove salt and pepper noise. Wiener filtering can better filter out salt and pepper noise through the local statistical characteristics of the image according to the minimum mean square error. In Reference [16], an improved Wiener filtering method is proposed to denoise satellite images. Different noise and filtering techniques are studied in detail, and the results show that the denoising effect of the filter depends on the type of noise present in the image. Compared with the existing linear and nonlinear filtering methods, the performance of the proposed improved Wiener filter in most noise models is relatively better. But the Wiener filter needs the spectrum information of the original signal and noise, and it can only achieve good results when the signal is sufficiently smooth [17]. The spatial domain denoising algorithm directly performs corresponding processing on the image pixels. According to the type of filter, it can be divided into linear filters and nonlinear filters. Images generally have such a property: the pixel value at any position in an image is closely related to the pixel value in the neighborhood where the pixel is located, such as spatial correlation or similar pixel values [18].

For the transform domain image denoising algorithm, the image is first transformed into other domains, and then, the coefficients are processed by the properties in the transform domain. Finally, the inverse transform is used to reconstruct the coefficients to obtain the denoised image [19]. There are many ways to transform an image from the spatial domain to the transform domain, such as Fourier transform, cosine transform, and Walsh-Hadamard transform. One of the most commonly used is wavelet transform. The wavelet domain spatial adaptive FIR Wiener filtering method proposed in [20] only performs Wiener filtering in each scale. Although it can quickly remove noise, the processing of different types of noise needs to be strengthened [21].

In addition to traditional spatial and frequency domain image denoising algorithms, literature [22] proposes a dual-domain image denoising (DDID) algorithm, using the idea of combining the space domain and the frequency domain, combined with bilateral filtering and short-time Fourier transform to denoise the image hierarchically. After the algorithm is denoised, the image can retain the original image details to a greater extent and obtain better visual effects. The DDID algorithm divides the noisy image into the base layer and the detail layer to process separately and realizes the image detail enhancement with good effect. Therefore, since the algorithm was proposed, it has been widely used in image denoising. However, the boundary information on the multimode image cannot effectively suppress the negative ringing effect, and the detail preservation needs to be improved [23].

In recent years, deep learning algorithms have been rapidly developed, and the convolution neural network (CNN) method has been proposed to use a layer-by-layer initialization method to overcome difficult problems in the training

process. CNN has been widely used in the field of computer vision and has significant effects on solving image classification, target detection, and image recognition problems [24]. With the needs of research, the neural network structure is continuously deepened to build a higher and more accurate deep learning network. Because CNN can directly input the original image, it is simple and easy to use, is widely recognized by the academic community, and has been successfully used in the field of image denoising. Reference [25] comprehensively studies the most advanced image denoising methods using CNN. A denoising prior driven network (PDNN) is proposed to remove fixed-level Gaussian noise. In the BSD-68 and Set-12 datasets, PDNN shows good denoising results in terms of PSNR. In some cases, combining neural networks with traditional methods can achieve better results. For example, literature [26] uses the powerful nonlinear fitting ability of a neural network to combine a neural network with a wavelet method. A neural network is used to find the optimal coefficients of Gabor wavelet, and a neural network can adaptively select wavelet parameters; a combination of RBF network and wavelet is trained. This method has a good effect in the field of sound signal detection, and the theoretical effect of image denoising needs further practical verification.

Because the image is often contaminated by noise and becomes blurred, the image obtained by the above algorithm still has the problem of a blurred image and unsatisfactory effect. Therefore, a multifeature extraction algorithm based on the depth residual network is proposed for medical MR image denoising. The innovation of the proposed method is as follows:

(1) In order to improve the learning ability of the image denoising algorithm and reduce the time of model training, multifeature extraction technology is adopted. Three different convolution kernels of different sizes are used to extract features from the input image in a way that the center position is unchanged and synchronized to obtain richer image features and achieve a better denoising effect

(2) Due to the large number of layers of the deep neural network, backward propagation easily leads to the disappearance of the gradient, making the training result poor. The proposed algorithm introduces residual learning and learns image noise residuals based on function mapping, thereby achieving image denoising and accelerating and optimizing the deep network. Especially under high-intensity noise levels, the denoising advantage is more prominent

3. Denoising Network

3.1. Multifeature Extraction Residual Network Denoising Framework. Since image denoising is a typical restoration problem, the proposed denoising network model uses a fully convolutional layer network, which removes the pooling layer in a conventional network. The function of dimensionality reduction and parameter reduction is accomplished by learning residual mapping. The multifeature extraction residual network denoising framework is shown in Figure 1.

The first layer uses multiple feature extraction layers to extract different feature information and map the input image from image space to feature space. After that, a series of residual units are connected in series to form a chain structure, and then the residual image containing only noise information is generated by the last convolution layer reconstruction. Finally, the final denoising result is calculated by the global skip connection. For the loss function in training, the joint loss function which combines the traditional mean square error loss and the perceptual loss is used. The perceptual loss is defined and calculated by a pretrained network [27].

3.2. Residual Learning. For the general convolution neural network, image denoising problems are adopted by learning the clean image way. The image result calculated by the network is compared with the standard image, and the cost function is to minimize this error. The formula is as follows:

$$J(\theta) = \frac{1}{m}\sum_{i}^{m}\left(\frac{1}{2}\left\|h_{w,b}\left(x^{(i)}\right) - y^{(i)}\right\|^2\right), \tag{1}$$

where x stands for the standard image, y represents the image after network calculation, and m represents the number of images. For the entire network, the cost function obtained by learning noise images is determined by $h_{w,b}(x^{(i)}) - y^{(i)}$.

For the whole model, residual learning is a very important technology. Many network models predict a clean image without pollution by learning mapping functions [28]. For residual learning, it is based on $R(y) \approx v$ function mapping learning image noise residual and then realizes image denoising. Finally, by using the expected noise residual value and the estimated noise residual value calculated by the convolutional network, the mean square error is obtained, and the cost function of the entire network is obtained. The formula is as follows:

$$L(\vartheta) = \frac{1}{2n}\sum_{1}^{n}\left\|R(y_i;\vartheta) - (y_i - x_i)\right\|_F^2, \tag{2}$$

where ϑ represents the training parameter and n represents the image block divided into the entire input image. In the network model, this residual learning strategy is selected because it is easier to be optimized by convolutional neural networks and can be very well applied to deep networks. Deep networks face many difficulties, and they are generally very difficult to train. It is mainly reflected in the fact that when the number of layers increases, the backward propagation easily causes the gradient to disappear, making the training result very poor. And this deep residual network solves this problem in the neural network structure, so that the network can be very deep, and the gradient will not disappear. When the noise level is relatively low, the feature map $F(y) = x$ of learning clean images is closer to the initial identity map than the feature map $R(y) \approx v$ of residual learning. It is equivalent to the mapping relationship of $F(y) = x$ during

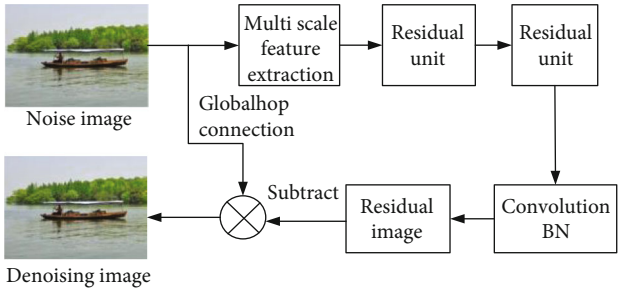


Figure 1: Multifeature extraction residual network denoising framework.

deep network training, which will make the gradient disappear more easily. Therefore, the residual training will be more conducive to the acceleration and optimization of the deep network.

3.3. Batch Normalization. Batch normalization (BN) is an adaptive heavy parameter method, which is mainly used to solve the problem of internal covariate shift (ICS) in deep learning. The model is shown in Figure 2.

The BN model selects m inputs because the whole is approximately the same distribution, and h obtained as a whole is put into the neural network for batch training. This batch processing method is very beneficial for deep learning, making the model simpler, and solving the problem of difficulty in training after the number of network layers becomes deeper.

For ICS problems that appear in deep network training, the conditional probabilities are consistent, and the marginal probabilities are inconsistent. For all $x \in X$, the relationship between the conditional probability P_t and the marginal probability P_b is

$$\begin{aligned} P_b(Y|X=x) &= P_t(Y|X=x), \\ P_b(X) &\neq P_t(X). \end{aligned} \tag{3}$$

Before passing the input vector x to the neuron, it should be noted that x here not only refers to the input of the input layer but also refers to any input layer or hidden layer in the network. X can also be understood as a small batch activation function of a certain layer that needs to be standardized. Translate and scale it first, so that the distribution of x becomes a standard distribution normalized in a fixed interval. For the general change framework, it is as follows:

$$h = f\left(g \cdot \frac{x-\gamma}{\kappa} + \upsilon\right), \tag{4}$$

where γ is the translation parameter, κ is the scaling factor, υ is the retranslation parameter, and g is the rescaling parameter. Through the scaling and translation transformation of γ and κ, we get the following results:

$$\hat{x} = \frac{x-\gamma}{\kappa}. \tag{5}$$

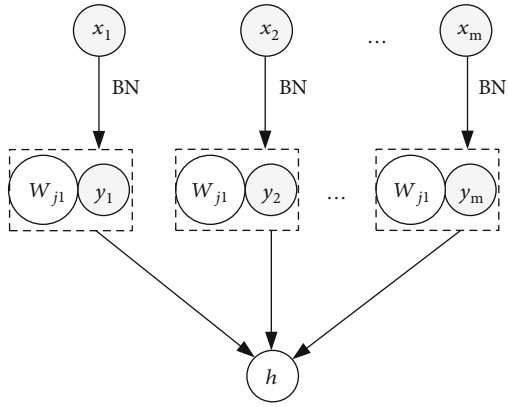


Figure 2: The structure of BN model.

The data obtained satisfy the standard distribution with a mean of 0 and a variance of 1. Substituting formula (5) into formula (4), it is easier to get further transformation:

$$y = g \cdot \hat{x} + \upsilon. \tag{6}$$

Finally, a distribution that satisfies the mean υ and the variance g^2 is obtained.

This normalization of BN is for a single neuron. When training the network, train a small batch of overall data and then calculate the mean and variance of neuron x_i as

$$\begin{aligned} \mu_i &= \frac{1}{m}\sum x_i, \\ \sigma_i &= \sqrt{\frac{1}{m}\sum (x_i - \mu_i)^2 + \varepsilon}, \end{aligned} \tag{7}$$

where m is the size of the small batch and ε is a very small positive value.

3.4. Multiple Feature Extraction. In the neural network model, usually in the first layer of connection, multiple identical convolution kernels are generally used to extract shallow features of the image. There are more commonly used convolution kernels such as 1×1, 3×3, 5×5, 7×7, and 9×9. These convolution kernels are all odd convolution kernels, because this kind of convolution kernel has a common feature. They all have a center point, and even-numbered convolution kernels do not have this property. In shallow feature extraction, usually a larger convolution kernel can learn richer feature information. But the cost is that once the convolution kernel increases, it will adversely affect the training efficiency and speed of the deep network [29].

The proposed algorithm integrates several convolution kernels of different sizes, and its principle is shown in Figure 3. Use multifeature extraction technology to scan the input image block. Since different convolution kernels can extract image features with different information, the one with the largest convolution kernel is used as the moving reference for synchronous movement. In this way, a variety of different feature information will be obtained, and the size of this output is also the same. In the first layer, this feature

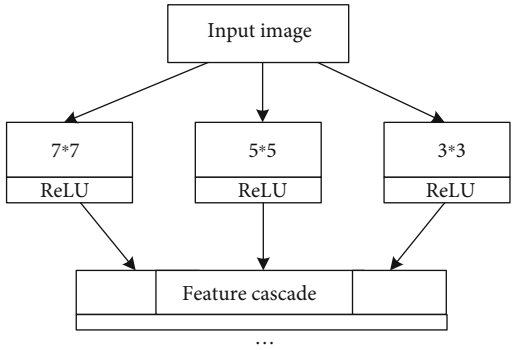


Figure 3: Structure of multifeature extraction model.

information is connected in series. Because more feature information is obtained, this multifeature extraction technology greatly improves the training speed in actual model training and greatly improves the convergence speed of the model.

In the first layer of the network, three convolution kernels of different sizes are used, 3×3, 5×5, and 7×7, and the numbers of these three convolution kernels are 12, 20, and 32, respectively. The number has increased sequentially because the 7×7 convolution kernel can learn relatively richer features. There are a total of 64 convolution kernels of three different sizes. The three convolution kernels of 3×3, 5×5, and 7×7 are used to scan to the right with a sliding step of 7 in a way that the center position is superimposed. According to this rule, three mapping feature maps with the same size should be obtained [30]. The proposed algorithm passes through the first layer and obtains 64 feature maps of the same size. In the first-level feature extraction, the design of such multiple feature extraction can extract richer image features, which greatly improves the convergence speed and training speed of the network model.

3.5. Activation Function. At present, most deep neural networks use a Rectified Linear Unit (ReLU) activation function to accelerate the convergence of the network. The ReLU activation function is not a symmetric function, and the average response of the sequence in ReLU will be less than 0. In addition, even if the input or weight obeys a symmetrical distribution, the response distribution is still asymmetrical. These properties will directly affect the convergence and performance of the network. Therefore, an improved activation function Parametric Rectifier Linear Unit (PReLU) is used. PReLU function is defined as follows:

$$f(y_i) = \begin{cases} y_i, y_1 > 0, \\ k_i y_i, y_i \le 0, \end{cases} \tag{8}$$

where y_i is the input of the nonlinear activation function f on the ith channel and k_i controls the slope of the negative part. When k_i is equal to 0, it becomes the ReLU activation function.

3.6. Joint Loss. The pixel-by-pixel loss is usually used as loss function, and the difference between the pixels between the output image and the input image is calculated as the minimization target to obtain a higher signal-to-noise ratio index. However, the output image using the method of comparing the difference pixel by pixel is prone to problems such as excessive smooth and blurry edges. In response to this problem, there are currently many studies showing that comparing the perceived loss of semantic feature level with the image quality perceived by the human eye shows a good uniformity. More edge texture details in real images can be reconstructed, thereby improving some of the problems of the pixel-by-pixel method. However, simply using this loss may also cause problems such as slight color artifacts due to uneven pixel space coverage. Therefore, by combining the two losses to obtain better results, the joint loss $L_{\&}$ is proposed as

$$L_{\&} = L_{\text{MSE}} + \lambda L_{\text{VGG}}, \tag{9}$$

where L_{MSE} and L_{VGG}, respectively, represent the mean square error (MSE) function of pixel-by-pixel comparison and the perceptual loss of feature comparison.

3.6.1. MSE Loss Function. The pixel-by-pixel loss function uses the traditional MSE method to calculate the MSE of the real target and the predicted target. By comparing each pixel to learn the difference between the two, the optimal solution is obtained. The formula is as follows:

$$L_{\text{MSE}} = \frac{1}{N} \sum_{n=1}^{N} \| F(x) - (x - y) \|_2^2, \tag{10}$$

where N is the total number of training samples, x is the noise input, y is a clean real label, and F is the optimal mapping function obtained after training.

3.6.2. Perceptual Loss Function. The realization of the perceptual loss needs to effectively extract the rich and abstract semantic feature information in the image. A pretrained classification network Visual Geometry Group (VGG) is connected in series as a loss network to extract the required feature map definition L_{VGG}. After the loss network is determined, the loss comparison learning method can input the output $x - F(x)$ and the real noise-free label image y through the initial processing of the front-end multifeature extraction residual network into the VGG network. Extract the feature images of the two from one of the convolutional layers φ and calculate the Euclidean distance between the semantic features of the two according to the formula. The formula is as follows:

$$L_{\text{VGG}} = \frac{1}{N} \sum_{n=1}^{N} \| \varphi_i(x - F(x)) - \varphi_i(y) \|_2^2, \tag{11}$$

where φ is the PReLU function after the ith convolutional layer in the loss network. It is used to extract feature maps, using rich edge texture features and semantic information

for comparison. The application of joint perception loss in the algorithm is shown in Figure 4.

For joint perception loss, first input the noise image to be processed into the built VGG and train the network through MSE loss. Compare the difference between the learning prediction result and the true label map from the pixel level. At this time, the image output by the network has completely removed the noise points, but the edge information is fuzzy [31]. So, the fuzzy denoising result and label are passed through the pretrained VGG again. The feature maps of the two are obtained from the activation function after the specific convolutional layer for comparison. Minimize the perceptual loss as the training goal for network training, so that the output image contains more edge information features. It is possible to restore the originally blurred area during reconstruction, obtain clearer and sharper image edges, and obtain clearer image denoising results.

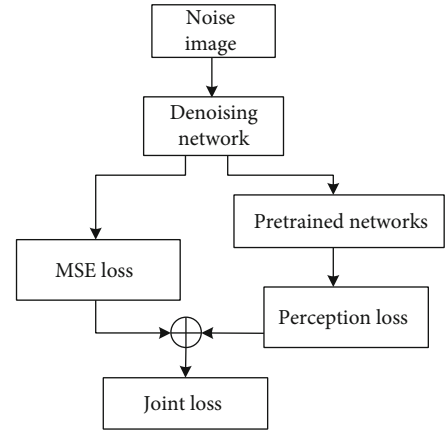


Figure 4: Flow chart of joint loss.

4. Network Parameter Setting

In the process of neural network training, it is necessary to learn a set of optimal parameters to minimize the result of loss function, so a suitable optimization algorithm needs to be added. So far, the commonly used neural network optimization algorithm is the gradient descent algorithm, which is used to find the minimum parameter of the loss function. In the training process, the error is gradually reduced and the local minimum of the function is found. Differentiate the loss function to get the gradient of a given point [32]. The positive and negative values of the gradient indicate an increase or decrease in the value of the loss function. Select the direction that reduces the cost function value, that is, the negative gradient direction, multiply the updated amount of the parameter calculated by the learning rate, and update the parameter.

The single-step weights and biases are updated as follows:

$$\omega_k \to \omega_k = \omega_k - \partial \frac{\delta C}{\delta \omega_k}, \qquad b_l \to b_l = b_l - \partial \frac{\delta C}{\delta b_l}. \tag{12}$$

The main problem of the gradient descent algorithm is that if the location of the initial point is unreasonably selected, the network is easy to fall into a local optimum, and it is difficult to find the global optimum. In addition, if the step size of the single step down is too small, the calculation amount will be too large, and if the number of iterations is too large, the step size may be too large to skip the optimal solution. And the gradient descent algorithm is not fast enough when the amount of data is large. Therefore, the stochastic gradient descent (SGD) algorithm is used to calculate the gradient using a single sample to speed up the calculation. However, the SGD method uses individuals to represent the overall change trend and cannot ensure that each iteration tends to the global optimal solution, and it cannot guarantee that each iteration will reduce the loss function result. Therefore, an adaptive moment estimation (Adam) optimization algorithm is proposed to replace the SGD algorithm [33]. The SGD algorithm maintains a single learning rate, updates all weights, and keeps the learning rate during the training process unchanged. And Adam iteratively updates the neural network weights by calculating the first-order moment estimation and the second-order moment estimation of the gradient. The adaptive learning rate is calculated for each parameter to solve the problem of high-intensity noise or sparse gradient [34]. The basic steps of the Adam optimization algorithm are as follows:

Step 1. Suppose $f(\theta)$ is the noise objective function, which is a random scalar function with differentiable parameters θ.

Step 2. Update the exponential moving average s_t and squared gradient v_t of the gradient. The moving average is estimated using the first-order moment and the second-order original moment of the gradient, and the step size ζ is selected reasonably.

Step 3. Initialize the deviation correction term to obtain the gradient of the random objective function f. Then, use the exponential moving average of v_t and the decay rate τ_2 to estimate the second-order original moment. That is, eliminate v_t:

$$v_t = (1 - \tau_2) \sum_{i=1}^{t} \tau_2^{t-i} \cdot q_i^2, \tag{13}$$

where τ_2 is the exponential decay rate estimated by the second moment, t is the time step, and $q_1, \cdots, q_T$ is the gradient on the time step sequence. In order to understand how the expected value penalty $E[v_t]$ of the exponential moving average at time step t is related to the true second moment, the deviation between these two quantities is corrected, as follows:

$$E[v_t] = E\left[(1-\tau_2)\sum_{i=1}^{t}\tau_2^{t-i}\cdot q_i^2\right] = E\left[q_t^2\right]\cdot(1-\tau_2)\sum_{i=1}^{t}\tau_2^{t-i} + \varepsilon = E\left[q_t^2\right]\cdot\left(1-\tau_2^t\right) + \varepsilon. \tag{14}$$

If the second moment $E[q_t^2]$ is static, then $\varepsilon = 0$.

In addition, for the parameter settings during training, the input of the network is to randomly cut out 45×45 image blocks from the training set images. The first layer of the network consists of convolution kernels of multiple sizes. Among them, there are 12 convolution kernels of 3×3, 20 convolution kernels of 5×5, 32 convolution kernels of 7×7 , and a total of 64 channels. For the convolutional layer in the following residual module, 64 convolution kernels of 3×3 are uniformly used. The last reconstruction layer uses c convolution kernels of 3×3 ($c = 1$ for grayscale images, $c = 3$ for color images). The algorithm to optimize the regression target uses the Adam method, and the momentum parameter is 0.9. When the training function is joint loss, the training batch size is smaller than 18, and the initial learning rate is 10-4. The learning rate is halved after every 2.0×105 iterations.

In the batch normalization used, a sufficient comparison experiment was done on the settings of the minibatch parameters. In the comparative experiment, the minibatch size is set to 32, 64, and 128, respectively, and the model convergence speed and the final denoising effect achieved under these parameters are compared. It is found from the experimental results that the overall effect of image denoising is the best when the minibatch value is 64. Therefore, in the training process of the denoising model, the value of batch-normalized minibatch is set to 64.

5. Experiment Scheme and Result Discussion

In order to evaluate the proposed algorithm, we must first have an evaluation standard and do different experiments on different test datasets, fully contrast with other excellent image denoising algorithms, and finally draw a conclusion.

5.1. Experimental Dataset. Since the use of the deep residual network algorithm requires a large amount of training image data, the experimental data selected by the proposed algorithm uses the internationally published glioblastoma multiforme (GBM) multimodal MR image dataset TCGA. Among them, the foreign population GBM experimental test library (TCGA-GBM) and the Chinese population GBM experimental test library (CH-GBM) are established. Some image examples of the dataset are shown in Figure 5.

In the experiment, 227 images were randomly selected from the TCGA-GBM dataset, of which 200 were used as training images and the remaining 27 were used as test images. The image size is set to 180×215 and 64×77. 115 images were randomly selected from the CH-GBM dataset, 100 were used as training images, and the remaining 15 were used as test images. The image size is set to 180×215 and 64×77. The image block size is set to 32×32, and the image block step size is 10. A total of about 86,000 training image blocks can be obtained. Then, the tested images are evaluated using the Peak Signal Noise Ratio (PSNR) and Structural Similarity Index (SSIM).

In the proposed algorithm, the size of the training set is artificially increased by methods such as image translation and flipping. Because in the training process, when the amount of data is small, it will cause the model to overfit, so that the training error is small and the test error is large. Therefore, the occurrence of overfitting can be effectively suppressed by adding a regular term after the cost function.

5.2. Evaluation Standard of Denoising Effect. The proposed algorithm mainly uses two objective evaluation indexes: PSNR and SSIM. Given a reference image, it is represented by f, and the test image is represented by d. The size of these two images is $(M \times N)$; then, the PSNR between the images is defined as

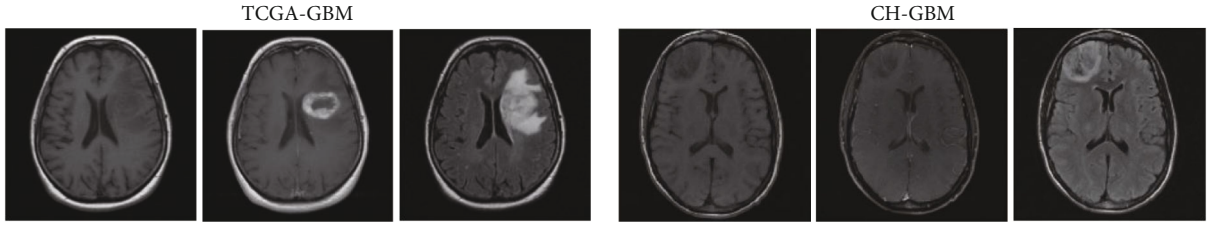
$$\begin{aligned} \mathrm{PSNR}(f,d) &= 10\log_{10}\frac{255^2}{\mathrm{MSE}(f,d)}, \\ \mathrm{MSE}(f,d) &= \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left(f_{ij}-d_{ij}\right)^2. \end{aligned} \tag{15}$$

When MSE tends to zero, the value of PSNR tends to infinity. This means that the higher the PSNR value, the better the image quality. The smaller the PSNR value, the greater the difference between the two images.

SSIM is a quality evaluation model that takes into account the brightness distortion, contrast distortion, and related loss of the image. It is defined as

$$\begin{aligned} \mathrm{SSIM}(f,d,l) &= l(f,d)c(f,d)s(f,d), \\ &\begin{cases} l(f,d) = \dfrac{2\mu_f\mu_d + C_1}{\mu_f^2 + \mu_d^2 + C_1}, \\ c(f,d) = \dfrac{2\sigma_f\sigma_d + C_2}{\sigma_f^2 + \sigma_d^2 + C_2}, \\ s(f,d) = \dfrac{\sigma_{fd} + C_3}{\sigma_f\sigma_d + C_3}, \end{cases} \end{aligned} \tag{16}$$

where $l(f,d)$ is the brightness comparison function, used to calculate the similarity of the average brightness μ_f and μ_d of the two images. When $\mu_f = \mu_d$, $l(f,d)$ takes the maximum value. $c(f,d)$ is the contrast comparison function. Measure the similarity of the contrast of two images. Contrast is measured by standard deviation σ_f and σ_d. It is only when σ_f equals σ_d that $c(f,d)$ has a maximum value of 1. $s(f,d)$ is a structure comparison function, used to represent the correlation between f and d pixels of two images. σ_{fd} is the covariance between f and d, and the value range of SSIM is [0,1]. A value of 0 means that there is no correlation between the two images, and a value of 1 means that f is equal to d. Constants C_1, C_2, and C_3 are used to avoid the phenomenon of a zero denominator.



(a) TCGA-GBM (b) CH-GBM

Figure 5: Partial image examples of the dataset.

5.3. Analysis of Network Iteration Algorithm. For the proposed network, the training iterations are all 50. When image denoising, the learning rate that is too small will lead to slower convergence speed, and the long-term slowness in feature learning will cause weak noise in the parameter update process, which will affect the quality of the denoised image. However, when the learning rate is large, the network system will be unstable. Considering comprehensively, the learning rate of the proposed network model is 0.01. The MatConvNet toolkit is used to train the model in the network. Since the training time of each model is different, the proposed training model runs for two days on average when different noises are added.

In order to effectively obtain the spatial information of the image, a 15-layer network is set up, including the convolutional layer, activation function, pooling layer, and BN layer. Under the condition that the basic parameters of the experiment remain unchanged, the influence of the two optimization algorithms of Adam and SGD on the denoising results is studied. When the image is added with 7% noise, the average PSNR value of the test image after denoising using the model optimized by Adam and SGD algorithm is shown in Figure 6.

It can be seen from Figure 6 that the proposed network model using the Adam algorithm optimized model to remove the noise in the medical MR image is better than the SGD algorithm optimized model. The average PSNR value has increased by about 1 dB, and the denoising result is relatively stable. When the number of iterations is 20, the average PSNR value tends to be stable. Therefore, in the proposed network model, the optimization algorithm uses Adam.

5.4. Qualitative Comparison of Denoising Effects. In order to compare the effect of medical MR image size on the denoising effect, the same TCGA-GBM dataset images with sizes of 190×215 and 62×77 were selected in the experiment. Train the denoising algorithm by manually adding 7% of noisy images. Finally, four different MR images are selected as test images to obtain the final denoising image. Figure 7 shows the average PSNR curve after denoising the two different sizes of test images.

It can be seen from Figure 7 that image training models of different pixels have a relatively large impact on the denoising effect. When training the model with images with larger pixels, the PSNR of the denoised image is significantly higher than that of images with smaller pixels. When the pixels are small, the average PSNR of the training model after denoising the image is relatively low, the whole process is unstable, and the convergence speed is slow. When the pixels are small, this is equivalent to reducing the amount of training image data. This leads to overfitting of the model, which makes the training error smaller and the test error larger. Therefore, when training the model, choosing a larger amount of data will help improve denoising performance of the algorithm. Through analyzing the influence of the above-mentioned selected network parameters on the image denoising performance, in the verification experiment of the proposed algorithm, both the training image and the test image are selected in the size of 190×215, and the optimization algorithm is Adam.

In order to further visually demonstrate the denoising effect of the proposed algorithm on the TCGA-GBM and CH-GBM datasets, compare it with literature [14, 20, 26]. The result is shown in Figure 8.

It can be seen from Figure 8 that, compared with the other three denoising algorithms, the image after denoising by the proposed algorithm is more thorough and the edge preservation effect is better. In contrast, the image denoised by the algorithm in literature [14, 20] obviously has some residual noise, and the noise removal is not very thorough. In the literature [14], the image edge blur after denoising using the improved MMF algorithm is larger. In particular, the blurred residual noise can be clearly seen around the edge of the brain image. And for the smoother part of the original image, the image becomes unsmooth after denoising, and fuzzy noise blocks of varying degrees appear. However, in the literature [20], the adaptive FIR algorithm is used to denoise the denoised image twice, and the phenomenon of filtering will appear. But for the Rician noise in medical MR images, the denoising image produces blurring, which makes the image noise removal incomplete. The main manifestation is that there are noise points around the edges of the image, but it is improved compared to the improved MMF algorithm. In Reference [26], combining the neural network and wavelet method, the image after denoising is significantly improved compared with the previous two methods. However, this method smoothes the details of the image, especially the contour area of the brain image, and the phenomenon of excessive filtering occurs, thereby losing useful information in the details. The proposed algorithm has a relatively good effect on removing noise in medical MR images, avoids the problem of incomplete denoising of the first two denoising methods, and can also reduce oversmoothing. Compared with the three comparison algorithms, it has a better overall effect on noise removal.

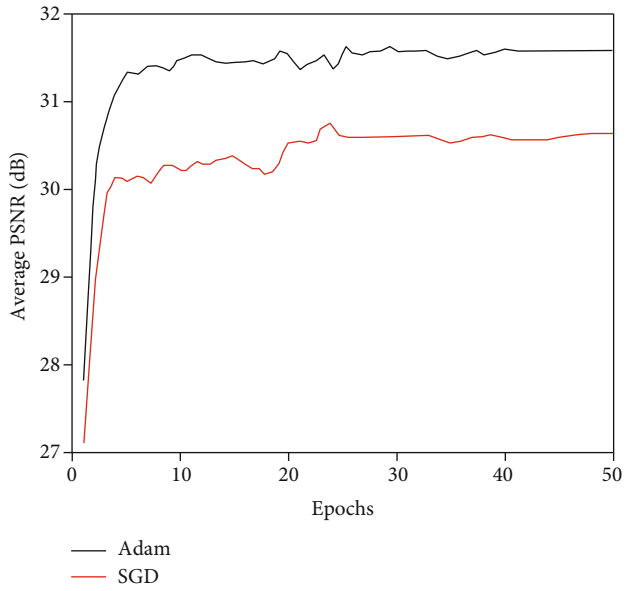


Figure 6: Denoising results using different optimization algorithms.

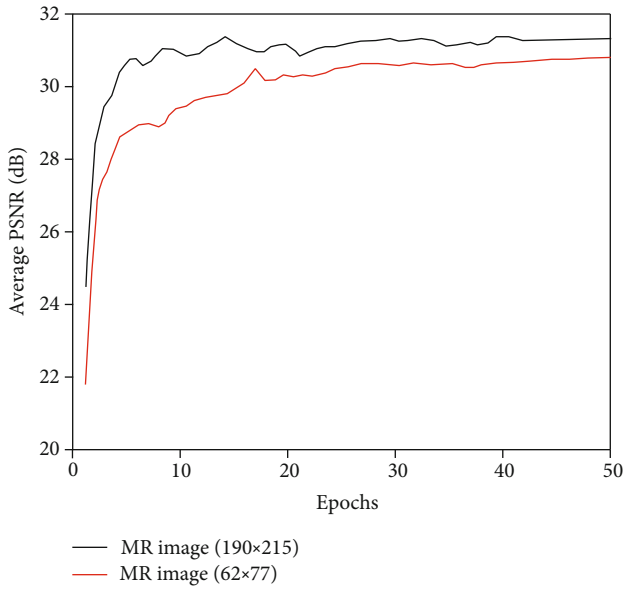


Figure 7: Denoising results of different size images.

5.5. Quantitative Comparison of Denoising Effects. Aiming at the problem of denoising in medical MR images, the denoising effects of different denoising algorithms are clearly analyzed. The proposed algorithm and the algorithm in literature [14, 20, 26] are quantitatively evaluated in two indicators: average PSNR and SSIM. Among them, the average SSIM of the four denoising algorithms after denoising the medical MR image with different noise intensities is shown in Table 1.

It can be seen from Table 1 that the SSIM value of the proposed algorithm is significantly higher than that of the algorithm in [14, 20], and the denoising effect is consistent with the subjective observation results. Comparing the proposed algorithm with the algorithm in [26], both have improved. Especially when the noise intensity is relatively weak, the SSIM value of the proposed algorithm has been greatly improved, showing a very obvious advantage. Therefore, the proportion of useful information in the denoised image of the proposed algorithm is closer to that of the original noise-free image. When 1% noise is added, SSIM can reach 0.9941 after denoising. Compared with other algorithms, the values are improved, and more satisfactory results are obtained, with less information loss.

On the TCGA-GBM and CH-GBM datasets, for different noise levels, the proposed algorithm is quantitatively compared with the MR image denoising effect of the algorithms in [14, 20, 26]. The average PSNR and SSIM curves obtained are shown in Figures 9 and 10.

It can be seen from Figures 9(a) and 10(a) that under different noise intensities, the average PSNR of the image after denoising using the proposed algorithm is significantly higher than that of the image processed by other denoising algorithms. And when denoising weakly noisy images, the average PSNR of the proposed algorithm and other contrasting algorithms are relatively small. In the case of strong noise denoising, the average PSNR is quite different, indicating that the proposed algorithm has better effect on removing strong noise in MR images. The denoised image has a very high similarity to the original noiseless image.

At the same time, it can be seen from Figures 9(b) and 10(b) that the four denoising algorithms have little difference in the average SSIM of the image after denoising when the noise intensity is weak. When the noise intensity is high, the average SSIM of the proposed algorithm after denoising shows obvious advantages. In terms of structural similarity, it is closer to the original image.

In summary, the proposed algorithm is better than the other three denoising algorithms in denoising simulated medical MR images. Especially under strong noise, it can effectively remove the noise in the simulated medical MR image.

6. Conclusions

Currently, medical image information can be obtained through a variety of technical means, among which MRI is a relatively common medical image acquisition technology. However, MR images will be interfered by random noise during the acquisition process, which reduces the useful information in images. Decreasing the accuracy and effectiveness of imaging will directly affect the correct diagnosis and treatment of clinicians. Therefore, a denoising algorithm for medical MR images based on multifeature extraction based on a deep residual network is proposed. The feature extraction layer is constructed by combining three convolution kernels of different sizes to obtain multiple shallow features for fusion and combined with batch normalization and residual learning technology to accelerate and optimize the deep network. In addition, a joint loss function is defined by combining the perceptual loss and the traditional mean square error loss to generate a clearer target image. Based on the MATLAB simulation platform, the TCGA-GBM and CH-

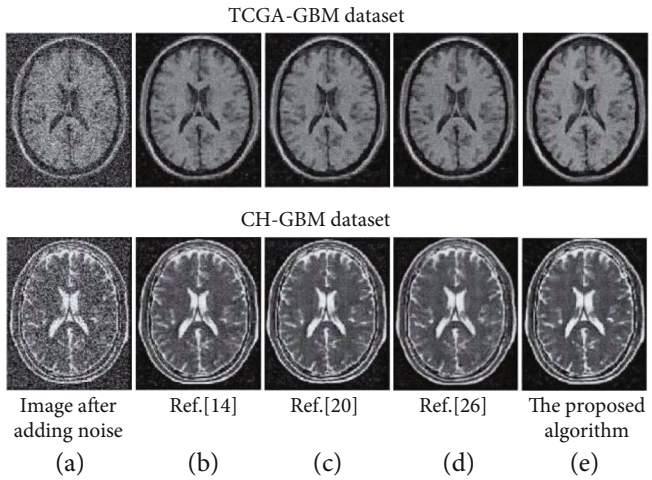


Figure 8: Comparison of denoising effects of different algorithms.

Table 1: SSIM comparison of different denoising algorithms.

Dataset	Algorithm	1%	4%	7%	10%	13%
TCGA-GBM	Ref. [14]	0.9903	0.9584	0.9153	0.8612	0.8109
	Ref. [20]	0.9901	0.9592	0.9276	0.8928	0.8529
	Ref. [26]	0.9918	0.9643	0.9374	0.9022	0.8657
	The proposed algorithm	0.9941	0.9668	0.9417	0.9152	0.8893
CH-GBM	Ref. [14]	0.9829	0.9622	0.9238	0.8874	0.8378
	Ref. [20]	0.9923	0.9637	0.9246	0.8984	0.8557
	Ref. [26]	0.9935	0.9626	0.9415	0.9043	0.8736
	The proposed algorithm	0.9978	0.9689	0.9426	0.9037	0.8809

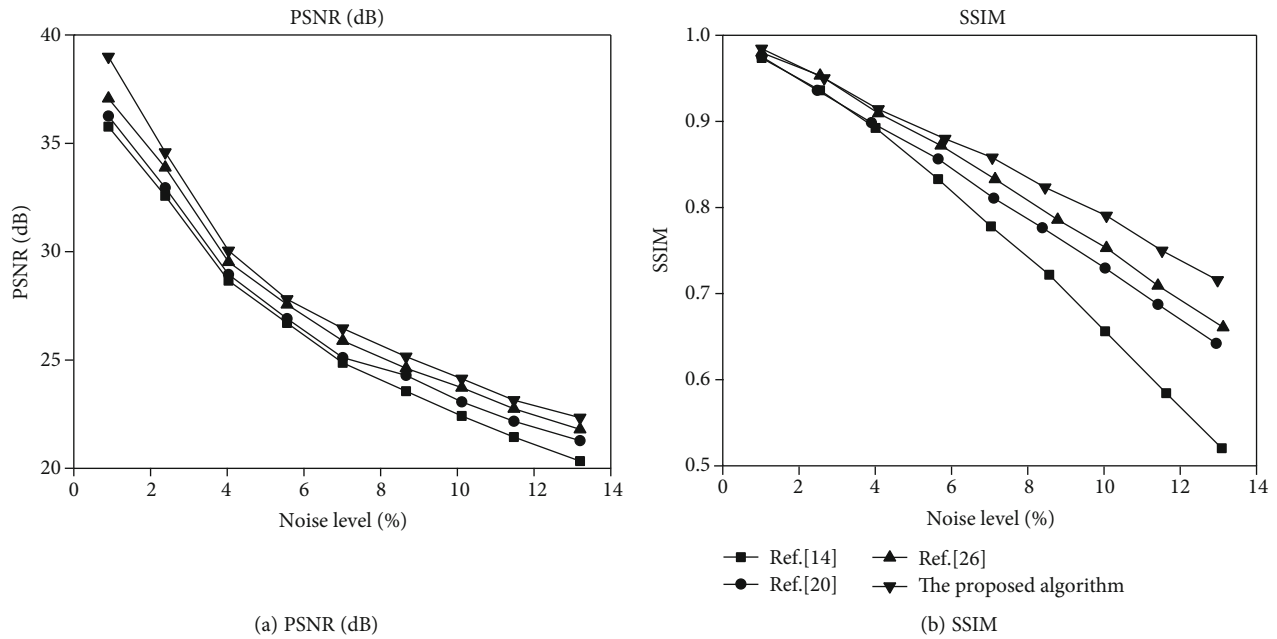
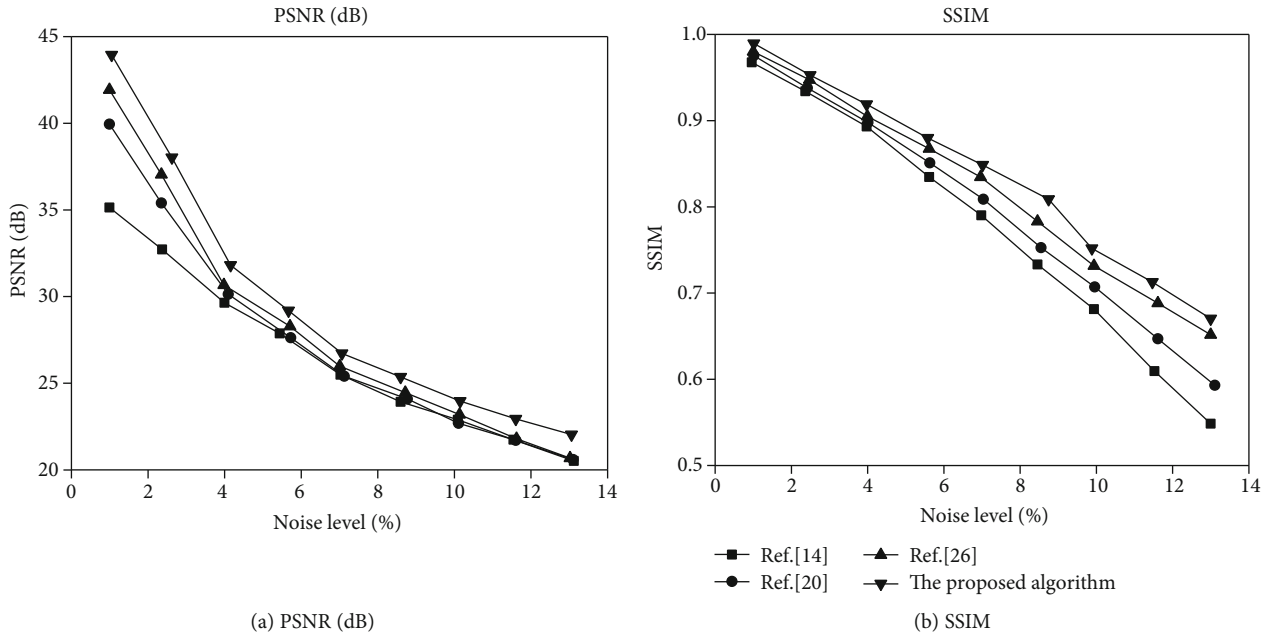


Figure 9: Comparison of denoising effects of different algorithms in the TCGA-GBM dataset.



(a) PSNR (dB) (b) SSIM

Figure 10: Comparison of denoising effects of different algorithms in the CH-GBM dataset.

GBM datasets are used to experimentally demonstrate the proposed algorithm. The results show that the performance of the proposed algorithm is the best when the image size is set to 190×215 and Adam is selected as the optimization algorithm. And the two indexes of PSNR and SSIM of the proposed algorithm are significantly higher than other comparison algorithms. As the noise level increases, the difference between the values becomes larger, which shows that it is suitable for processing high-intensity noise MR images.

Regardless of whether it is in the classification or denoising of deep residual learning, the parameter adjustment of the deep network has always been a key step, and it is more difficult. Therefore, further research is needed. In addition, the deep residual network can achieve better results when the amount of training data is large. But when the amount of data is small, the denoising effect needs to be improved. At the same time, the amount of data is large, which is slower than traditional algorithms. Therefore, the training speed of the network needs to be further improved while ensuring the denoising effect.

Data Availability

The data included in this paper are available without any restrictions.

Conflicts of Interest

The authors declare that they have no conflicts of interest to report regarding the present study.

Acknowledgments

This work was supported by the Directing project of the Hubei Provincial Department of Education (No.B2018251) and Hubei Polytechnic University teaching reform research project (No.2018C19).

References

[1] I. Sack and T. Schaeffter, "Quantification of Biophysical Parameters in Medical Imaging Extracellular matrix-specific molecular MR imaging probes for the assessment of aortic aneurysms," vol. 10, no. 16, pp. 373–381, 2018.

[2] J. Wang, X. Tao, Y. Yuan, and Y. Wu, "Application of MR fingerprinting technology in medical imaging," *Chinese Journal of Interventional Imaging and Therapy*, vol. 15, no. 1, pp. 59–62, 2018.

[3] J. Sultana, M. S. Islam, M. R. Islam, and D. Abbott, "High numerical aperture, highly birefringent novel photonic crystal fibre for medical imaging applications," *Electronics Letters*, vol. 54, no. 2, pp. 61-62, 2018.

[4] M. R. Santos, A. Silva, and N. P. Rocha, "DICOM metadata quality auditing for medical imaging stakeholders characterisation: a pilot study," *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, vol. 7, no. 1-2, pp. 188–196, 2019.

[5] M. R. Khosravi, "Improving the scientific influence of international journals: a guideline for guest editors of current medical imaging reviews," *Current Medical Imaging Reviews*, vol. 14, no. 4, pp. 487-488, 2018.

[6] A. V. Dalca, K. L. Bouman, W. T. Freeman, N. S. Rost, M. R. Sabuncu, and P. Golland, "Medical image imputation from image collections," *IEEE Transactions on Medical Imaging*, vol. 38, no. 2, pp. 504–514, 2019.

[7] L. Buckley, G. Heddon, I. Byrne, and C. Angers, "Improved X-ray safety, quality control, and resource management in medical imaging using QATrack+," *Journal of Medical Imaging and Radiation Sciences*, vol. 51, no. 1, pp. 22–28, 2020.

[8] O. Oktay, W. Bai, R. Guerrero et al., "Stratified decision forests for accurate anatomical landmark localization in cardiac

images," *IEEE Transactions on Medical Imaging*, vol. 36, no. 1, pp. 332–342, 2017.

[9] B. K. Choi, N. Madusanka, H. K. Choi et al., "Convolutional neural network-based MR image analysis for Alzheimer's disease classification," *Current Medical Imaging Reviews*, vol. 16, no. 1, pp. 27–35, 2020.

[10] Y. Huo, Z. Xu, H. Moon et al., "SynSeg-Net: synthetic segmentation without target modality ground truth," *IEEE Transactions on Medical Imaging*, vol. 38, no. 4, pp. 1016–1025, 2019.

[11] W. Jifara, F. Jiang, S. Rho, M. Cheng, and S. Liu, "Medical image denoising using convolutional neural network: a residual learning approach," *Journal of Supercomputing*, vol. 75, no. 2, pp. 704–718, 2019.

[12] C. Kumar and R. Prakash, "Ultrasound medical image denoising using threshold based wavelet transformation method," *Journal of Medical Imaging and Health Informatics*, vol. 10, no. 8, pp. 1825–1830, 2020.

[13] Y. Zheng, M. Li, J. Zhang, and J. Wang, "Selection of regularization parameter in GMM based image denoising method," *Multimedia Tools and Applications*, vol. 77, no. 22, pp. 30121–30134, 2018.

[14] M. Aghajarian, J. E. Mcinroy, and C. H. G. Wright, "Salt-and-pepper noise removal using modified mean filter and total variation minimization," *Journal of Electronic Imaging*, vol. 27, no. 1, pp. 168–175, 2018.

[15] T. Qiu, C. Wen, K. Xie, F. Q. Wen, G. Q. Sheng, and X. G. Tang, "Efficient medical image enhancement based on CNN-FBB model," *IET Image Processing*, vol. 13, no. 10, pp. 1736–1744, 2019.

[16] K. Kalaivani and V. P. Y. Asnath, "Modified wiener filter for restoring Landsat images in remote sensing applications," *Pertanika Journal of Science & Technology*, vol. 26, no. 3, pp. 1005–1018, 2018.

[17] Y. Chang, L. Yan, M. Chen, H. Fang, and S. Zhong, "Two-stage convolutional neural network for medical noise removal via image decomposition," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 6, pp. 2707–2721, 2020.

[18] Z. HosseinKhani, M. Hajabdollahi, N. Karimi et al., "Real-time removal of impulse noise from MR images for radiosurgery applications," *International Journal of Circuit Theory and Applications*, vol. 47, no. 3, pp. 406–426, 2019.

[19] H. S. M. Muslim, S. A. Khan, S. Hussain, A. Jamal, and H. S. A. Qasim, "A knowledge-based image enhancement and denoising approach," *Computational & Mathematical Organization Theory*, vol. 25, no. 2, pp. 108–121, 2019.

[20] W. Fan, W. S. Xiao, and W. S. Xiao, "Image denoising based on wavelet thresholding and Wiener filtering in the wavelet domain," *The Journal of Engineering*, vol. 2019, no. 19, pp. 6012–6015, 2019.

[21] S. Kollem, K. R. L. Reddy, and D. S. Rao, "Denoising and segmentation of MR images using fourth order non-linear adaptive PDE and new convergent clustering," *International Journal of Imaging Systems & Technology*, vol. 29, no. 3, pp. 195–209, 2019.

[22] Z. Huang, Y. Zhang, Q. Li, T. Zhang, N. Sang, and H. Hong, "Progressive dual-domain filter for enhancing and denoising optical remote-sensing images," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 5, pp. 759–763, 2018.

[23] R. Abazari and M. Lakestani, "Non-subsampled shearlet transform and log-transform methods for despeckling of medical ultrasound images," *Informatica*, vol. 30, no. 1, pp. 1–19, 2019.

[24] M. Azhar, H. Dawood, H. Dawood, G. I. Choudhary, A. K. Bashir, and S. H. Chauhdary, "Detail-preserving switching algorithm for the removal of random-valued impulse noise," *Journal of Ambient Intelligence and Humanized Computing*, vol. 10, no. 10, pp. 3925–3945, 2019.

[25] R. S. Thakur, R. N. Yadav, and L. Gupta, "State-of-art analysis of image denoising methods using convolutional neural networks," *IET Image Processing*, vol. 13, no. 13, pp. 2367–2380, 2019.

[26] V. Nasir, J. Cool, and F. Sassani, "Intelligent machining monitoring using sound signal processed with the wavelet method and a self-organizing neural network," *IEEE Robotics and Automation Letters*, vol. 4, no. 4, pp. 3449–3456, 2019.

[27] B. N. Anoop, J. Joseph, J. Williams, J. S. Jayaraman, A. M. Sebastian, and P. Sihota, "A prospective case study of high boost, high frequency emphasis and two-way diffusion filters on MR images of glioblastoma multiforme," *Australasian Physical & Engineering Sciences in Medicine*, vol. 41, no. 2, pp. 415–427, 2018.

[28] S. G. Javed, A. Majid, and Y. S. Lee, "Developing a bio-inspired multi-gene genetic programming based intelligent estimator to reduce speckle noise from ultrasound images," *Multimedia Tools and Applications*, vol. 77, no. 12, pp. 15657–15675, 2018.

[29] R. Kashyap and V. Tiwari, "Active contours using global models for medical image segmentation," *International Journal of Computational Systems Engineering*, vol. 4, no. 2-3, pp. 195–201, 2018.

[30] F. Hashimoto, H. Ohba, K. Ote, A. Teramoto, and H. Tsukada, "Dynamic PET image denoising using deep convolutional neural networks without prior training datasets," *IEEE Access*, vol. 7, no. 99, pp. 96594–96603, 2019.

[31] I. S. Klyuzhin, J. C. Cheng, C. Bevington, and V. Sossi, "Use of a tracer-specific deep artificial neural net to denoise dynamic PET images," *IEEE Transactions on Medical Imaging*, vol. 39, no. 2, pp. 366–376, 2020.

[32] K. Gong, J. Guan, K. Kim et al., "Iterative PET image reconstruction using convolutional neural network representation," *IEEE Transactions on Medical Imaging*, vol. 38, no. 3, pp. 675–685, 2019.

[33] K. Kyungsang, W. Dufan, G. Kuang et al., "Penalized PET reconstruction using deep learning prior and local linear fitting," *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, pp. 1478–1487, 2018.

[34] F. Hashimoto, H. Ohba, K. Ote, and H. Tsukada, "Denoising of dynamic sinogram by image guided filtering for positron emission tomography," *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 2, no. 6, pp. 541–548, 2018.