Research on Diagnostic Imaging Method of COVID-19 Based on 3D U-Net

Xin Yuan

School of AI and Advanced Computing, Xi'an Jiaotong-liverpool University, Suzhou, Jiangsu, 215000, China

xin.yuan2002@student.xjtlu.edu.cn

Keywords: COVID-19; CT images; 3D U-Net; Diagnostic imaging

Abstract: With the global spread of the COVID-19 epidemic, the diagnosis method of COVID-19 based on CT images has received extensive attention. This paper studies the diagnostic imaging method for COVID-19 based on 3D U-Net. The main contents include the introduction of the basic theory and structure of 3D U-Net, the analysis of the characteristics and modeling methods of COVID-19 infections, the optimization strategy of the convolutional neural network, the comparison of different loss functions and principal component analysis on the segmentation results, and the evaluation of the influence of noise error on the segmentation performance. This paper concludes that the COVID-19 imaging diagnosis method based on 3D U-Net can effectively identify and segment the focus of COVID-19, improving the accuracy and efficiency of diagnosis. It is of great importance for preventing and controlling epidemics and public health protection.

1. Introduction

Imaging diagnosis is one of the main contents of medical imaging and is also a general term for medical imaging technology. It can be divided into basic and non-basic imaging diagnoses, composed of conventional and unique imaging technology, respectively. To improve the accuracy and efficiency of diagnosis, medical imaging entrusts computer science to assist in imaging diagnosis. Since the COVID-19 outbreak, the diagnostic method of COVID-19 based on CT images has become the key to preventing and controlling the epidemic, and the diagnostic imaging method of COVID-19 based on 3D U-Net has become an evaluation index. Compared with the traditional two-dimensional image diagnosis, the COVID-19 image diagnosis method based on 3D U-Net emphasizes three-dimensional information, deep learning, and automatic segmentation. Therefore, we propose the issue of the COVID-19 imaging diagnosis method based on 3D U-Net, and the convolutional neural network provides a new optimization strategy for it.

Convolutional neural network originates from deep learning with artificial neural networks as the core, and its optimization strategy contains mathematical principles. From the perspective of convolutional neural network structure, convolutional neural network pursues high efficiency and realizes image feature extraction and classification through the combination of convolutional layer, pooling layer, and fully connected layer. However, it only happens ideally. Today, convolutional neural networks have implemented a unique path that adapts to different tasks and datasets. The advancement of convolutional neural networks has rewritten image processing, reflected the development level of artificial intelligence, and will change the field of medical imaging, posing a challenge to public health. Therefore, researchers must have an interdisciplinary vision and pattern to discuss convolutional neural networks. We propose a convolutional neural network optimization strategy in the COVID-19 image diagnosis method based on 3D U-Net.

In conclusion, the optimization strategy of the convolutional neural network is the critical condition and guarantee to realize the COVID-19 image diagnosis method based on 3D U-Net. From the existing research, the optimization strategy of convolutional neural networks has made progress, but there are also areas for improvement. Scientists have yet to find a practical path fully applicable to COVID-19 lesion segmentation and are still trying to move forward. Therefore, the optimization strategy of the convolutional neural network needs to be innovated and improved, which is conducive

  DOI: 10.25236/icceme.2023.014

to improving the diagnostic level of COVID-19 and meeting the needs of developing medical imaging.

Based on the above background, this paper proposes a convolutional neural network optimization strategy to optimize the COVID-19 image diagnosis method based on 3D U-Net. The loss function principal component analysis theory and noise error evaluation method are used to solve the problem of inconsistent and inaccurate segmentation results. The main contents include comparing the influence of different loss functions, principal component analysis on segmentation results, and evaluating the influence of noise error on segmentation. In addition, it effectively reduces the risk brought by CT image value range and data set quality, which has theoretical and practical significance.

2. The Basic Theory of 3D U-Net

3D U-Net is a convolutional neural network for 3D volume segmentation based on the original U-Net architecture, replacing all 2D operations with 3D operations [1]. The basic theory is as follows:

(1) The network structure of 3D U-Net consists of an encoding path and a decoding path, each with four resolution levels.

(2) The encoding path consists of two 3×3×3 convolutional layers per layer, each followed by a ReLU layer, followed by a 2×2×2 max pooling layer with a stride of 2 in each direction.

(3) Each layer of the decoding path consists of a 2×2×2 deconvolution layer with a stride of 2, followed by two 3×3×3 convolutional layers, each followed by a ReLU layer.

(4) Through shortcuts, layers of the same resolution in the encoding path are passed to the decoding path, giving them the original high-resolution features.

(5) The last layer is a 1×1×1 convolutional layer, which can reduce the number of output channels, and the final number of output channels is the number of categories of labels.

(6) Batch normalization was used before ReLU.

(7) Using the weighted softmax loss function, the network can be trained using sparsely annotated data. By setting the weight of unlabeled pixels to zero, the network can learn from labeled pixels and extend to the entire stereo data.

Here is a formula to represent the loss function of 3D U-Net:

$$L = -\sum_{i=1}^{N} w_i \log(p_i) \tag{1}$$

In the formula, N is the number of pixels, w_i is the weight of the i-th pixel, and p_i is the predicted probability of the corresponding category of the i-th pixel.

3. COVID-19 Lesion Modeling

COVID-19 lesion modeling is an effective method for diagnosing and treating COVID-19, focusing on the impact of the COVID-19 virus on the lung organs. Three-dimensional visualization and quantitative analysis directly reflect the extent of infection and patient rehabilitation. Several components of COVID-19 lesion modeling are being gradually developed, including various scoring systems, such as the 2.5-dimensional multi-view model and the C-SEIR infection dynamics model, which are gradually gaining attention [2]. Among them, the multi-view 2.5D model uses an efficient segmentation algorithm that can quickly complete accurate segmentation of inflammatory regions. The formula is expressed as (2):

$$S = \arg\min_S \sum_{i=1}^{N} w_i \left(\frac{1}{2} \left(1 - \frac{\langle I_i, S_i \rangle}{\| I_i \| \ \| S_i \|} \right) + \lambda \left(1 - S_i \right) \right) \tag{2}$$

The C-SEIR epidemic dynamic model considers the impact of regional mobility, isolation and other prevention and control measures on the transmission process. The formula is as follows:

$$
\begin{aligned}
&\frac{dS}{dt} = -\beta SI - \mu S + \sigma \\
&R\frac{dE}{dt} = \beta SI - (\alpha + \mu) \\
&E\frac{dI}{dt} = \alpha E - (\gamma + \mu) \\
&I\frac{dR}{dt} = \gamma I - (\sigma + \mu)R
\end{aligned}
\tag{3}
$$

S represents the susceptible population, E represents the incubation period population, I represents the infection period population, R represents the recovered population, β represents the infection rate, α represents the conversion rate during the incubation period, γ represents the recovery rate, μ represents the mortality rate, and σ represents the reinfection rate.

The modeling of COVID-19 lesions is shown in Figure 1.

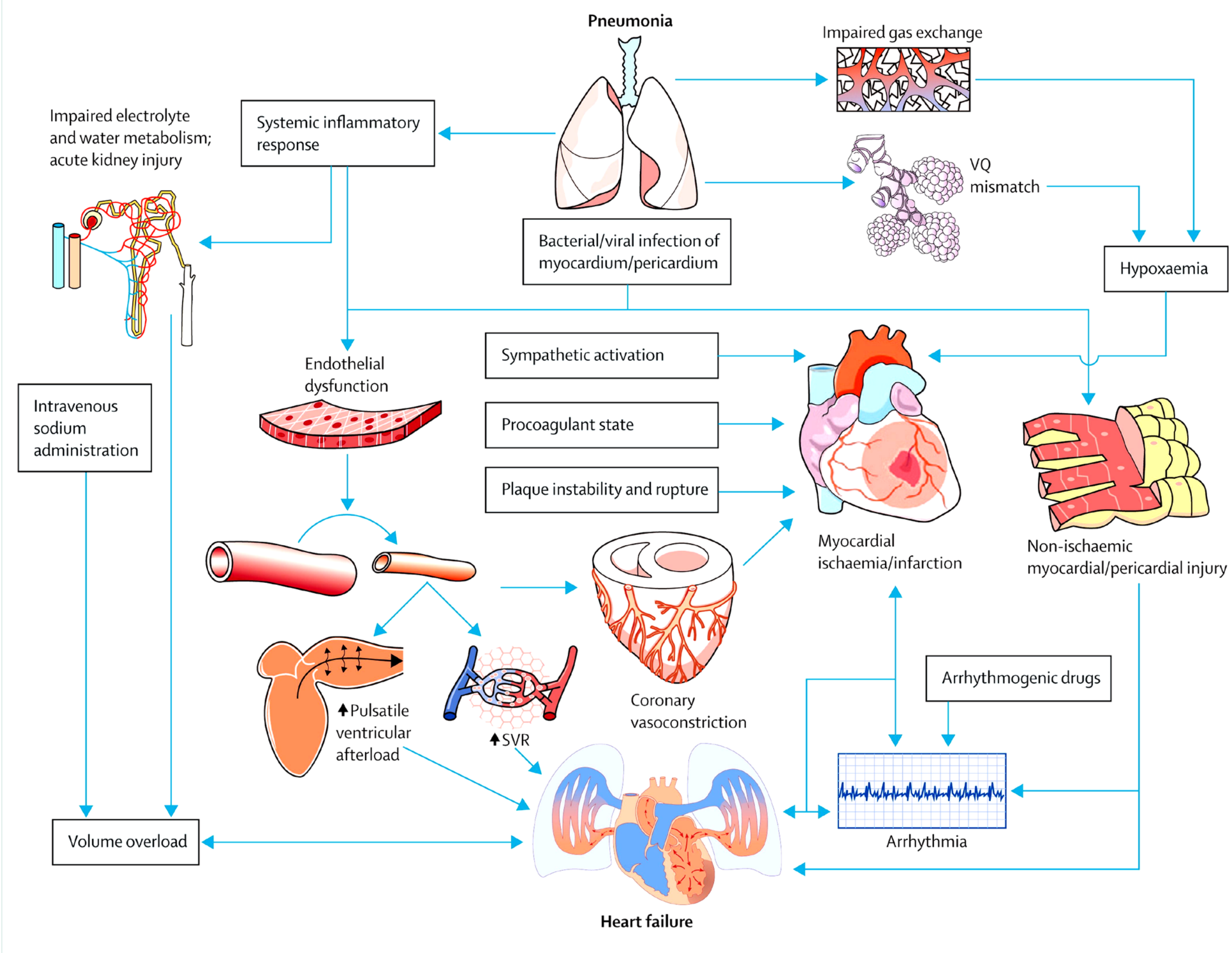

Figure 1 The modeling of COVID-19 lesions

However, from the existing research, some COVID-19 lesion modeling practices still need to be in the classic SEIR model stage, contrary to the propagation characteristics and generation mechanism of COVID-19. It leads to problems such as parameter estimation, data sampling, and model generalization.

4. Convolutional Neural Network

A convolutional neural network (CNN) is a deep learning model that can effectively process high-dimensional data such as images and extract features and patterns. In related image classification research, previous studies mainly considered fully-connected neural networks. They ignored the spatial structure of images, resulting in inaccurate feature extraction, too many parameters, and complex training [3]. Based on the above problems, we have improved the convolutional neural network for image classification, mainly using convolution technology and pooling technology, and deeply optimized the traditional neural network framework structure to make image classification's accuracy and efficiency higher. Specifically, the input image is first convoluted using multiple convolution kernels to obtain various feature maps. Second, the activation function is used to enhance the nonlinear expression ability of the network. Next, the pooling operation is used to reduce the feature map's space size, calculation amount, and risk of overfitting. Finally, the fully connected layer changes the feature map into vectors and outputs the classification results [4]. In the research, the convolution kernel is a filter used to extract local features in the image. The function of the convolution kernel is represented as follows.

$$y = f(\sum_{i=1}^{m}\sum_{j=1}^{n} w_{ij}x_{ij} + b) \tag{4}$$

where x_{ij} is the pixel value in the input image, w_{ij} is the weight value in the convolution kernel, b is the bias term, f is the activation function, and y is the output feature value. The convolutional neural network for COVID-19 imaging diagnosis based on 3D U-Net is shown in Figure 2.

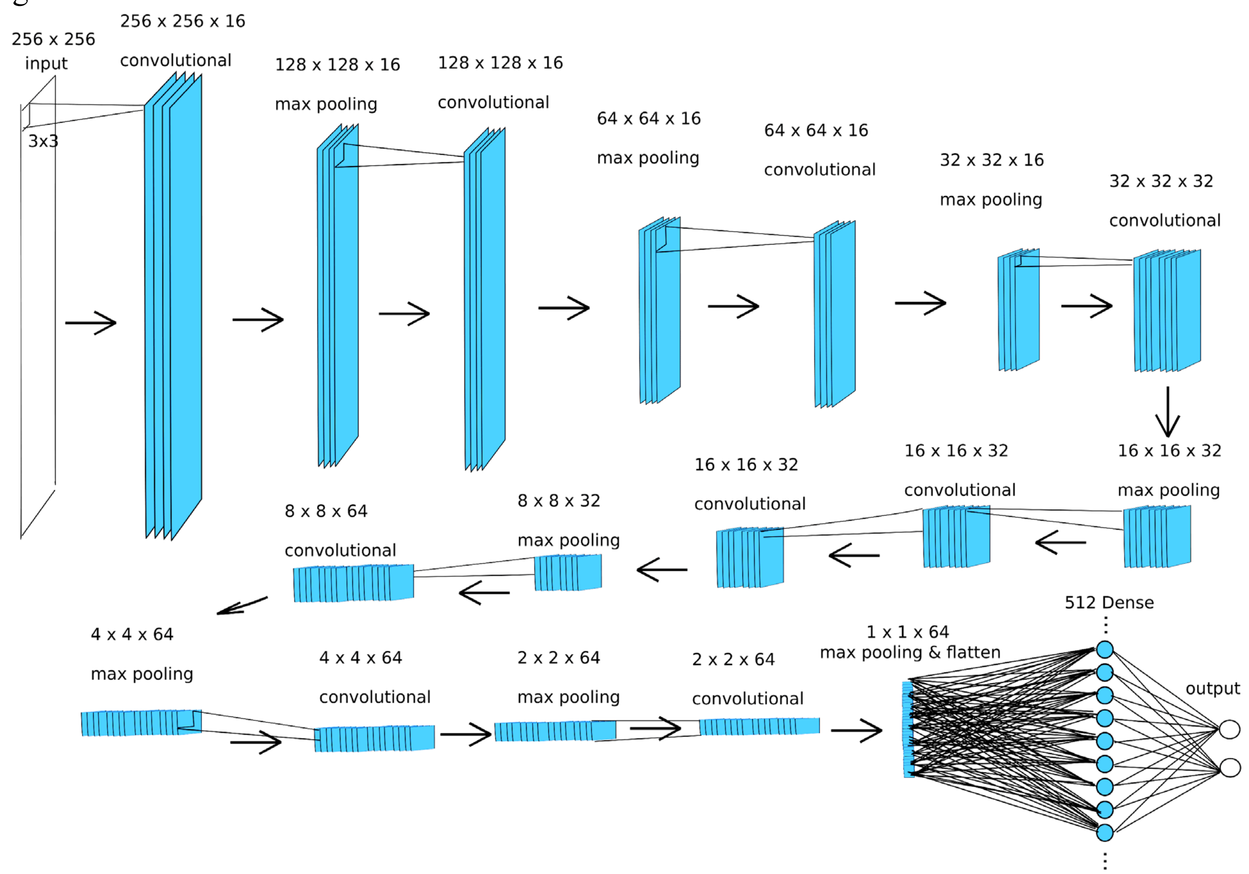

Figure 2 The convolutional neural network for COVID-19 imaging diagnosis based on 3D U-Net

5. Loss Function Principal Component Analysis

Loss Function Principal Component Analysis (LFPCA) is a data dimension reduction method based on loss function, which can be used to process functional or high-dimensional data. Compared with traditional principal component analysis (PCA), LFPCA focuses on the relationship between data structure and variability and has better flexibility and adaptability. Although some scholars have questioned that the loss function may not directly relate to data variability, most scholars advocate that LFPCA can rationally evaluate data. Scientists proposed a classical optimization model of LFPCA, including the square error and penalty terms. Since then, the model has become a typical tool of LFPCA and developed the LFPCA based on different loss functions. They believe the loss function is measurable and an “optimization criterion.” LFPCA will converge when the loss function satisfies certain conditions. Therefore, the loss function is the core of LFPCA. In addition, LFPCA is summarized into two types of models, namely the projection-based LFPCA model and the kernel-based LFPCA model. The former focuses on the linear transformation of data, and the latter focuses on the nonlinear transformation of data. LFPCA has experienced some practical failures, but from a theoretical point of view, it can provide a general dimension reduction framework, and the concept of loss function has gradually become the consensus of dimension reduction research and practice [5].

6. Experiments and Analysis

6.1 COVID-19 Lesion Identification

Semantic segmentation based on deep learning was mainly considered in the related research process of COVID-19 lesion recognition. In contrast, the texture features and variability of different lesions were ignored, so the problem of poor lesion segmentation appeared. For the above problems, we have improved the recognition of COVID-19 lesions, mainly applying texture analysis technology and a multi-task learning model, and deeply optimized the traditional UNet ++ framework to improve the recognition effect. Specifically, the texture features in CT images are first extracted, and texture descriptors of various directions and scales are obtained using methods such as gray level co-occurrence matrix (GLCM) and local binary pattern (LBP). Second, the texture features are fused with the original image features and input into the multi-task learning model. At the same time, the lesion segmentation and classification tasks are performed, and the correlation between the two tasks is used to improve the model's generalization ability. Based on the multi-task learning model, the attention mechanism and residual connection are introduced to enhance the attention and feature extraction of the lesion area. Finally, experiments are performed on public datasets to evaluate the segmentation and classification of the model. Specifically, the multi-task learning model is a machine learning that uses multiple related tasks to share information, which is used to improve the learning effect of a single task. The function of the multi-task learning model is expressed as follows:

$$L(\theta) = \sum_{i=1}^{n} \alpha_i L_i(\theta_i) + \beta R(\theta) \tag{5}$$

In the function, $L(\theta)$ is the total loss function, θ is the model parameter, n is the number of tasks, α_i is the weight coefficient of the i-th task, $L_i(\theta_i)$ is the loss function of the i-th task, β is the weight coefficient of the regularization term, and $R(\theta)$ is the regularization term .

6.2 The Influence of CT Image Value Range on Segmentation Results

From the perspective of image preprocessing, the value of a CT image cannot accurately provide the target features required by the segmentation model. The Dice coefficient mainly evaluates the segmentation results of the segmentation model. Still, the CT image lacks the relevant information and adjustment mechanism of the windows and window level, and the focus of this problem may be the CT value. In CT images, the CT value is usually described as the Hounsfield unit, and the density of the tissue structure directly reflects the degree of X-ray absorption. However, most of the CT values are about water and air, and the CT values of targets are relatively scarce. Usually, the CT value of the target tissue is challenging to obtain or measure [6]. The asymmetry of CT value and the imperfection of the window and window level directly lead to impairment of the segmentation effect. The segmentation results of the CT image value are shown in Figure 3.

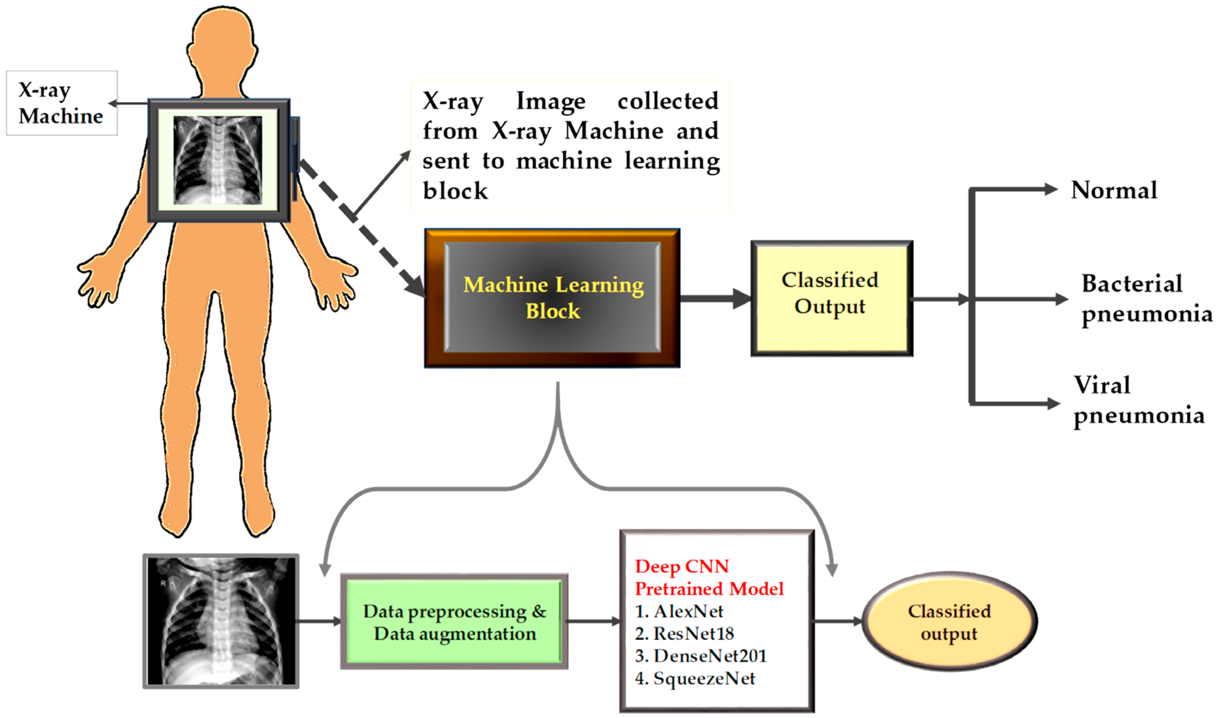

Figure 3 The segmentation results of the CT image value

6.3 The Influence of Loss Function on Segmentation Results

From the perspective of model training, the loss function has restricted the ability of the segmentation model for a long time. Since the 21st century, the segmentation model integrating deep learning has reshaped the segmentation effect through convolutional neural networks. However, the drawbacks of traditional cross-entropy loss still play a restrictive role. The cross-entropy loss is sensitive to the unbalanced number of samples and does not distinguish the problematic and easy samples, so the segmentation model needs to be improved. Under the premise of image segmentation, Dice loss is regarded as a direct way to measure the overlap between segmentation results and real labels. However, the practical effect of the segmentation model based on Dice loss remains to be discussed. In addition, due to the difficulty of Dice loss, the segmentation model needs more optimization efficiency. Therefore, Dice loss does not consistently achieve the goal of segmentation. It is not only a technical problem but also needs to improve performance and efficiency.

6.4 The Effect of PCA on the Loss Function

In fact, COVID-19 has a diagnostic problem of infectious diseases in medical imaging. Segmentation is a standard and effective image processing tool in the imaging diagnosis mechanism, which plays an essential role in lung CT images. This also makes segmentation not only related to technology but also to clinical. Therefore, image diagnosis based on accurate segmentation has become an auxiliary diagnostic mechanism for COVID-19. The actual derivation of 3D U-Net-based segmentation methods is generally based on deep learning, but the process involves many trials. From 2D to 3D, U-Net is closely around the segmentation effect. 3D U-Net should be committed to improving segmentation accuracy to meet the requirements of COVID-19 imaging diagnosis. However, insufficient data volume and noise interference also bring overfitting. To sum up, the segmentation speed and stability of 3D U-Net need to be improved, and the loss function needs to be further improved. They are essential tasks for researching COVID-19 image diagnosis based on 3D U-Net.

6.5 The Influence of Noise Error on Segmentation Results

This section proposes an Unet network based on principal component analysis. This method uses PCA to reduce the gray value of the COVID-19 CT image and extract the feature vector of the lung region. Then it is used as the input of the Unet network to train the convolutional neural network to identify the boundaries and categories of the lung area. This method can effectively segment the influence of noise error and occlusion area in CT images and has high accuracy and robustness [7]. The principal component analysis is a statistical method that is used to reduce the data dimension and retain the maximum amount of information. The function/model of principal component analysis is expressed as follows:

Given a data matrix $X \in R^{n\times d}$, where each row $x_i \in R^d$ represents a sample, solve a projection matrix $W \in R^{d\times k}$ so that the projected data matrix $Y = XW \in R^{n\times k}$ satisfies the following conditions:

First, the mean value of each column of Y is zero, $\frac{1}{n}\sum_{i=1}^{n} y_i = 0$

Second, the variance of each column of Y is maximized, $\frac{1}{n}\sum_{i=1}^{n} y_i^2 = \max_W$

Third, the covariance between different columns of Y is zero, $\frac{1}{n}\sum_{i=1}^{n} y_i y_j^T = 0, i \neq j$

Among them, n represents the number of samples, d represents the original feature dimension, k represents the feature dimension after dimensionality reduction, x_i represents the i-th sample, and y_i represents the result of the i-th sample projection.

7. Conclusions

COVID-19 has become a global public health crisis, which poses new challenges and requirements for medical imaging diagnosis. Imaging diagnosis reflects the level of medical imaging technology and is an important means of medical diagnosis. Moreover, it meets the urgent need to prevent and control the epidemic, maintain human health, and reflect the inherent medical imaging requirements. The segmentation method based on 3D U-Net is to construct the theoretical analysis framework and practical mechanism of COVID-19 image diagnosis under the guidance of deep learning. In recent years, modern information technologies such as principal component analysis have promoted the development of image processing. This study empowers the 3D U-Net network via principal component analysis and improves the accuracy and robustness of segmentation results. Its value fits the internal logic of COVID-19 imaging diagnosis. Therefore, a new approach is provided for a 3D U-Net network based on principal component analysis. In summary, the segmentation method based on 3D U-Net can be continuously improved to better diagnose COVID-19 and ensure human health.

References

[1] Çiçek Ö, Abdulkadir A, Lienkamp S S, et al. 3D U-Net: learning dense volumetric segmentation from sparse annotation[C]//Medical Image Computing and Computer-Assisted Intervention–MICCAI 2016: 19th International Conference, Athens, Greece, October 17-21, 2016, Proceedings, Part II 19. Springer International Publishing, 2016: 424-432.

[2] Daniel S J. Education and the COVID-19 pandemic[J]. Prospects, 2020, 49(1): 91-96.

[3] Wu J. Introduction to convolutional neural networks[J]. National Key Lab for Novel Software Technology. Nanjing University. China, 2017, 5(23): 495.

[4] Li Z, Liu F, Yang W, et al. A survey of convolutional neural networks: analysis, applications, and prospects[J]. IEEE transactions on neural networks and learning systems, 2021.

[5] de Almeida F A, Santos A C O, de Paiva A P, et al. Multivariate Taguchi loss function optimization based on principal components analysis and normal boundary intersection[J]. Engineering with Computers, 2022: 1-17.

[6] Diwakar M, Kumar M. A review on CT image noise and its denoising[J]. Biomedical Signal Processing and Control, 2018, 42: 73-88.

[7] Papadopoulos A. The noise error component in stochastic frontier analysis[J]. Empirical Economics, 2022: 1-35.