



Research article

COVID-19 disease identification network based on weakly supervised feature selection

Jingyao Liu^{1,2}, Qinghe Feng⁴, Yu Miao^{1,3}, Wei He¹, Weili Shi^{1,3,*} and Zhengang Jiang^{1,3,*}

¹ School of Computer Science and Technology, Changchun University of Science and Technology, Changchun, Jilin 130022, China

² School of Computer and Information Engineering, Chuzhou University, Chuzhou 239000, China

³ Zhongshan Institute of Changchun University of Science and Technology, Zhongshan 528437, China

⁴ School of Intelligent Engineering, Henan Institute of Technology, Xinxiang 453003, China

* **Correspondence:** Email: shiweili@cust.edu.cn, jiangzhengang@cust.edu.cn.

Abstract: The coronavirus disease 2019 (COVID-19) outbreak has resulted in countless infections and deaths worldwide, posing increasing challenges for the health care system. The use of artificial intelligence to assist in diagnosis not only had a high accuracy rate but also saved time and effort in the sudden outbreak phase with the lack of doctors and medical equipment. This study aimed to propose a weakly supervised COVID-19 classification network (W-COVNet). This network was divided into three main modules: weakly supervised feature selection module (W-FS), deep learning bilinear feature fusion module (DBFF) and Grad-CAM++ based network visualization module (Grad-V). The first module, W-FS, mainly removed redundant background features from computed tomography (CT) images, performed feature selection and retained core feature regions. The second module, DBFF, mainly used two symmetric networks to extract different features and thus obtain rich complementary features. The third module, Grad-V, allowed the visualization of lesions in unlabeled images. A fivefold cross-validation experiment showed an average classification accuracy of 85.3%, and a comparison with seven advanced classification models showed that our proposed network had a better performance.

Keywords: Classification; COVID-19; deep learning; feature selection; weakly supervised

1. Introduction

Respiratory infections are a major cause of disease and death worldwide. Further, pneumonia affects millions of people every year, posing a significant risk to children, adults aged 65 years and older, and individuals with health problems such as diabetes, obesity, and hypertension. Pneumonia has more than 30 different causes, but its main etiological factors are viruses and bacteria. Coronavirus disease 2019 (COVID-19) spread rapidly, resulting in a large number of deaths among those infected, and affecting economies and health systems worldwide [1]. Although a larger number of people were vaccinated, 6,833,388 deaths were reported by February 10, 2023 [2]. Computed tomography (CT) images of the chest are usually used to determine the exact area and size of the lesion. At present, CT diagnosis relies heavily on physicians' superior skills. Radiologists can make mistakes even after long clinical training and professional guidance due to the complex pathology and subtle textural changes in the images of different lung lesions. Therefore, developing intelligent classification methods for CT images to support disease identification is extremely important. Usually, these classification methods are of three types: 1) attention-guided methods; 2) methods using convolutional neural networks; and 3) label correlation methods. Attention-guided methods usually direct attention to the primary lesion region, but lack attention to the suspicious region. Methods using convolutional neural networks extract different features mainly by changing the network structure, but these methods cannot reject the background information and the acquired features are redundant. Tag correlation methods can obtain the connection between multiple tags and their interdependencies, but the images often have some noisy information that affects the multi-label learning ability. Therefore, this study was significant in examining the images after removing the background, thus effectively avoiding the interference of background information.

In recent years, artificial intelligence and swarm intelligence optimization algorithms have been increasingly used to solve biological problems [3,4], perform disease diagnosis [5–7] and address combinatorial optimization problems [8]. The rapid development and application of deep learning have enabled its extensive use in medical image processing for lesion segmentation or detection [9–12], disease classification [13], noise induction [14], image annotation [15], alignment, regression, and so forth. The continuous improvement in computer hardware and software and the emergence of more functional and powerful neural network models have made deep learning models more capable of extracting image features for better image classification. In this study, we used deep learning to study the CT classification task of COVID-19.

Deep learning has developed rapidly in recent years in the field of artificial intelligence and played an important role in diagnosing COVID-19. Most researchers are using it for intelligent diagnosis [16–18]. Wang et al. [19] proposed a model based on ELU (exponential linear unit) called ELUCNN and developed a mobile application based on it. The backbone network of this model was a 10-layer CNN which could help diagnose COVID-19. The experimental results showed that the proposed ELUCNN performed better than 14 other advanced methods. Zhang et al. [20] proposed a deep learning model SNELM using SqueezeNet as the backbone network. This model was combined with extreme learning machine as the classifier. The results of 10 cross-validations on two datasets showed that the model outperformed seven state-of-the-art classification networks. Zhang et al. [21] proposed a new deep learning network called ANC. This model could visualize Gradient-weighted Class Activation Mapping (Grad-CAM) heat map for deep learning networks and expand the training set using 18 data augmentation methods. The validation results on two COVID-19 classification

datasets showed that this model performed better than all nine currently popular classification networks. Lai et al. [22] used the segmentation network U-Net for disease diagnosis, which achieved 94.1% accuracy in diagnostic experiments on 1000 COVID-19 images due to the better performance of the U-Net network in segmentation and its less application in classification. Irmak et al. [23] proposed two new CNNs as backbone networks. One CNN could perform COVID-19 and non-COVID-19 binary classification, and the second CNN can perform triple classification. Experiments on COVID-19 chest x-ray images demonstrated its effectiveness. All the aforementioned studies used the whole image as the target of the study, and most of them used a single network to extract features without removing redundant data and extracting single features. While classifying lung diseases, the lesion region is usually located in the central region of the image, and the noise area of the edge background is extremely large, reducing the efficiency and accuracy of feature extraction. The CT images of patients with COVID-19 usually show ground glass-like changes. Also, the size of the lesion region varies depending on the disease, which requires removing the background and focusing on the lesion region. Hence, we performed weakly supervised preprocessing using the data. Also, we used two identical networks to extract complementary features and then combined them linearly to obtain better classification results considering the strong deep learning feature extraction capability. The main contributions of the method proposed in this study are as follows:

- ① A weakly supervised COVID-19 classification network (W-COVNet) was proposed, which was mainly divided into three modules: weakly supervised feature selection module (W-FS), deep learning bilinear feature fusion module (DBFF) and Grad-CAM++ based network visualization module (Grad-V).
- ② W-FS mainly learned the lung area features by training Data2 to get the segmentation model and then performed feature selection on Data1 using the segmentation model. The new dataset Data1-Seg was obtained through multiple attempts of merit seeking. Five multiple-way data augmentation methods were used for data enhancement to expand the training set.
- ③ DBFF mainly used migration learning methods with the lightweight network VGG as the backbone network. The network used two identical groups of networks focusing on different regions to extract different features for linear fusion, improving the feature extraction capability. Also, four DBFF schemes were proposed.
- ④ Grad-V focused on the visual display of deep learning networks using Grad-CAM++. With this module, it was possible to show whether the network could accurately identify focal areas and interpret the classification results favorably.
- ⑤ The proposed network, W-COVNet, was compared with the seven most popular current classification networks using a fivefold cross-validation method on the same dataset, and the results showed that W-COVNet outperformed the other networks.

This manuscript is structured as follows. In Section 1, we describe diagnostic methods and deep learning-related research of COVID-19. In Section 2, recent studies on COVID-19 disease diagnosis are also discussed. In Section 3, weakly supervised dataset preprocessing methods are introduced, and new models for classification networks and Grad-CAM++ visualization methods are proposed. In Section 4, the specific steps, methods and results used for the experiments are explained and compared with other methods. In Section 5, the results are discussed and analyzed based on the results obtained from the model, and finally conclusions are presented.

2. Related works

The research of the intelligent diagnosis of diseases can help doctors accurately diagnose various diseases. Xie et al. [24] found that the aberrations in certain genes affect adjacent genes, thus leading to the development of cancer. Further, the small size of cancer samples does not allow easy detection of cancer-causing genes. Therefore, a conceptual learning model was proposed to explore the characteristics of genes effectively via a dimensionality reduction process. Huang et al. [25] proposed an algorithm based on BI-RADS and CART using a decision tree algorithm to classify breast ultrasound data and identify malignant and benign tumors. This solved the problem of the inability of CAD systems to handle data from different sources and the problem of limited transparency to physicians. AI is not interpretable, and therefore AI-based drug recommendations are less credible, Xi et al. [7] used the proposed evaluation metric of traceability rate in an interpretable AI drug recommendation model, leading to an improvement in model performance. Dong et al. [26] proposed MorbidGCN, a GCN-based multi-disease prediction network, taking into account the differences between patients with single and multiple diseases and the relationship between multiple diseases. The network integrated population phenotype information and disease network information and performed better on two large datasets.

Usually, image classification and recognition are performed using global features. For example, Li and Liu [27] used a combination of wavelet packet Tsallis entropy, feedforward neural network and real number coding biogeography-based optimization (RCBBO) methods for the pathological examination of the brain and found optimal weights and biases. Fulton et al. [28] built a machine learning model using ResNet-50 to classify Alzheimer's disease. The prediction accuracy was 99% for the validation set triple classification and 99.34% for the training set.

However, different diseases and people have different sizes of diseased areas, and their locations are difficult to predict. Therefore, a large number of noisy areas exist when using global images for feature extraction. If the lesion area can be found accurately, its features can be extracted more easily, thus improving the classification accuracy. Hence, two aspects can be considered. First, methods such as the attention mechanism can be used to guide the network to focus on the lesion area; second, useless background areas can be removed so that the network can narrow its focus and improve its ability to obtain valid features. Guan et al. [29] proposed a three-branch attention-guided ResNet50 as the backbone of the convolution neural network (AG-CNN), which first acquired the mask of the focal area under the attention heat map guidance, cut it out as the branch feature extraction region and fused the global and local features. The final experimental result on the ChestX-ray14 dataset was an average Area under curve (AUC) value of 0.841. Ypsilantis et al. [30] mimicked the human visual attention mechanism using recurrent attention. Pesce et al. [31] proposed two network architectures, one consisting of a recurrent attention model and the other using weak markers and manually circling bounding boxes, which could accurately circle lung nodules. Togacar et al. [32] fused features from SqueezeNet and MobileNetV2 and used a new feature selection and combination method: social mimic optimization. Finally, the classification accuracy of COVID-19 reached 100% using a support vector machine for classification, and this model could be applied to mobile devices. Cohen et al. [33] proposed a COVID-19 virus severity network (CSSNet). First, seven large non-COVID-19 chest x-ray datasets were pretrained to obtain correlation features. The severity of infection was then evaluated on a small number of COVID-19 datasets. Ni et al. [34] used the proposed NiNet to segment and detect lesion regions on CT images of a few patients with COVID-19. Ko et al. [35] used a migration learning

approach to find the optimal one from four commonly used deep learning networks as the backbone network. The proposed deep learning framework FCONet was used to tri-classify the enhanced images with an accuracy of 96.97%. Wang et al. [36] first obtained a lung segmentation model in an unsupervised manner and then proposed a new lightweight method based on the weakly supervised detection of COVID-19. Khan et al. [37] fused two different datasets, pretrained them using migration learning, and then performed the classification and detection of COVID-19 with 98.2% accuracy for triple classification and 95% accuracy for quadruple classification. They named the proposed network as Coro-Net. Hussain et al. [38] used a fivefold cross-validation for COVID-19 and normal CXR dichotomous classification with accuracy, sensitivity, and specificity results of 100%.

The aforementioned related studies for lung segmentation mainly focused on x-ray images, with fewer studies on the segmentation of CT images. Most of the new methods for the classification of COVID-19 extracted global features, but with the problem of feature redundancy. In this study, a weakly supervised method was used to obtain lung fields for COVID-19 CT images and perform feature selection. We improved the deep learning VGG network and proposed a new bilinear method to obtain effective features.

3. Methodology

3.1. Dataset and preprocessing

Two main datasets were used in this study, both of which were obtained from open-source websites. The first dataset (Data1) contained 397 non-COVID-19 CT images and 349 COVID-19 CT images. This dataset was validated by radiologists at Tongji Hospital in Wuhan, China [39]. The second dataset (Data2), COVID-19 CT segmentation [40], had 100 axial CT images from more than 40 patients. This dataset had COVID-19 lung CT masks that were labeled, it was mainly used for effective semi-supervised lung region segmentation. All CT images were collected by the Italian Society of Medical and Interventional Radiology.

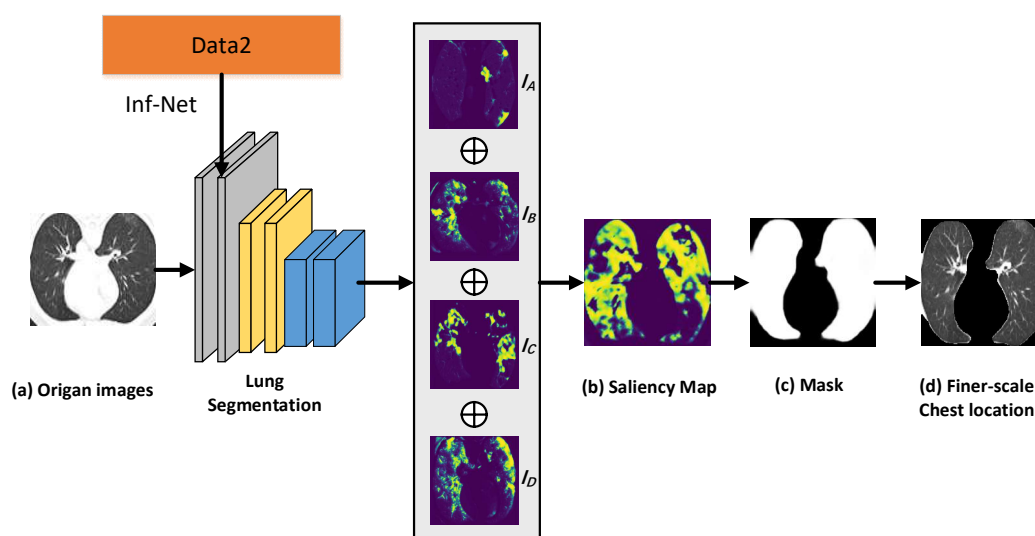


Figure 1. Flow chart of optimization-seeking weakly supervised segmentation preprocessing module.

In this study, we introduced the Inf-Net [41] segmentation model. The backbone of this model was Res2Net, which could be trained on a small number of labeled images via a reverse attention mechanism and showed edge detection attention. It could effectively perform lesion segmentation and showed the best performance compared with five currently popular segmentation networks. Six different metrics were used for performance evaluation in the COVID-19 lesion segmentation dataset, namely, *Dice*, Specificity (*Spec.*), Sensitivity (*Sen.*), Structure Measure (S_α), Enhanced-alignment Measure (E_ϕ^{mean}) and Mean Absolute Error (*MAE*) (Table 1). As lung segmentation was required in this study, the first step involved obtaining the edge image from the ground truth (GT) map of the lung region of the Data2 dataset with lung region markers, and the second step used the GT map with the edge map to apply Inf-Net training to obtain the lung segmentation. The third step employed the obtained lung segmentation via the semi-supervised segmentation network Semi-Inf-Net [41] using the lung segmentation parameters to segment the lungs of Data1. In the fourth step, the weakly supervised feature selection module (W-FS) method was proposed because some bias existed in the training process, and some images had poor lung segmentation, as shown in Figure 1. Further, features within the lung region were removed. Data1 was trained three times for lung region segmentation, and the image with the best result was selected for data optimization. Finally, a new complete lung region image dataset Data1-Seg was obtained.

Table 1. Performance comparison of Inf-Net with five popular segmentation networks [41].

Methods	<i>Dice</i> ↑	<i>Sen.</i> ↑	<i>Spec.</i> ↑	S_α ↑	E_ϕ^{mean} ↑	<i>MAE</i> ↓
U-Net [36]	0.439	0.534	0.858	0.622	0.625	0.186
Attention-UNet [37]	0.583	0.637	0.921	0.744	0.739	0.112
Gated-UNet [38]	0.623	0.658	0.926	0.725	0.814	0.102
Dense-UNet [39]	0.515	0.594	0.840	0.655	0.662	0.184
U-Net++ [35]	0.581	0.672	0.902	0.722	0.720	0.120
Inf-Net [34]	0.682	0.692	0.943	0.781	0.838	0.082

Note: ↑ denotes that higher is better, and ↓ denotes that lower is better.

3.2. Data augmentation on the training set

Data1-Seg contained two types of CT images: COVID-19 and non-COVID. We resized the set D^1 of the Data1-Seg to a consistent size of 224×224 to obtain the new set D^2 of images, as shown in Eq (1).

$$D^2 = \text{Resize}(D^1, [224, 224]) = \{d_1^1, d_2^1, d_3^1, \dots, d_n^1\} \quad (1)$$

The dataset was randomly divided into two subsets by the random hold-out method: the testing set (X: 20%) and the training set (Y: 80%). The relevant data partitioning information is shown in Table 2. The relevant sizes related to these subsets satisfied Eq (2). To address the issue of limited data, the MDA [42] technique was applied to expand the data and reduce the occurrence of overfitting. Five

data enhancement methods were applied to the training set, turning it from the original one image to six images.

$$|D^2| \xrightarrow{RHO} |X| + |Y| = |\{x_1, \dots, x_i\}| + |\{y_1, \dots, y_j\}| \quad (2)$$

where i is the number of X (testing set) images, j is the number of Y (training set) images, and $|\cdot|$ refers to the cardinality of a set.

Table 2. Distribution of Data1-Seg dataset in the model.

Dataset	COVID-19	non-COVID	Total
Training (80%)	280	318	598
Testing (20%)	69	79	148
Total (100%)	349	397	746

It was assumed that K_{MDA} MDA techniques ($K_{MDA} = \{K_1, K_2, K_3, K_4, K_5\}$) were used. In this study, K_{MDA} MDA techniques included gamma correction, rotation, noise injection, brighter and darker. The N_{tra} training set images were generated using each MDA technique, and finally, $K_{MDA} \times N_{tra}$ images were generated for all MDAs. The following five MDA techniques were mainly used in this study:

① Gamma correction (Ga_Co)

The gamma correction factor $r^{G-C} = 1.5$ was used to produce new images as follows:

$$\overline{y^{K_1}(i)} = Ga_Co[y(i)] = [y_1^{K_1}(1, r^{Ga_Co}), \dots, y_{N_{tra}}^{K_1}(i, r^{Ga_Co})] \quad (3)$$

② Rotation (Rot)

The rotation angle $\theta^{Rot} = 90^\circ$ was applied to the images:

$$\overline{y^{K_2}(i)} = Rot[y(i)] = [y_1^{K_2}(1, \theta^{Rot}), \dots, y_{N_{tra}}^{K_2}(i, \theta^{Rot})] \quad (4)$$

③ Noise injection (No_In)

Gaussian noise was injected into the image to get a new image:

$$\overline{y^{K_3}(i)} = No_In[y(i)] = [y_1^{K_3}(1), \dots, y_{N_{tra}}^{K_3}(i)] \quad (5)$$

④ Brighter (Bri)

$$\overline{y^{K_4}(i)} = Bri[y(i)] = [y_1^{K_4}(1), \dots, y_{N_{tra}}^{K_4}(i)] \quad (6)$$

⑤ Darker (Dar)

$$\overline{y^{K_5}(i)} = Dar[y(i)] = [y_1^{K_5}(1), \dots, y_{N_{tra}}^{K_5}(i)] \quad (7)$$

$$\overline{y^K(i)} = concat[\overline{y^{K_1}(i)}, \overline{y^{K_2}(i)}, \overline{y^{K_3}(i)}, \overline{y^{K_4}(i)}, \overline{y^{K_5}(i)}] \quad (8)$$

where $\overline{y^K(i)}$ means that the data augmentation was the concatenation of five MDA results.

$$y(i) \xrightarrow{MDA} \text{concat} \left[y(i), \overline{y^K(i)} \right] \quad (9)$$

where $y(i)$ means that the training set consisted of the original and augmented images.

3.3. Bilinear feature extraction

Transfer learning uses old knowledge to learn new knowledge, and the main goal is to transfer the learned knowledge to a new domain quickly. In this study, we used the idea of transfer learning to migrate the network after training in ImageNet, which could reduce the network parameters and converge quickly. Further, we used VGG for migration learning, a simple and robust network with four main structures: VGG11, VGG13, VGG16 and VGG19. We conducted experiments on each of them. Figure 2 shows the structure of VGG16.

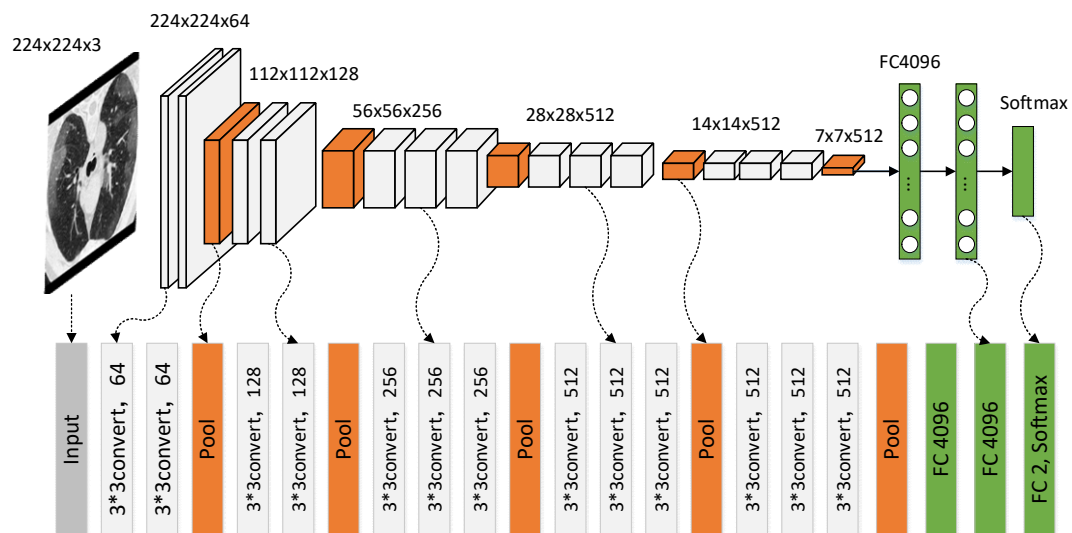


Figure 2. Diagram of the VGG16 structure.

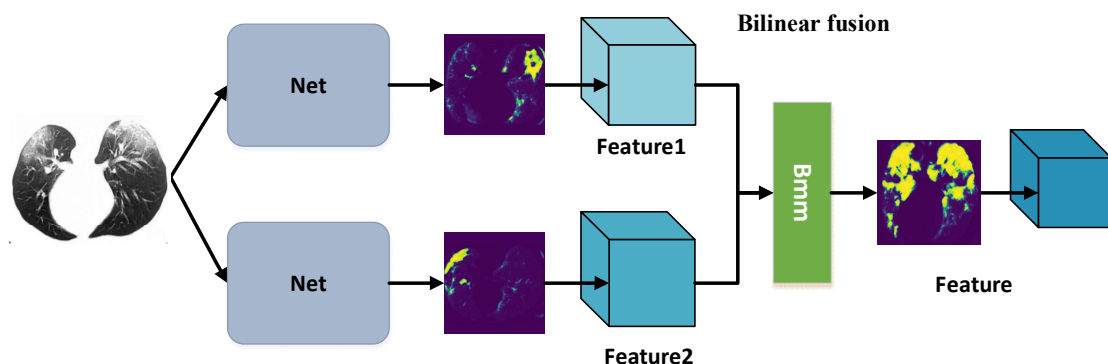


Figure 3. Flow chart of the symmetric complementary bilinear module.

The symmetric complementary bilinear classification network had two feature extractors, and the two sets of extracted features had local complementary characteristics for fine-grained images with scattered lesion areas such as COVID-19. Then, linear multiplication was performed to combine the features efficiently and obtain richer features. Figure 3 shows the symmetric complementary bilinear module using two VGG networks.

3.4. Grad-CAM++ based network visualization

Deep learning is not interpretable and is usually understood as a black box. Grad-CAM (Gradient-weighted Class Activation Mapping) [43] is a heat map that can be loaded with network parameters and model structures to visualize deep learning networks. Usually, the network feature maps are processed by using global average pooling (Figure 4), and each feature map contained information of different classes.

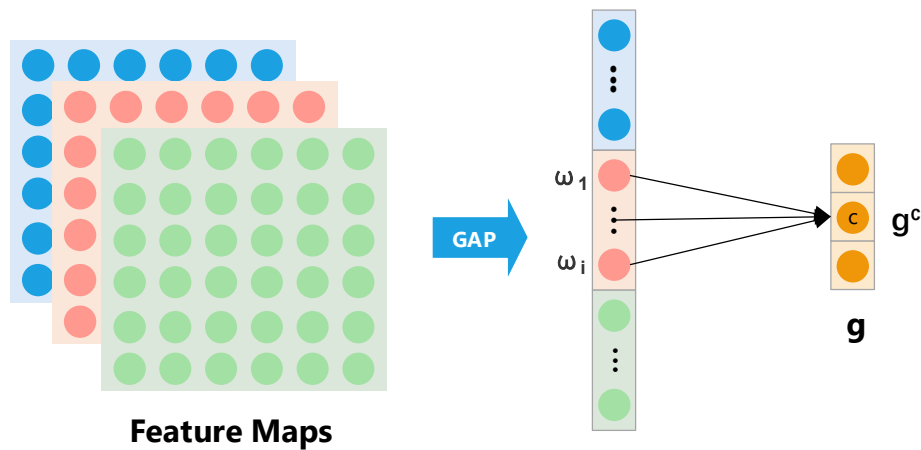


Figure 4. Global average pooling operation.

In Grad-CAM, the weights are calculated by averaging the gradients of the feature maps globally. The more accurate features are further back in the network. Grad-CAM usually shows the features of the last convolutional layer, as shown in Eq (10).

$$\omega_k^c = \frac{1}{A} \sum_i \sum_j \frac{\partial g^c}{\partial B_{ij}^k} \quad (10)$$

where ω_k^c denotes the weight of the k th feature map corresponding to category c , and by using the heat map, gradient is calculated. A denotes the number of feature map pixels, g^c denotes the gradient score of category c , and B_{ij}^k denotes the pixel value at the (i, j) coordinate in the k th feature map.

The heat map is obtained by first obtaining the category weights corresponding to all the feature maps and then performing a weighted summation, as shown in Eq (11).

$$L_{Grad-CAM}^c = \text{relu} \left(\sum_i \omega_k^c B^k \right) \quad (11)$$

where B^k denotes the k th feature map. Grad-CAM++ [44] is more effective than Grad-CAM, mainly because the weights are more complex and accurately obtained, and the positioning is precise.

3.5. Proposed approach

In this study, we proposed a weakly supervised COVID-19 classification network (W-COVNet), which consisted of three main parts: Part 1, the weakly supervised feature selection module, which was mainly used for image lung segmentation; Part 2, the deep learning bilinear feature fusion module; and Part 3, the Grad-CAM++ based network visualization module. W-COVNet could perform feature selection and focus on effective regions, thus improving classification accuracy while visualizing focal areas. Figure 5 shows the overall framework. To elaborate further, pseudo code for the W-COVNet algorithm is given in Algorithm 1.

Algorithm 1. Pseudo code of our W-COVNet algorithm.

Phase I: W-FS	
Step 1	Input: Original image set Data1 and Data2.
Step 2	Train Data2 to get lung field segmentation set L.
Step 3	Use weakly supervised learning method to train the Data1 by L.
Step 4	Get the lung field of Data1 set L_F.
Step 5	Segment the lung field of the original image Data1 by L_F using crop.
Step 6	Obtain the Data1-Seg through the lung field with background removed set D ¹ .
Step 7	Resizing: Resize the image D ¹ to [224, 224] and get dataset D ² . See Eq. (1).
Step 8	Y _{RHO} : testing set (X: 20%) and training set (Y: 80%). See Eq. (2).
Step 9	MDA(Y): Ga_Co, No_In, Rot, Bri, and Dar to augment training set Y.
Phase II: DBFF	
Step 10	Read one raw pretrained model VGG16.
Step 11	Train two groups of VGG16 in parallel to obtain different features.
Step 12	Obtain two sets of features F1 and F2.
Step 13	Bmm (F1, F2).
Step 14	Generate DBFF.
Phase III: Grag-V	
Step 15	Read a DBFF bilinear COVID-19 classification model set M.
Step 16	Read a COVID-19 image set I.
Step 17	Grad-CAM++ visualization of image I using model M.
Step 18	Obtain the heat map I _{Grad} after visualization.
Step 19	Generate Grag-V.
Step 20	Output: The model W-COVNet and its performances.

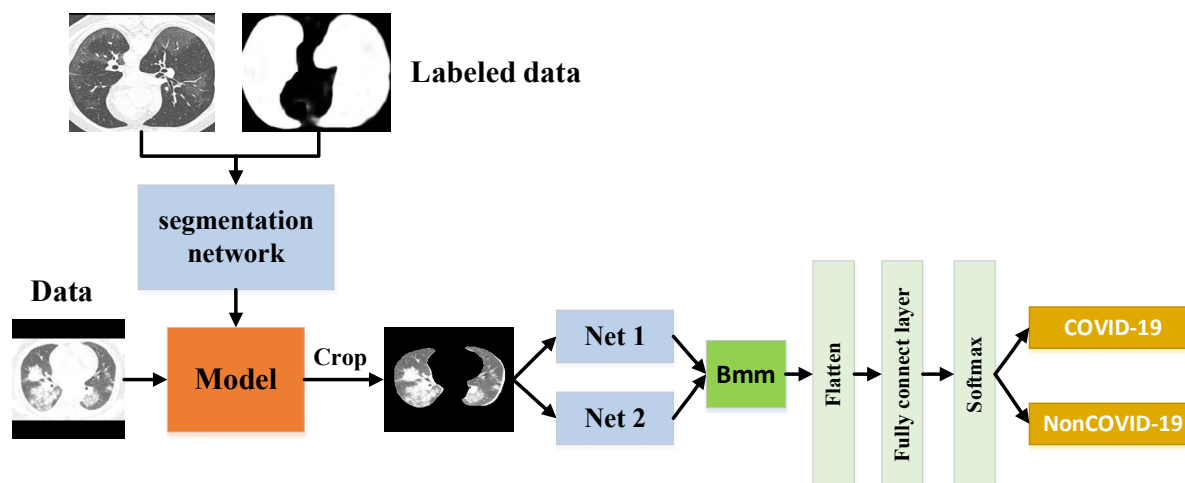


Figure 5. Structure of the proposed W-COVNet.

4. Experiments and results

4.1. Experimental platform

The study was executed in a Linux environment on a 32 GB Tesla V100 graphics card-equipped NVIDIA DGX Station deep learning workstation. The complete code, including data preprocessing and algorithm implementation, was implemented in Python. It made use of libraries such as NumPy and PyTorch, a deep learning toolset. The model's tuned hyper-parameters were the epochs, learning rate (LR), batch size (BS), and dropout rate (DR). The best experimental results were obtained when epochs = 50, LR = 0.003, BS = 8 and DR = 0.4.

4.2. Classification performance

We used various metrics in the testing set, namely, accuracy (Acc), precision (Pre), sensitivity (Sen), specificity (Spe), recall (Rec) and F1-score ($F1-sc$), to evaluate the performance of the proposed W-COVNet method. The corresponding equations were as follows.

$$Acc = \frac{TP(Y) + TN(Y)}{TP(Y) + TN(Y) + FP(Y) + PN(Y)} \quad (12)$$

$$Pre = \frac{TP(Y)}{TP(Y) + FP(Y)} \quad (13)$$

$$Sen = \frac{TP(Y)}{TP(Y) + FN(Y)} \quad (14)$$

$$Spe = \frac{TN(Y)}{TN(Y) + FP(Y)} \quad (15)$$

$$Rec = \frac{TP(Y)}{TP(Y) + FN(Y)} \quad (16)$$

$$F1-sc = \frac{2 * Pre(Y) * Rec(Y)}{Pre(Y) + Rec(Y)} \quad (17)$$

For the four forms of VGG, we proposed four network schemes with VGG11, VGG13, VGG16 and VGG19 as the backbone networks, as shown in Table 3. We created a histogram display of the results, which are shown in Figure 6. Accuracy refers to the ratio of the number of correctly predicted samples to the total number of predicted samples. Precision is the proportion of positive samples that are actually predicted to be positive. The larger the value of sensitivity is, the greater the “sickness is judged as sickness” and the smaller the “missed detection” (FN). The larger the value of specificity is, the larger the “healthy are judged as healthy” and the smaller the “false detection” (FP). The f1-score (equilibrium average) is the result of the calculation that takes into account the accuracy and completeness of the model, and the value is more inclined to the smaller value of the index. The f1-score indicates how many of the data with the correct real value can be predicted correctly. Table 3 shows that VGG16 results were balanced and performed best in overall prediction ability, with the ability to predict positive samples correctly, and VGG16 also had the smallest false detection rate. The best results were obtained when VGG16 was used as the backbone network. In conclusion, the performance in terms of accuracy, sensitivity, recall and F1-score on the testing set was good. Therefore, the W-COVNet proposed in this study was effective.

Table 3. Classification of W-COVNet networks after using two kinds of testing set (%).

Network	Class	Acc	Sen	Pre	Rec	Spe	F ₁ -sc
[VGG11] ²	COVID-19	84.5	82.6	83.8	82.6	86.1	83.2
	non-COVID	84.5	86.1	85.0	86.1	82.6	85.5
[VGG13] ²	COVID-19	82.4	82.6	80.3	82.6	82.3	81.4
	non-COVID	82.4	82.3	84.4	82.3	82.6	83.3
[VGG16] ²	COVID-19	86.5	87.0	84.5	87.0	86.1	85.7
	non-COVID	86.5	86.1	88.3	86.1	87.0	87.2
[VGG19] ²	COVID-19	83.1	84.1	80.6	84.1	82.3	82.3
	non-COVID	83.1	82.3	85.5	82.3	84.1	83.9

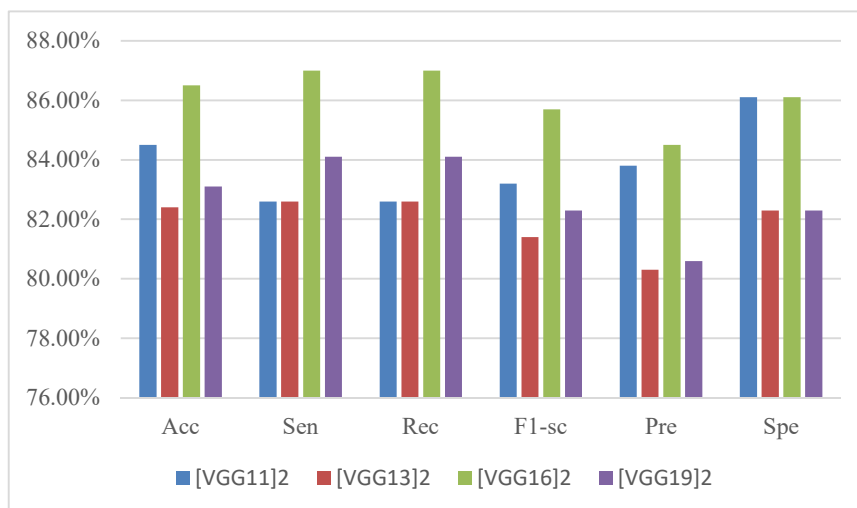


Figure 6. Comparison of the experimental results of the four schemes.

Figure 6 depicts the W-COVNet training process. Scheme 3 with the best results was selected to plot the training loss and the testing accuracy under each epoch. The horizontal axis is the epoch, the left vertical axis is the test accuracy, and the right vertical axis is the training loss. Figure 7 shows that the training loss became smaller with the iteration of epoch until it reached 0.1 and no longer changed, while the test accuracy became larger with the increase in epoch until it reached 0.86 and no longer changed, indicating that the network converged around the tenth epoch and performed well with high test accuracy.

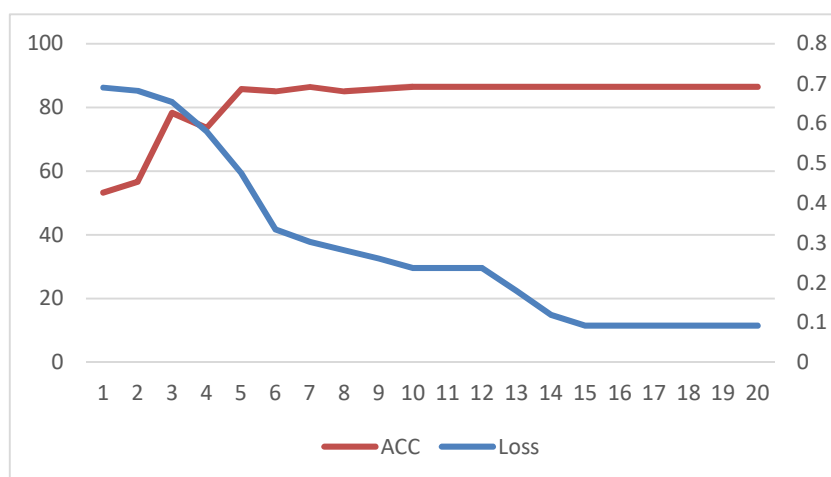


Figure 7. Test accuracy and training loss of W-COVNet.

4.3. Confusion matrix

We used the confusion matrix [45] to display the test results of the experimental data. For each class $C = 1, 2$ (1: COVID-19, 2: non-COVID), we set the COVID-19 class to “positive” and the other class to “negative”. The true negatives, true positives, false negatives and false positives were used to

identify the diagnosis of CT images with the model. Figure 8 shows the confusion matrix of the W-COVNet with VGG11, VGG13, VGG16 and VGG19 as the backbone proposed in this study.

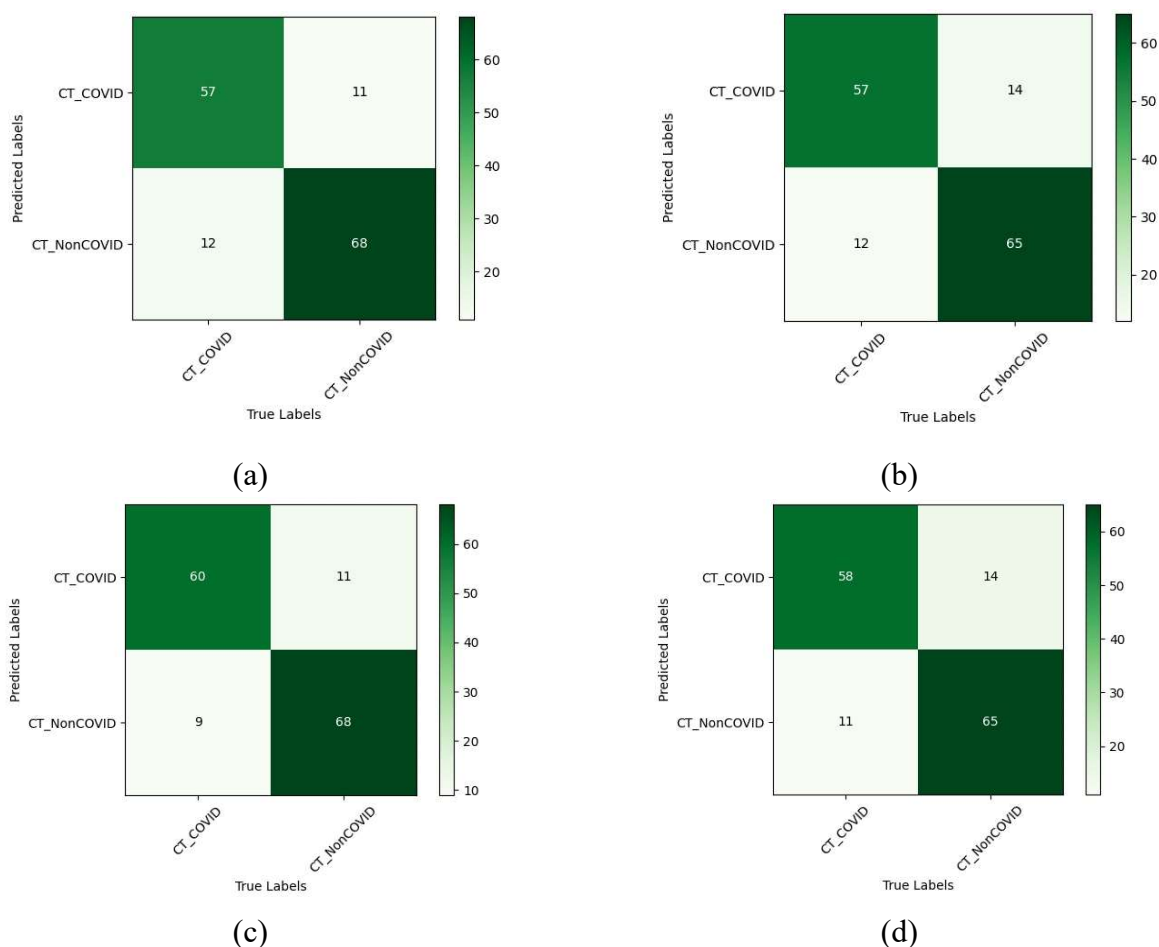


Figure 8. Confusion matrix of W-COVNet with VGG11 (a), VGG13 (b), VGG16 (c) and VGG19 (d) as the backbone.

4.4. 5-Folds cross-validation

We divided the data into a training set and a testing set due to the small amount of data trained and to prevent overfitting problems during the study. The data were divided into five equal parts; one part of each experiment was used for testing, and the remaining parts were used for training. The experiment was repeated five times to find the average. The first experiment used 20% of the data in the red area as the test set and the rest as the training set. The second experiment used 20% of the second red area as the test set and the rest as the training set, and so on, so that the experimental results were more accurate. A 5-Folds multiple partitioning format allowed the use of the full dataset. Finally, the averaging method could accurately represent the model performance and verify the generalization performance of the network.

We experimentally determined that option 3 was the best result. Therefore, W-COVNet selected VGG16 as the backbone and performed a 5-folds cross-validation using the new dataset Data1-Seg. The results are shown in Table 4.

Table 4. Classification of W-COVNet networks after using two kinds of the testing set (%).

Run	Class	Acc	Sen	Pre	Rec	Spe	F ₁ -sc
R1	COVID-19	85.8	85.5	84.3	85.5	86.1	84.9
	non-COVID	85.8	86.1	87.2	86.1	85.5	86.6
R2	COVID-19	86.5	87.0	84.5	87.0	86.1	85.7
	non-COVID	86.5	86.1	88.3	86.1	87.0	87.2
R3	COVID-19	83.3	80.0	83.6	80.0	86.3	81.8
	non-COVID	83.3	86.3	83.1	86.3	80.0	84.7
R4	COVID-19	84.6	87.1	81.3	87.1	82.3	84.1
	non-COVID	84.6	82.3	87.8	82.3	87.1	85.0
R5	COVID-19	86.5	87.0	84.5	87.0	86.1	85.7
	non-COVID	86.5	86.1	88.3	86.1	87.0	87.2
Mean ±	COVID-19	85.3 ± 1.37	85.3 ± 3.04	83.6 ± 1.35	85.3 ± 3.04	85.4 ± 1.72	84.4 ± 1.64
SD	non-COVID	85.3 ± 1.37	85.4 ± 1.72	87.0 ± 2.19	85.4 ± 1.72	85.3 ± 3.04	86.1 ± 1.22

4.5. Ablation study

We conducted ablation experiments to verify the effectiveness of the proposed W-COVNet method. We used VGG16 as the backbone network to conduct ablation experiments. Experiment No.1 used the backbone network VGG16 to classify the Data1-Seg data after feature selection. Experiment No.2 is the result after training the Data1-Seg using feature fusion. Experiment No.3 was the result after preprocessing the images and the training set had enhanced data. Experiment No.4 showed the result after performing feature selection, data enhancement and feature fusion. As shown in Table 5, experiment No.1, No.2, No.3 and No.4 validated the effectiveness of several improvements. Among these, Experiment No.2 and No.3 validate the effectiveness of DBFF data fusion and MDA data enhancement, respectively. Finally No.4 showed the performance of our proposed method W-COVNet. As shown in Table 5, the performance of W-COVNet was the best, and the method was effective.

Table 5. Ablation studies of our W-COVNet (%).

Methods	Acc	Sen	Pre	Rec	Spe	F ₁ -sc
(No.1) Backbone	74.47	65.22	78.95	65.22	83.33	71.43
(No.2) Backbone + DBFF	79.43	71.01	84.48	71.01	87.50	77.17
(No.3) Backbone + MDA	81.76	86.96	76.92	86.96	77.22	81.63
(No.4) Backbone + MDA + DBFF	86.49	86.96	84.51	86.96	86.08	85.71

4.6. Grad-CAM++ based network visualization results

The objective of this study was to use the image set labeled data without pixel labeling of the lesion area. However, the deep learning network could find the lesion by learning, allowing for accurate disease diagnosis. Therefore, the Grad-CAM++ technique was used for visualization operation. Three images were selected for each category, namely COVID-19 and non-COVID. Each image was divided into three groups: the first group was the original CT image; the second group was the image of the lung area after removing the background and keeping only the feature area; and the

third group was the heat map visualized using Grad-CAM++. As shown in Figure 9, the red area is the feature area, and the blue area is the non-feature area. If the features are denser, the corresponding feature values increase, leading to a brighter color. The displayed results showed that using Grad-CAM++ could not only provide an effective interpretation of the deep learning network but also help doctors find the lesion areas.

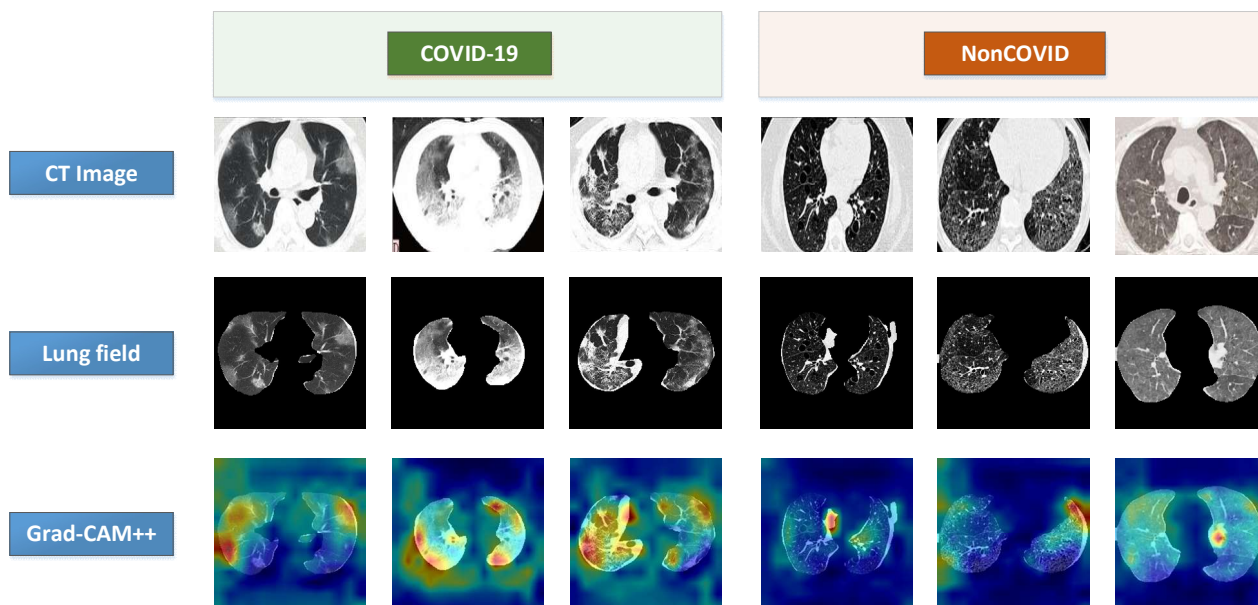


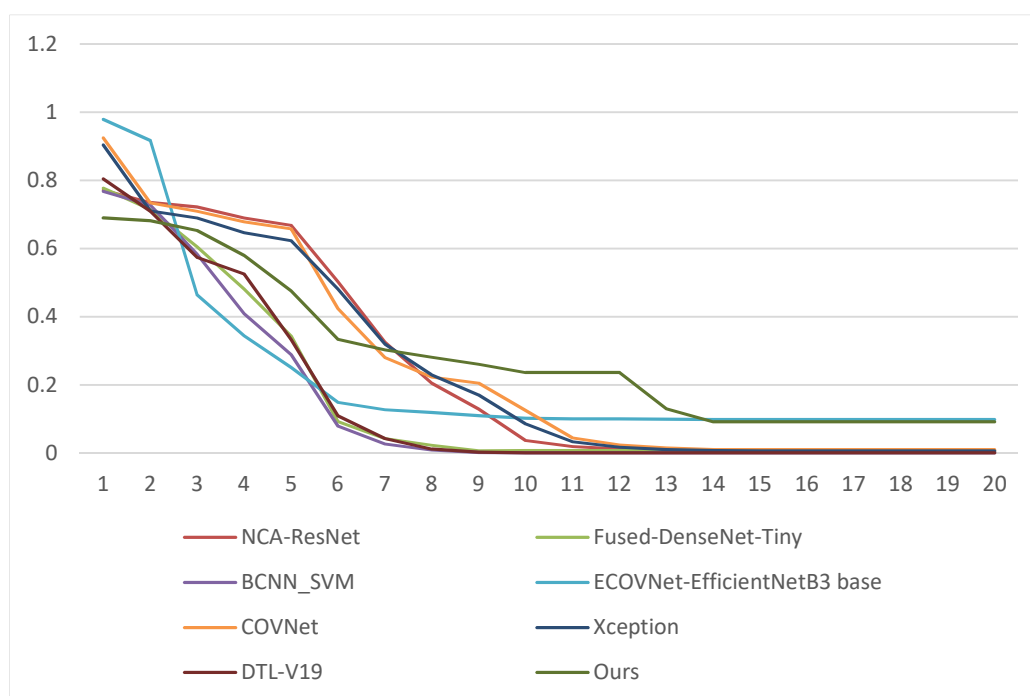
Figure 9. Use Grad-CAM++ + to visualize the feature area.

4.7. Comparison experiments

We compared the W-COVNet method with seven types of better-performance classification methods: NCA-ResNet [46], Fused-DenseNet-Tiny [47], BCNN_SVM [48], ECOVNet-EfficientNetB3 base [49], COVNet [35], Xception [50] and DTL-V19 [51] to demonstrate its effectiveness. All experiments used the dataset Data1-Seg and the same data preprocessing and performed a fivefold cross-validation. The experimental results are shown in Table 6. Figure 10 and Figure 11 show the test accuracy and training loss trend of our method with these seven comparison experiments, and it can be seen that all converge within 20 epochs, and our proposed W-COVNet has the highest accuracy after convergence. It was observed that our proposed W-COVNet method performed the best, with an accuracy of 85.3%. The extracted features were more accurate using the images with the background removed. The use of two symmetric networks at the same time allowed focus on different features and improved the feature richness. The effectiveness of the network was verified by the experiments. The use of MDA data augmentation prevented the overfitting of the model, thus improving its performance. In addition, the BCNN_SVM backbone network was VGG16 and VGG19, and the DTL-V19 backbone network comprised VGG19. In this study, we also used the VGG network as the backbone network. Comparing the experimental results, we could see that VGG performed better than other networks on this dataset.

Table 6. Performance comparison of the proposed W-COVNet with other studies (%).

Method	Class	Pre	Spec	F1-sc	Acc
NCA-ResNet	COVID-19	65.9 ± 4.35	72.7 ± 5.31	62.3 ± 2.53	66.4 ± 2.85
	non-COVID	67.0 ± 1.96	59.2 ± 2.89	69.7 ± 3.23	
Fused-DenseNet-Tiny	COVID-19	64.1 ± 3.41	70.5 ± 5.18	61.9 ± 6.86	65.7 ± 3.95
	non-COVID	67.3 ± 4.61	60.4 ± 4.18	68.6 ± 3.11	
BCNN_SVM	COVID-19	80.9 ± 2.99	83.1 ± 4.98	80.2 ± 4.84	81.7 ± 3.26
	non-COVID	83.3 ± 6.13	80.2 ± 9.86	82.9 ± 2.49	
ECOVNet-EfficientNetB3 base	COVID-19	71.1 ± 2.70	74.0 ± 3.64	71.8 ± 3.64	73.4 ± 2.63
	non-COVID	75.8 ± 4.25	72.7 ± 6.70	74.9 ± 1.98	
COVNet	COVID-19	62.7 ± 4.39	76.8 ± 10.10	49.1 ± 10.47	60.7 ± 2.24
	non-COVID	60.8 ± 3.66	42.6 ± 14.95	67.4 ± 2.78	
Xception	COVID-19	64.1 ± 3.60	65.3 ± 5.34	66.9 ± 4.02	67.6 ± 3.70
	non-COVID	71.4 ± 4.03	70.1 ± 5.80	68.1 ± 4.03	
DTL-V19	COVID-19	79.5 ± 4.63	82.6 ± 5.97	76.9 ± 3.96	79.1 ± 2.94
	non-COVID	79.4 ± 4.23	75.1 ± 7.74	80.8 ± 2.80	
	COVID-19	83.6 ± 1.34	85.4 ± 1.72	84.4 ± 1.64	85.3 ± 1.37
W-COVNet (this work)	COVID-19	83.6 ± 1.34	85.4 ± 1.72	84.4 ± 1.64	85.3 ± 1.37
	non-COVID	87.0 ± 2.19	85.3 ± 3.04	86.1 ± 1.21	

**Figure 10.** Training loss of our method with other studies.

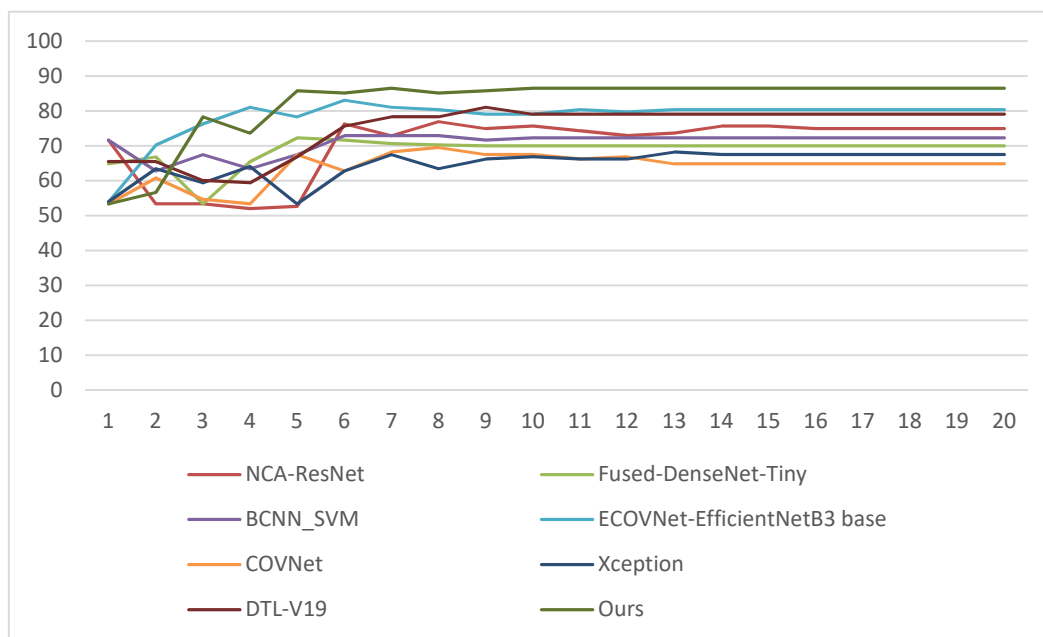


Figure 11. Test accuracy of our method with other studies.

5. Conclusions

Lung diseases are a major cause of death, and the emergence of COVID-19 in the last 3 years has taken a huge toll on humanity. The sudden outbreak has led to the failure of the medical system because of the high contagiousness of the disease. The doctors and equipment are not sufficient for rapid diagnosis and treatment, and hence artificial intelligence–assisted diagnosis is especially important. In this study, we proposed a weakly supervised COVID-19 classification network (W-COVNet). Given a large amount of redundant information in CT data, W-COVNet could perform feature selection to eliminate unnecessary information and retain only the core feature regions, avoiding interference in the deep learning training process. Meanwhile, bilinear fusion provided a stronger feature representation than a single linear model. Because of the poor visualization of deep learning, W-COVNet also used the currently popular Grad-CAM++ technique to visualize the deep learning process. We conducted extensive experiments to verify the performance of the network and achieved an average classification accuracy of 85.3% on the dataset with background removed, which was better than that of other COVID-19 classification networks and classical classification networks. W-COVNet could extract effective features to help physicians quickly diagnose and discover areas of lesions.

This study had certain limitations. Fewer data led to insufficient training. Moreover, the proposed W-COVNet had two disadvantages: 1) Since the segment was trained on the lung region of CT images, the network could not perform weakly supervised segmentation of CXR (chest X-ray) images. 2) Also, it could not judge the severity of the disease.

Acknowledgments

The data that support the findings of this study are openly at <https://github.com/UCSD-AI4H/COVID-CT> [39]. <http://medicalsegmentation.com/covid19/> [40].

This study was supported by the Jilin Provincial Key Laboratory of Intelligent Computing in Medical Image (No. YDZJ202102CXJD026), the Innovation Team Projects of Universities in Guangdong (No. 2022KCXTD057), the Higher Education Natural Science Foundation of Anhui Province (No. 2022AH051099), the Key Project of Science Research in Universities of Anhui Province of China (KJ2021A1066) and the open Foundation of Anhui Engineering Research Center of Intelligent Perception and Elderly Care, Chuzhou University (No. 2022OPA03).

Conflict of interest

The authors report no declarations of interest.

References

1. O. A. Ataguba, J. E. Ataguba, Social determinants of health: the role of effective communication in the COVID-19 pandemic in developing countries, *Global Health Action*, **1** (2020), 1788263. <https://doi.org/10.1080/16549716.2020.1788263>
2. World Health Organization, WHO coronavirus disease (COVID-19) Dashboard, 2022. Available from: https://covid19.who.int/?gclid=Cj0KCQjwZ7BRDzARIsAGjbK2ZXWRpJROEl97HGmSOx0_ydkVbc02Ka1FlcysGjEI7hnaIeR6xWhr4aAu57EALw_wcB.
3. Q. Huang, X. Huang, Z. Kong, X. Li, D. Tao, Bi-phase evolutionary searching for biclusters in gene expression data, *IEEE Trans. Evol. Comput.*, **5** (2018), 803–814. <https://doi.org/10.1109/TEVC.2018.2884521>
4. Q. Huang, J. Yao, J. Li, M. Li, M. R. Pickering, X. Li, Measurement of quasi-static 3-D knee joint movement based on the registration from CT to US, *IEEE Trans. Ultrason. Free*, **6** (2020), 1141–50. <https://doi.org/10.1109/TUFFC.2020.2965149>
5. J. Xi, Z. Miao, L. Liu, X. Yang, W. Zhang, Q. Huang, et al., Knowledge tensor embedding framework with association enhancement for breast ultrasound diagnosis of limited labeled samples, *Neurocomputing*, **468** (2022), 60–70. <https://doi.org/10.1016/j.neucom.2021.10.013>
6. Q. Huang, F. Pan, W. Li, F. Yuan, H. Hu, J. Huang, et al., Differential diagnosis of atypical hepatocellular carcinoma in contrast-enhanced ultrasound using spatio-temporal diagnostic semantics, *IEEE J. Biomed. Health*, **10** (2020), 2860–2869. <https://doi.org/10.1109/JBHI.2020.2977937>
7. J. Xi, D. Wang, X. Yang, W. Zhang, Q. Huang, Cancer omic data based explainable AI drug recommendation inference: A traceability perspective for explainability, *Biomed. Signal Process.*, **79** (2023), 104144. <https://doi.org/10.1016/j.bspc.2022.104144>
8. W. Shi, W. N. Chen, S. Kwong, J. Zhang, H. Wang, T. Gu, et al., A coevolutionary estimation of distribution algorithm for group insurance portfolio, *IEEE Trans. Syst. Man CY-S*, **11** (2021), 6714–28. <https://doi.org/10.1109/TSMC.2021.3096013>
9. Y. Yuan, M. Chao, Y.C. Lo, Automatic skin lesion segmentation using deep fully convolutional networks with jaccard distance, *IEEE Trans. Med. Imaging*, **9** (2017), 1876–1886. <https://doi.org/10.1109/TMI.2017.2695227>
10. P. Liskowski, K. Krawiec, Segmenting retinal blood vessels with deep neural networks, *IEEE Trans. Med. Imaging*, **11** (2016), 2369–2380. <https://doi.org/10.1109/TMI.2016.2546227>

11. H. Fu, J. Cheng, Y. Xu, D. W. K. Wong, J. Liu, X. Cao, Joint optic disc and cup segmentation based on multi-label deep network and polar transformation, *IEEE Trans. Med. Imaging*, **7** (2018), 1597–1605. <https://doi.org/10.1109/TMI.2018.2791488>
12. H. Fu, Y. Xu, S. Lin, X. Zhang, D. W. K. Wong, J. Liu, et al., Segmentation and quantification for angle-closure glaucoma assessment in anterior segment OCT, *IEEE Trans. Med. Imaging*, **9** (2017), 1930–1938. <https://doi.org/10.1109/TMI.2017.2703147>
13. M. Anthimopoulos, S. Christodoulidis, L. Ebner, A. Christe, S. Mougiakakou, Lung pattern classification for interstitial lung diseases using a deep convolutional neural network, *IEEE Trans. Med. Imaging*, **5** (2016), 1207–1216. <https://doi.org/10.1109/TMI.2016.2535865>
14. J. M. Wolterink, T. Leiner, M. A. Viergever, I. Išgum, Generative adversarial networks for noise reduction in low-dose CT, *IEEE Trans. Med. Imaging*, **12** (2017), 2536–2345. <https://doi.org/10.1109/TMI.2017.2708987>
15. S. Albarqouni, C. Baur, F. Achilles, V. Belagiannis, S. Demirci, N. Navab, Aggnet: Deep learning from crowds for mitosis detection in breast cancer histology images, *IEEE Trans. Med. Imaging*, **5** (2016), 1313–1321. <https://doi.org/10.1109/TMI.2016.2528120>
16. X. Zhang, G. Wang, S. G. Zhao, CapsNet-COVID19: Lung CT image classification method based on CapsNet model, *Math. Biosci. Eng.*, **19** (2022), 5055–5074. <https://doi.org/10.3934/mbe.2022236>
17. M. J. Horry, S. Chakraborty, B. Pradhan, M. Fallahpoor, H. Chegeni, M. Paul, Factors determining generalization in deep learning models for scoring COVID-CT images, *Math. Biosci. Eng.*, **18** (2021), 9264–9293. <https://doi.org/10.3934/mbe.2021456>
18. A. Singh, K. K. Singh, M. Greguš, I. Izonin, CNGOD-An improved convolution neural network with grasshopper optimization for detection of COVID-19, *Math. Biosci. Eng.*, **19** (2022), 12518–12531. <https://doi.org/10.3934/mbe.2022584>
19. S. H. Wang, S. C. Satapathy, M. X. Xie, Y. D. Zhang, ELUCNN for explainable COVID-19 diagnosis, *Soft Comput.*, **2023** (2023), 1–17. <https://doi.org/10.1007/s00500-023-07813-w>
20. Y. Zhang, M. A. Khan, Z. Zhu, S. Wang, SNELM: SqueezeNet-guided ELM for COVID-19 recognition, *Comput. Syst. Sci. Eng.*, **1** (2023), 13–26. <https://doi.org/10.32604/csse.2023.034172>
21. Y. Zhang, X. Zhang, W. Zhu, ANC: Attention network for COVID-19 explainable diagnosis based on convolutional block attention module, *CMES COMP Model. Eng.*, **3** (2021), 1037–1058. <https://doi.org/10.32604/cmes.2021.015807>
22. C. C. Lai, T. P. Shih, W. C. Ko, H. J. Tang, P. R. Hsueh, Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): The epidemic and the challenges, *Int. J. Antimicrob. Ag.*, **3** (2020), 105924. <https://doi.org/10.1016/j.ijantimicag.2020.105924>
23. E. Irmak, Implementation of convolutional neural network approach for COVID-19 disease detection, *Physiol. Genomics*, **12** (2020), 590–601. <https://doi.org/10.1152/physiolgenomics.00084.2020>
24. F. Xie, J. Xi, Q. Duan, Driver attribute filling for genes in interaction network via modularity subspace-based concept learning from small samples, *Complexity*, **2020** (2020), 1–12. <https://doi.org/10.1155/2020/6643551>
25. Q. Huang, F. Zhang, X. Li, A new breast tumor ultrasonography CAD system based on decision tree and BI-RADS features, *World Wide Web*, **21** (2018), 1491–1504. <https://doi.org/10.1007/s11280-017-0522-5>

26. G. Dong, Z. C. Zhang, J. Feng, X. M. Zhao, MorbidGCN: Prediction of multimorbidity with a graph convolutional network based on integration of population phenotypes and disease network, *Brief Bioinform.*, **4** (2022). <https://doi.org/10.1093/bib/bbac255>
27. S. Wang, P. Li, P. Chen, P. Phillips, G. Liu, S. Du, et al., Pathological brain detection via wavelet packet tsallis entropy and real-coded biogeography-based optimization, *Fund. Inform.*, **4** (2017), 275–291. <https://doi.org/10.3233/FI-2017-1492>
28. L. V. Fulton, D. Dolezel, J. Harrop, Y. Yan, C. P. Fulton, Classification of Alzheimer’s disease with and without imagery using gradient boosted machines and ResNet-50, *Brain Sci.*, **9** (2019), 212. <https://doi.org/10.3390/brainsci9090212>
29. Q. Guan, Y. Huang, Z. Zhong, Z. Zheng, L. Zheng, Y. Yang, Diagnose like a radiologist: Attention guided convolutional neural network for thorax disease classification, preprint, arXiv:1801.09927.
30. P. P. Ypsilantis, G. Montana, Learning what to look in chest X-rays with a recurrent visual attention model, preprint, arXiv:1701.06452.
31. E. Pesce, S. J. Withey, P. P. Ypsilantis, R. Bakewell, V. Goh, G. Montana, Learning to detect chest radiographs containing pulmonary lesions using visual attention networks, *Med. Image Anal.*, **53** (2019), 26–38. <https://doi.org/10.1016/j.media.2018.12.007>
32. M. Toğaçar, B. Ergen, Z. Cömert, COVID-19 detection using deep learning models to exploit Social Mimic Optimization and structured chest X-ray images using fuzzy color and stacking approaches, *Comput. Biol. Med.*, **121** (2020), 103805. <https://doi.org/10.1016/j.compbiomed.2020.103805>
33. J. P. Cohen, L. Dao, K. Roth, P. Morrison, Y. Bengio, A. F. Abbasi, et al., Predicting covid-19 pneumonia severity on chest x-ray with deep learning, *Cureus J. Med. Sci.*, **12** (2020), e9448. <https://doi.org/10.7759/cureus.9448>
34. Q. Ni, Z. Y. Sun, L. Qi, W. Chen, Y. Yang, L. Wang, et al., A deep learning approach to characterize 2019 coronavirus disease (COVID-19) pneumonia in chest CT images, *Eur. Radiol.*, **30** (2020), 6517–6527. <https://doi.org/10.1007/s00330-020-07044-9>
35. H. Ko, H. Chung, W. S. Kang, K. W. Kim, Y. Shin, S. J. Kang, et al., COVID-19 pneumonia diagnosis using a simple 2D deep learning framework with a single chest CT image: model development and validation, *J. Med. Int. Res.*, **6** (2020), e19569. <https://doi.org/10.2196/19569>
36. X. Wang, X. Deng, Q. Fu, Q. Zhou, J. Feng, H. Ma, et al., A weakly-supervised framework for COVID-19 classification and lesion localization from chest CT, *IEEE Trans. Med. Imaging*, **8** (2020), 2615–2625. <https://doi.org/10.1109/TMI.2020.2995965>
37. A. I. Khan, J. L. Shah, M. M. Bhat, CoroNet: A deep neural network for detection and diagnosis of COVID-19 from chest x-ray images, *Comput. Method Prog. Bio.*, **196** (2020), 105581. <https://doi.org/10.1016/j.cmpb.2020.105581>
38. L. Hussain, T. Nguyen, H. Li, A. A. Abbasi, K. J. Lone, Z. Zhao, et al., Machine-learning classification of texture features of portable chest X-ray accurately classifies COVID-19 lung infection, *Biomed. Eng. Online*, **19** (2020), 1–18. <https://doi.org/10.1186/s12938-020-00831-x>
39. J. Zhao, Y. Zhang, X. He, P. Xie, Covid-ct-dataset: a ct scan dataset about covid-19, 2020.
40. COVID-19 CT segmentation dataset. Available from: <https://medicalsegmentation.com/covid19/>.
41. D. P. Fan, T. Zhou, G. P. Ji, Y. Zhou, G. Chen, H. Fu, et al., Inf-net: Automatic covid-19 lung infection segmentation from ct images, *IEEE Trans. Med. Imaging*, **8** (2020), 2626–2637. <https://doi.org/10.1109/TMI.2020.2996645>

42. S. H. Wang, V. V. Govindaraj, J. M. Górriz, X. Zhang, Y. D. Zhang, Covid-19 classification by FGCNet with deep feature fusion from graph convolutional network and convolutional neural network, *Inform. Fusion*, **67** (2021), 208–229. <https://doi.org/10.1016/j.inffus.2020.10.004>
43. R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, D. Batra, et al., Grad-cam: Visual explanations from deep networks via gradient-based localization, *Int. J. Comput. Vision*, **2** (2020), 336–359. <https://doi.org/10.1007/s11263-019-01228-7>
44. A. Chattopadhyay, A. Sarkar, P. Howlader, V. N. Balasubramanian, Grad-cam++: Generalized gradient-based visual explanations for deep convolutional networks, in *2018 IEEE winter conference on applications of computer vision (WACV)*, (2018), 839–847. <https://doi.org/10.1109/WACV.2018.00097>
45. D. Chicco, N. Tötsch, G. Jurman, The Matthews correlation coefficient (MCC) is more reliable than balanced accuracy, bookmaker informedness, and markedness in two-class confusion matrix evaluation, *Biodata Min.*, **1** (2021), 1–22. <https://doi.org/10.1186/s13040-021-00244-z>
46. Q. Huang, Y. Lei, W. Xing, C. He, G. Wei, Z. Miao, et al., Evaluation of pulmonary edema using ultrasound imaging in patients with COVID-19 pneumonia based on a non-local Channel attention ResNet, *Ultrasound Med. Biol.*, **5** (2022), 945–953. <https://doi.org/10.1016/j.ultrasmedbio.2022.01.023>
47. F. J. P. Montalbo, Diagnosing Covid-19 chest x-rays with a lightweight truncated DenseNet with partial layer freezing and feature fusion, *Biomed. Signal Process.*, **68** (2021), 102583. <https://doi.org/10.1016/j.bspc.2021.102583>
48. R. Mastouri, N. Khelifa, H. Neji, S. Hantous-Zannad, A bilinear convolutional neural network for lung nodules classification on CT images, *Int. J. Comput. Ass Rad*, **16** (2021), 91–101. <https://doi.org/10.1007/s11548-020-02283-z>
49. A. Garg, S. Salehi, M. La Rocca, R. Garner, D. Duncan, Efficient and visualizable convolutional neural networks for COVID-19 classification using Chest CT, *Expert Syst. Appl.*, **195** (2022), 116540. <https://doi.org/10.1016/j.eswa.2022.116540>
50. F. Chollet, Xception: Deep learning with depthwise separable convolutions, in *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*, (2017), 1251–1258. <https://doi.org/10.48550/arXiv.1610.02357>
51. H. Panwar, P. Gupta, M. K. Siddiqui, R. Morales-Menendez, P. Bhardwaj, V. Singh, A deep learning and grad-CAM based color visualization approach for fast detection of COVID-19 cases using chest X-ray and CT-Scan images, *Chaos Soliton Fract.*, **140** (2020), 110190. <https://doi.org/10.1016/j.chaos.2020.110190>



AIMS Press

©2023 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>).