Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Computers in Biology and Medicine 137 (2021) 104834

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

ULNet for the detection of coronavirus (COVID-19) from chest X-ray images

Tianbo Wu [a], Chen Tang [a,*], Min Xu [a], Nian Hong [a], Zhenkun Lei [b]

[a] School of Electrical and Information Engineering, Tianjin University, Tianjin, 300072, China
[b] State Key Laboratory Analysis for Industrial Equipment, Dalian University of Technology, Dalian, 116024, China

ARTICLE INFO

Keywords:
Deep learning
ULNet
COVID-19
Classification
Chest X-ray images

ABSTRACT

Novel coronavirus disease 2019 (COVID-19) is an infectious disease that spreads very rapidly and threatens the health of billions of people worldwide. With the number of cases increasing rapidly, most countries are facing the problem of a shortage of testing kits and resources, and it is necessary to use other diagnostic methods as an alternative to these test kits. In this paper, we propose a convolutional neural network (CNN) model (ULNet) to detect COVID-19 using chest X-ray images. The proposed architecture is constructed by adding a new down-sampling side, skip connections and fully connected layers on the basis of U-net. Because the shape of the network is similar to UL, it is named ULNet. This model is trained and tested on a publicly available Kaggle dataset (consisting of a combination of 219 COVID-19, 1314 normal and 1345 viral pneumonia chest X-ray images), including binary classification (COVID-19 vs. Normal) and multiclass classification (COVID-19 vs. Normal vs. Viral Pneumonia). The accuracy of the proposed model in the detection of COVID-19 in the binary-class and multiclass tasks is 99.53% and 95.35%, respectively. Based on these promising results, this method is expected to help doctors diagnose and detect COVID-19. Overall, our ULNet provides a quick method for identifying patients with COVID-19, which is conducive to the control of the COVID-19 pandemic.

1. Introduction

COVID-19 is an acute pandemic and has spread around the world. At present, there are 223 countries and regions in the world with the COVID-19 pandemic, and more than 202 million people have been infected with COVID-19, which has caused more than 4 million deaths [1]. Once infected, patients with COVID-19 may experience various symptoms and signs of infection, including fever, cough, and respiratory diseases (such as flu). In severe cases, the infection can lead to pneumonia, breathing difficulties, multiple organ failure and death [2,3]. Therefore, it seriously affects people's production and life. In response to the pandemic, many countries have declared complete blockades and required their people to stay indoors and strictly avoid gatherings. At the same time, the research and development departments of different research centers are seeking possible solutions in the fields of medicine, biology, data science, and deep learning to prevent and control this pandemic [4–6].

A key and important step in combating COVID-19 is to effectively screen infected patients to isolate and treat those who are positive. At present, the main screening method used to detect COVID-19 is real-time reverse transcription polymerase chain reaction (RT–PCR) [7,8]. RT–PCR detection is the gold standard due to its high specificity. However, some viruses have mutated with the rapid development of the pandemic, and RT–PCR is not sensitive to some mutated viruses. In addition, RT–PCR is a very time-consuming, laborious and complex manual process, and it is in short supply in some countries. Another method can be based on chest radiographic images. Various studies published in the Journal of Radiology [9,10] indicate that chest scans may help to detect COVID-19. Researchers have found that the lungs of patients with COVID-19 symptoms have some fuzzy marks, such as ground glass-fuzzy dark spots, which can be used to distinguish COVID-19-infected patients from non-COVID-19-infected patients [10,11]. In addition, chest X-ray images have been widely used to detect COVID-19 [12,13]. Many deep learning-based systems have been proposed and have shown promising results in terms of the accuracy of using chest images to detect patients infected with COVID-19. In the method based on deep learning, the network is first trained on the chest images in the training set. After training, the chest images to be checked are simultaneously input into the trained network, and the corresponding classification results can be obtained easily. Compared with RT–PCR, these deep learning-based methods are faster, easier, and labor-saving. Moreover, these deep learning-based methods are more

* Corresponding author.
E-mail address: tangchen@tju.edu.cn (C. Tang).

https://doi.org/10.1016/j.compbiomed.2021.104834
Received 10 May 2021; Received in revised form 27 August 2021; Accepted 31 August 2021
Available online 4 September 2021

Table 1
Some published deep learning methods for detecting COVID-19 based on chest images.

Study	Dataset	Models	Performance
Apostolopoulos et al. [15]	1428 chest X-ray images including 224 images with confirmed COVID-19 cases, 700 images withconfirmed bacterial pneumonia cases and 504 images of normal cases.	VGG19	Accuracy of 98.75% for 2-classes and 93.48% for 3-classes
Hemdan et al. [16]	50 chest X-ray images including 25 cases with COVID-19 and 25 cases without any infections.	COVIDX-Net	Accuracy of 90% for 2-classes and F1-scores of 0.91 for COVID-19 and 0.89 for normal
Qzturk et al. [17]	1125 chest X-ray images comprising 125 with COVID-19, 500 normal and 500 pneumonia cases.	DarkCovidNet	Accuracy of 87.02% for 3-classes
Khan et al. [18]	3084 chest X-ray images and CT images comprising 290 COVID-19, 1203 normal, 931 viral pneumonia and 660 bacterial pneumonia.	CoroNet	Accuracy of 95% for 3-classes and 89.6% for 4-classes
Sethy et al. [19]	50 chest X-ray images comprising 25 cases with COVID-19 and 25 cases with without any infections.	Deep features from ResNet50 + SVM classifier	Accuracy of 95.38%
Suat et al. [20]	2331 chest X-ray images including 231 COVID-19 cases, 1050 pneumonia cases and 1050 no-findings cases.	Convolutional CapsNet	Accuracy of 97.24% for 2-classes and 84.22% for 3-classes
J. Zhang et al. [21]	1078 chest X-ray images comprising 1008 non-COVID-19 pneumonia and 70 COVID-19.	Deep CNN based on Backbone network	Sensitivity of 96.0% and specificity of 70.7% along with an AUC of 95.2%
B. Ghoshal et al. [22]	5941 chest X-ray images including 2786 bacterial pneumonia, 1583 normal, 68 COVID-19 and 1504 non-COVID-19 viral pneumonia.	Dropweights based Bayesian Convolutional Neural Networks	Accuracy of 92.90%
Harsh Panwar et al. [23]	284 chest X-ray images including 142 cases with COVID-19 and 142 cases of normal	nCOVnet	Accuracy of 88% for 2-classes and the accuracy rate of positive COVID-19 is 97%
Tawsifur Rahman et al. [24]	18479 chest X-ray images including 3616 COVID-19, 8851 Normal and 6012 Non-COVID.	DenseNet201	Accuracy of 95.11% for 3-classes
S.Wang et al. [25]	99 chest CT images including 55 viral pneumonia and 44 COVID-19.	M-Inception	Accuracy of 73.1%, along with a sensitivity of 74.0% and a specificity of 67.0%
Zheng et al. [26]	542 chest CT images including 313 COVID-19 and 229 no-findings.	DeCovNet	Accuracy of 90.1%
L. Li et al. [27]	4356 chest CT images including 1735 pneumonia, 1325 nonpneumonia and 1296 COVID-19.	COVNet	Specificity of 96%, sensitivity of 90%, and AUC of 0.96
Y. Song et al. [28]	275 chest CT images including 88 COVID-19, 101 bacteria pneumonia and 86 healthy.	DeepPneumonia	Accuracy of 94.0% for 2-classes and 86.0% for 3-classes

intelligent because they automatically learn features from the data itself rather than based on the manual extraction of features [14]. Therefore, researchers believe that a system based on chest radiology could be an effective tool for detecting, quantifying and tracking COVID-19 cases.

In this paper, we first review some advanced networks and propose a ULNet model that is constructed by downsampling, upsampling, skip connections and fully connected layers. The ULNet is an effective tool to help doctors quickly identify patients with COVID-19 by using chest X-ray images and thus delay the rapid spread of COVID-19. At the same time, the proposed model is trained on 2905 chest X-ray images, which are open source and can be used publicly to classify different types of chest X-ray images (COVID-19 vs. viral pneumonia vs. normal). In addition, we evaluated the performance of the ULNet model by precision, recall, F1-score, accuracy and receiver operating characteristic (ROC) curve and compared it with other published studies.

The contributions of the paper are summarized as follows:

- We build a new deep learning model (ULNet) and apply it to two classification and three classification tasks. Multiple experiments show that the use of ULNet can improve the classification accuracy.
- We evaluate the performance of our model through different evaluation metrics and an external dataset.
- The proposed ULNet model helps researchers continue to develop advanced deep learning technologies for computer-aided design systems to respond to the COVID-19 pandemic.

The remainder of the paper is organized as follows. In Section 2, some COVID-19 diagnosis systems based on deep convolution neural networks in recent years are reviewed. In Section 3, the proposed ULNet model is introduced in detail. Section 4 gives the dataset, the experimental setup and the performance of our model. Section 5 describes the general contributions of this paper and discusses some future directions.

2. Related work

Recently, a number of scientists have proposed a series of systems for diagnosing COVID-19 from chest X-ray images based on deep learning models, and we reviewed some of the advanced models. Apostolopoulos et al. [15] developed a diagnostic method for COVID-19 based on deep learning. In this study, both two-class and three-class analyses were carried out simultaneously. The accuracy of the proposed model was 98.75% for the two-class (COVID-19 vs. no findings) and 93.48% for the three-class (COVID-19 vs. no findings vs. pneumonia). Hemdan et al. [16] provided a COVIDX-Net model comprising seven different deep architectures (MobileNetV2, VGG19, InceptionV3, DenseNet201, ResNetV2, and Xception as well as InceptionResNetV2) to detect COVID-19 from chest X-ray images. In this study, the average accuracy

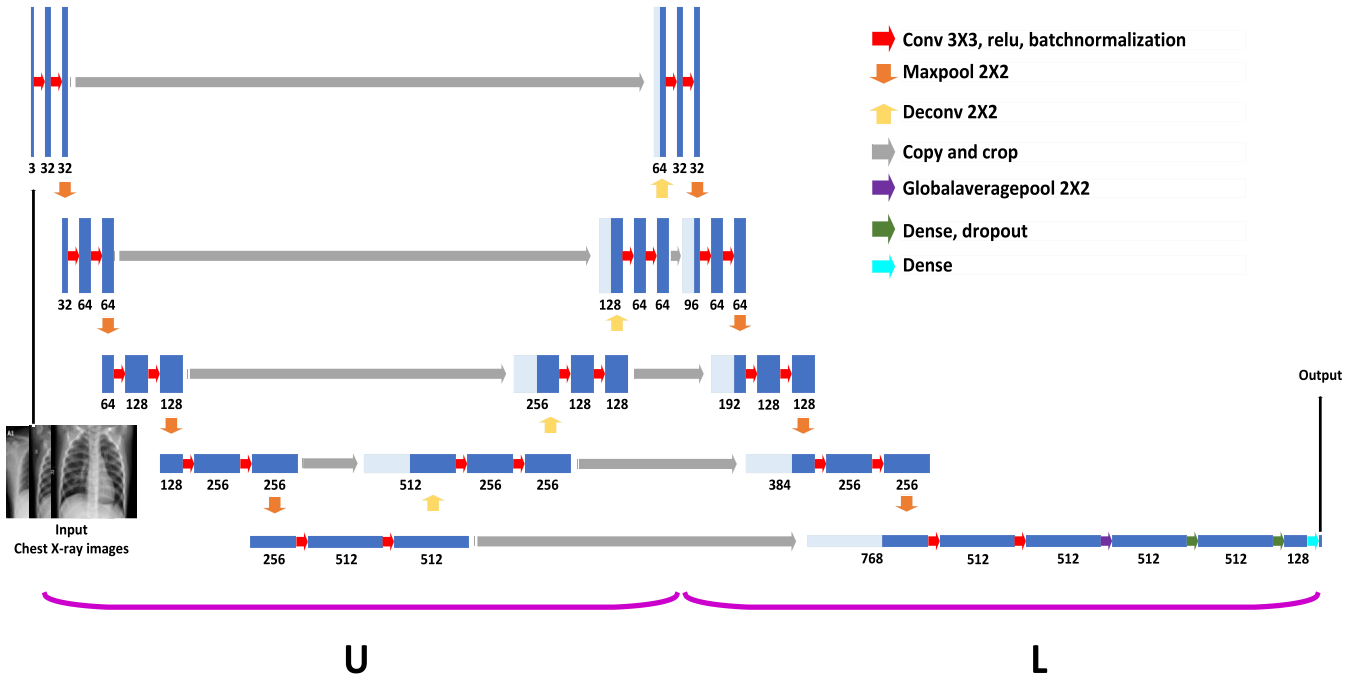

Fig. 1. The structure of the ULNet model.

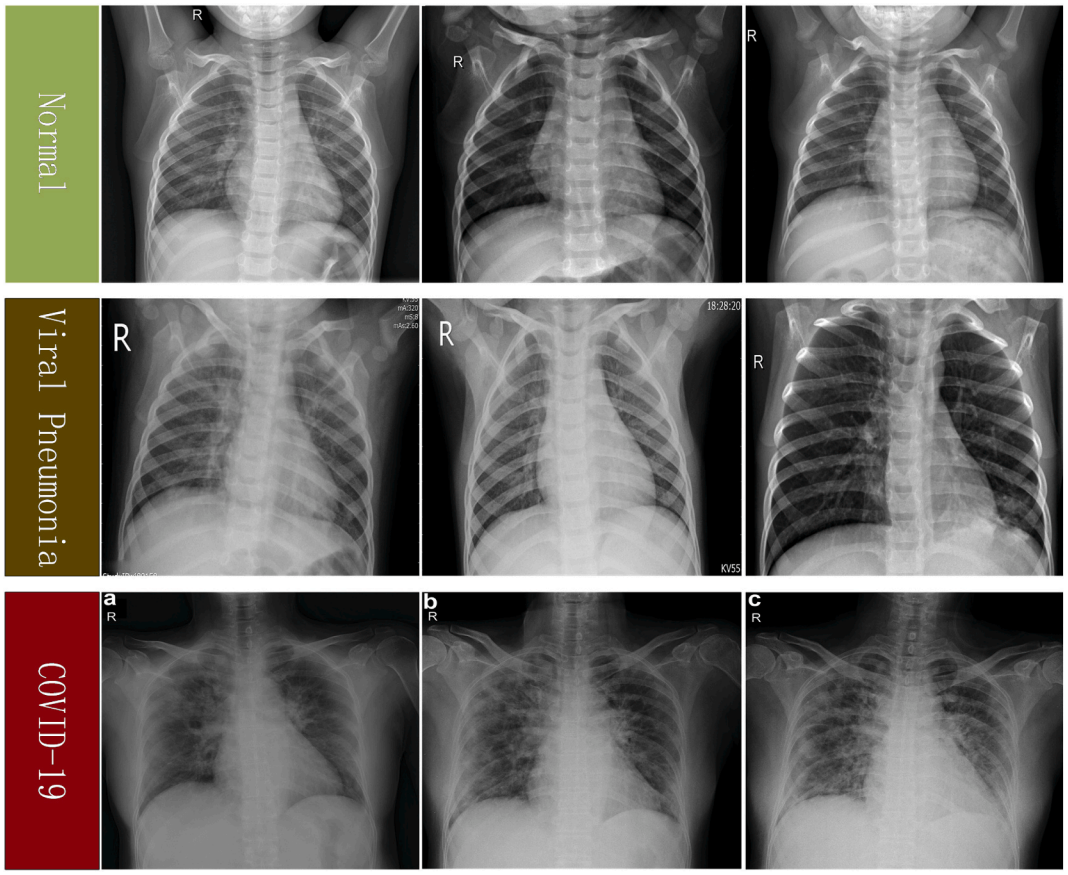

Fig. 2. Example of chest X-ray images in Kaggle's COVID-19 Radiography Dataset: (first row) Normal chest X-ray images, (second row) Viral Pneumonia chest X-ray images, (third row) COVID-19 chest X-ray images.

rate for the binary class problem was 90%, and the F1-scores for COVID-19 and normal conditions were 0.91 and 0.89, respectively. Qzturk et al. [17] introduced a deep CNN model (DarkCovid-Net) based on the DarkNet model. This model was trained on chest X-ray images and was designed to provide accurate diagnosis for two-class classification (COVID-19 vs. Normal) and three-class classification (COVID-19

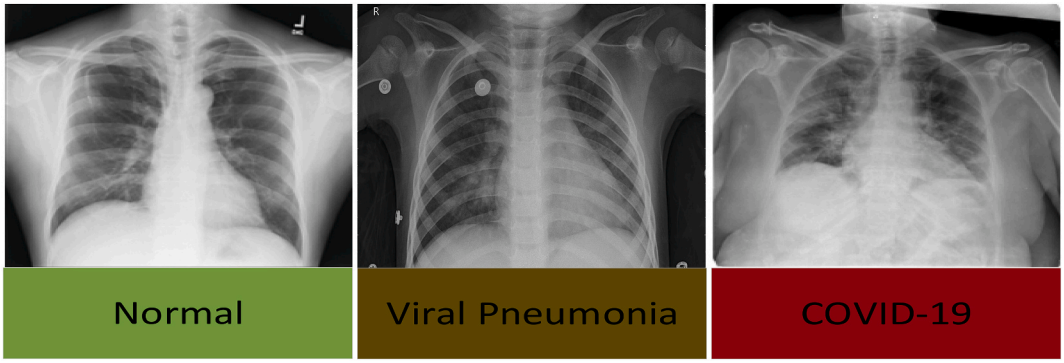

Fig. 3. Example of chest X-ray images in the QaTa-COV19 dataset: (first column) Normal chest X-ray images, (second column) Viral Pneumonia chest X-ray images, (third column) COVID-19 chest X-ray images.

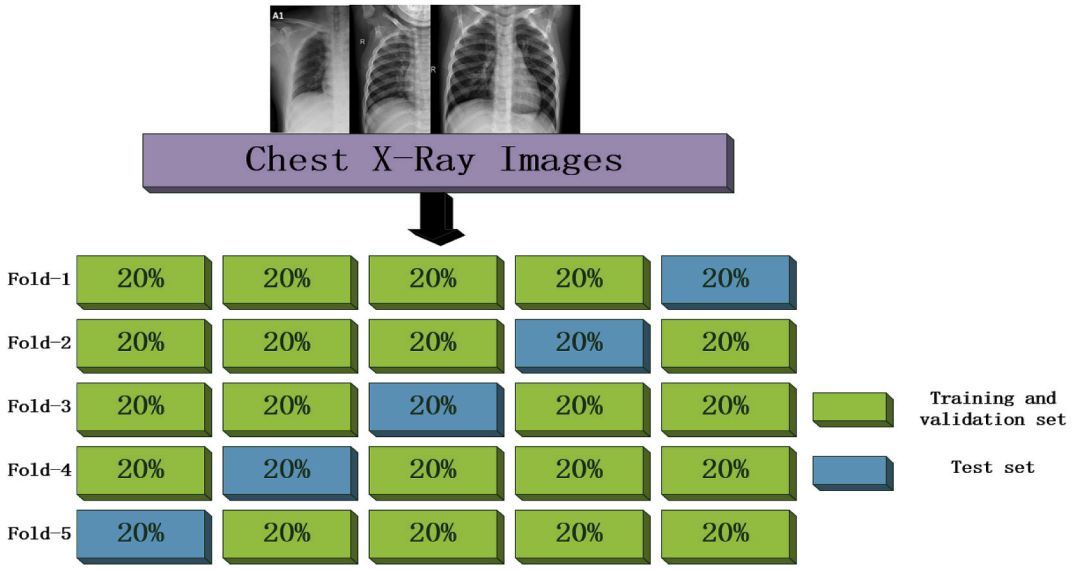

Fig. 4. Schematic diagram of Kaggle's COVID-19 Radiography Dataset used in the 5-fold cross-validation.

Table 2
Distribution of Kaggle's COVID-19 Radiography Dataset for binary classifications.

Class	Number of images	Training	Validation	Test
COVID-19	219	158	18	43
Normal	1341	965	107	269

Table 3
Distribution of Kaggle's COVID-19 Radiography Dataset for three classifications.

Class	Number of images	Training	Validation	Test
COVID-19	219	158	18	43
Normal	1341	965	107	269
Viral pneumonia	1345	969	107	269

vs. Normal vs. Pneumonia). The accuracy was 98.08% for the two-class model and 87.02% for the three-class model. Khan et al. [18] proposed the CoroNet model, which is based on the Xception architecture and pretrained on ImageNet, to diagnose COVID-19 using chest X-ray images and computed tomography (CT) images. This model achieved an overall accuracy of 95% and 89.6% for three classes (viral pneumonia vs. COVID-19 vs. bacterial pneumonia vs. normal) and four classes (pneumonia vs. COVID-19 vs. normal). Sethy et al. [19] used various CNN models along with a support vector machine (SVM) classifier to detect COVID-19 from chest X-ray images. The experimental results showed that the ResNet50 model with the SVM classifier achieved the highest accuracy of 95.38%. Suat et al. [20] introduced a convolutional capsule network, an artificial neural network for detecting COVID-19 from chest X-ray images. The developed model had an accuracy rate of 97.24% for binary-class classification (COVID-19 vs. no findings) and 84.22% for multiclass classification (COVID-19 vs. no findings vs. pneumonia). Zhang et al. [21] presented a deep CNN model based on a backbone network. The model was trained on chest X-ray images and achieved a sensitivity of 96% and specificity of 70.7% along with an AUC of 95.2%. B. Ghoshal et al. [22] proposed a Bayesian deep learning classifier based on the transfer learning method to diagnose COVID-19 from chest X-ray images. The experimental results showed that Bayesian inference improves the detection accuracy of VGG16 from 85.7% to 92.9%. Harsh Panwar et al. [23] proposed the nCOVnet model to detect COVID-19 by using chest X-ray images. The results showed that the proposed model correctly detected patients with positive COVID-19 with an accuracy of 97%, and the overall accuracy of nCOVnet was 88%. Tawsifur Rahman et al. [24] introduced a DenseNet201 model that was trained on chest X-ray images from Kaggle's COVID-19 Radiography Database. This model achieved an overall accuracy of 95.11% for three classes (COVID-19 vs. Normal vs. Non-COVID).

In addition, many researchers have proposed deep learning systems for detecting COVID-19 from CT images. S. Wang et al. [25] introduced

Fig. 5. Confusion matrix results of our ULNet for 2-class classification. (a) Overlapping CM, (b) Fold-1 CM, (c) Fold-2 CM, (d) Fold-3 CM, (e) Fold-4 CM, (f) Fold-5 CM.

Table 4
The experimental results in fold 1 for 2 classes of the proposed ULNet.

Class	Precision	Recall	F1-score	Accuracy
COVID-19	100%	97.67%	98.82%	99.68%
Normal	99.63%	100%	99.81%	

Table 5
The experimental results in fold 2 for 2 classes of the proposed ULNet.

Class	Precision	Recall	F1-score	Accuracy
COVID-19	91.49%	100%	95.56%	98.72%
Normal	100%	98.51%	99.25%	

the M-Inception model, which was trained on chest CT images to classify viral pneumonia and COVID-19. The results of the test dataset showed that the overall accuracy was 73.1%, the sensitivity was 74.0%, and the specificity was 67.0%. Zheng et al. [26] proposed a new deep learning model (DeCovNet), which was trained on chest CT images and reached an accuracy rate of 90.1%. L. Li et al. [27] introduced the COVNet model

Table 6
The experimental results in fold 3 for 2 classes of the proposed ULNet.

Class	Precision	Recall	F1-score	Accuracy
COVID-19	91.49%	100%	95.56%	98.72%
Normal	100%	98.51%	99.25%	

Table 7
The experimental results in fold 4 for 2 classes of the proposed ULNet.

Class	Precision	Recall	F1-score	Accuracy
COVID-19	100%	100%	100%	100%
Normal	100%	100%	100%	

Table 8
The experimental results in fold 5 for 2 classes of the proposed ULNet.

Class	Precision	Recall	F1-score	Accuracy
COVID-19	100%	100%	100%	100%
Normal	100%	100%	100%	

Table 9
Performance of the proposed ULNet for 2 classes on Kaggle's COVID-19 Radiography Dataset.

Folds	Precision	Recall	F1-score	Accuracy
Fold-1	99.82%	98.84%	99.32%	99.68%
Fold-2	95.75%	99.26%	97.41%	98.72%
Fold-3	95.75%	99.26%	97.41%	98.72%
Fold-4	100%	100%	100%	100%
Fold-5	100%	100%	100%	100%
Average	98.26%	99.47%	98.83%	99.42%

Table 10
Performance of the proposed ULNet for 2 classes on the QaTa-COV19 dataset.

Folds	Precision	Recall	F1-score	Accuracy
Fold-1	98.51%	98.50%	98.50%	98.50%
Fold-2	99.51%	99.50%	99.50%	99.50%
Fold-3	99.51%	99.50%	99.50%	99.50%
Fold-4	99.02%	99.00%	99.00%	99.00%
Fold-5	100%	100%	100%	100%
Average	99.31%	99.30%	99.30%	99.30%

based on ResNet50 trained on chest CT images. The experimental results showed that the specificity of the COVNet model in the classification of COVID-19 was 96%, the sensitivity was 90%, and the AUC was 0.96. Y. Song et al. [28] developed a deep learning model (DeepPneumonia) to detect COVID-19 from chest CT images. The overall accuracy of the model for the classification of COVID-19 vs. bacterial pneumonia was 86.0%, and the overall accuracy for the classification of COVID-19 vs. healthy individuals was 94.0%. More details are shown in Table 1.

In this paper, we further improve the accuracy of classification. We propose that the ULNet compensates for the loss of information by adding new skip connections and adds a new downsampling side structure to extract deep-level features, which improves the classification performance of the network.

3. ULNet model

In this section, we propose a new deep learning network ULNet based on the U-net [29] architecture to detect COVID-19 from chest X-ray images. The structure of the network is shown in Fig. 1.

Since U-net [29] was proposed in 2015, it has been widely used in the field of medical image segmentation due to its excellent network structure, but no one has used it or its improved structure to deal with classification tasks. We propose ULNet for classification on the basis of U-net, which is composed of a U-shaped network on the left and an L-shaped network on the right, as shown in Fig. 1. The U-shaped network is basically the same as the U-net structure, including the downsampling structure, upsampling structure and skip connections, but we add a batch normalization layer after each convolutional layer in the downsampling structure and the upsampling structure. The batch normalization layer is very important because it normalizes the mean and variance of the output of each layer of the network, which can accelerate the convergence rate of the network, control the gradient explosion, prevent the gradient from disappearing, and suppress overfitting [30]. At the same time, the same downsampling structure as the U-shaped network is added behind it, and then two fully connected layers are added to form an L-shaped network. The downsampling structure of the L-shaped network increases the depth of the network to capture deep-level features. At the same time, we make skip connections between the upsampling structure of the U-shaped network and the downsampling structure of the L-shaped network. Different levels of features can be merged by these skip connections to avoid a large amount of information loss caused by the pooling operation. In addition, the last fully connected layer is forwarded to a sigmoid function in the two-classification task, and it is forwarded to a softmax function in the three-classification task. The above is the complete structure of the ULNet. In the following, we introduce in detail the main components of the proposed model, such as convolutional layers, activation units, batch normalization, pooling, deconvolution and fully connected layers.

3.1. Convolutional layer

Convolution is the basic operation of a convolutional neural network, and the convolutional layer is the base layer of the CNN. It is responsible for determining the features of the pattern. In this layer, the input image is passed through a filter, and the values resulting from filtering consist of the feature map. This layer applies some kernels that slide through the pattern to extract low-level and high-level features in the pattern. The kernel is a $n \times n$-shaped matrix to be transformed with the input pattern matrix. The stride parameter is the number of steps tuned for shifting over the input matrix. The output of the convolutional layer can be given as

$$x_j^l = f\left(\sum_{a=1}^{N} w_j^{l-1} * y_a^{l-1} + b_j^l\right) \tag{1}$$

where x_j^i is the j-th feature map in layer l, w_j^{l-1} indicates the j-th kernel in layer $l-1$, y_a^{l-1} represents the a-th feature map in layer $l-1$, b_j^l indicates the bias of the j-th feature map in layer l, N is the number of total features in layer $l-1$, and * represents the vector convolution process.

3.2. Activation function

The activation function is used after each convolution to introduce nonlinearity into the model. As an activation function, the ReLU function has become very prevalent since it was proposed due to its advantages in overcoming gradient disappearance and accelerating training speed. The formula of the ReLU function can be expressed as

$$\mathrm{Re}LU(x) = \begin{cases} 0, & x < 0 \\ x, & x \geq 0 \end{cases} \tag{2}$$

If the input is negative, the output is 0, and if the input is nonnegative, the output is x.

3.3. Batch normalization layer

Given a mini batch $B = \{x_1, x_2, \ldots, x_m\}$ of size m, the normalized values $(\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_m)$ and their linear transformations $(y_1, y_2, \ldots, y_m)$. Batch normalization $BN_{\gamma,\beta}$ refers to the transform $BN_{\gamma,\beta}: x_1, x_2, \ldots, x_m \rightarrow y_1, y_2, \ldots, y_m$ and is computed as:

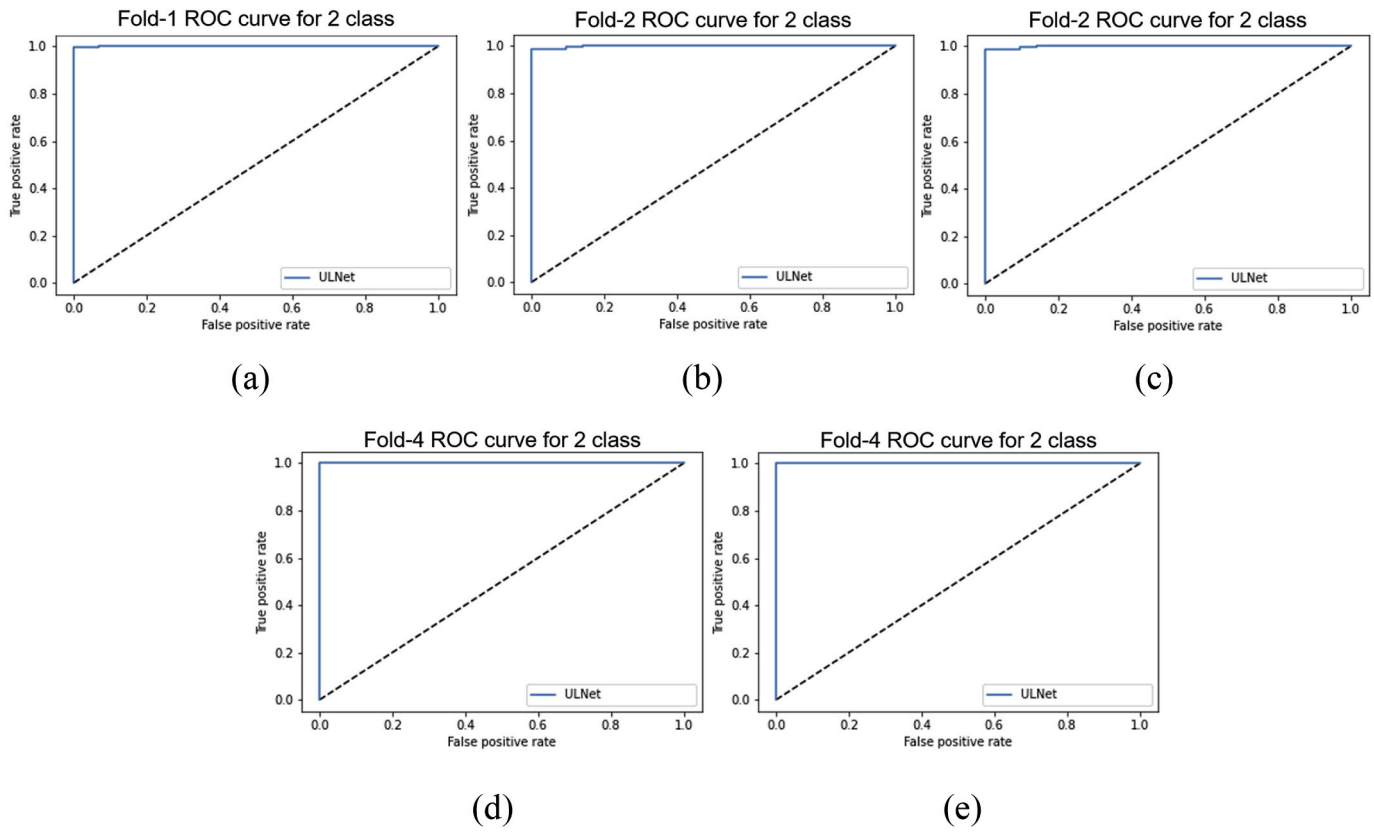

Fig. 6. Receiver operating curve of our ULNet for 2-class classification. (a) Fold-1 ROC, (b) Fold-2 ROC, (c) Fold-3 ROC, (d) Fold-4 ROC, (e) Fold-5 ROC.

$$\mu_B = \frac{1}{m}\sum_{i=1}^{m} x_i$$

$$\sigma_B^2 = \frac{1}{m}\sum_{i=1}^{m}(x_i - \mu_B)^2 \tag{3}$$

$$\widehat{x}_i = \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \varepsilon}}$$

$$y_i = \gamma \widehat{x}_i + \beta \equiv BN_{\gamma,\beta}(x_i)$$

where μ_B and σ_B^2 are the minibatch mean and variance, respectively, β and γ are parameters learnable via backpropagation, and ε is a small positive number to avoid division by zero [28].

3.4. Pooling layer

The pooling layer maintains the most relevant information while downsampling and compresses features to reduce the amount of calculation. In addition, the pooling layer can expand the receptive field. The pooling operation is a window with a size of $h_p \times h_p$ sliding on the feature map in steps. There are usually two methods: maximum pooling: returns the maximum value in each window; average pooling: returns the average value of each window.

3.5. Deconvolution layer

Deconvolution is a special forward convolution. It enlarges the size of the input image by adding 0 in a certain proportion, then rotates the convolution kernel and carries out forward convolution. When performing forward convolution, the relationship between the input and output dimensions is:

$$o = \frac{i + 2p - k}{s} + 1 \tag{4}$$

When performing deconvolution, the relationship between input and output sizes involves two situations. If $(o + 2p - k)\%s = 0$, at the time, the input and output size relationship of deconvolution is:

$$o = s(i-1) - 2p + k \tag{5}$$

If $(o + 2p - k)\%s \neq 0$, the input and output size relationship of deconvolution is:

$$o = s(i-1) - 2p + k + (o + 2p - k)\%s \tag{6}$$

where o is the size of the output, i is the size of the input, p is the size of the padding, k is the size of the convolution kernel, and s is the step size.

3.6. Fully connected layer

Each node in the fully connected layer is connected to all of the nodes of the previous layer and is used to integrate the features extracted from the front. In the CNN network, the fully connected layer maps the feature map generated by the convolution layer into a feature vector of fixed length, which is generally the number of image categories in the input image dataset. This feature vector contains the combined information of all features of the input image. Although the position information of the image is lost, the vector retains the most characteristic image features in the image to complete the image classification task. From the perspective of the image classification task, the computer only needs to judge the content of the image, calculate the specific category value of the input image (the category probability), and output the most likely category to complete the classification task.

In addition, the optimization function used in our ULNet is Adam, which is one of the most commonly used and effective methods in gradient descent optimization. At the same time, a dropout layer is

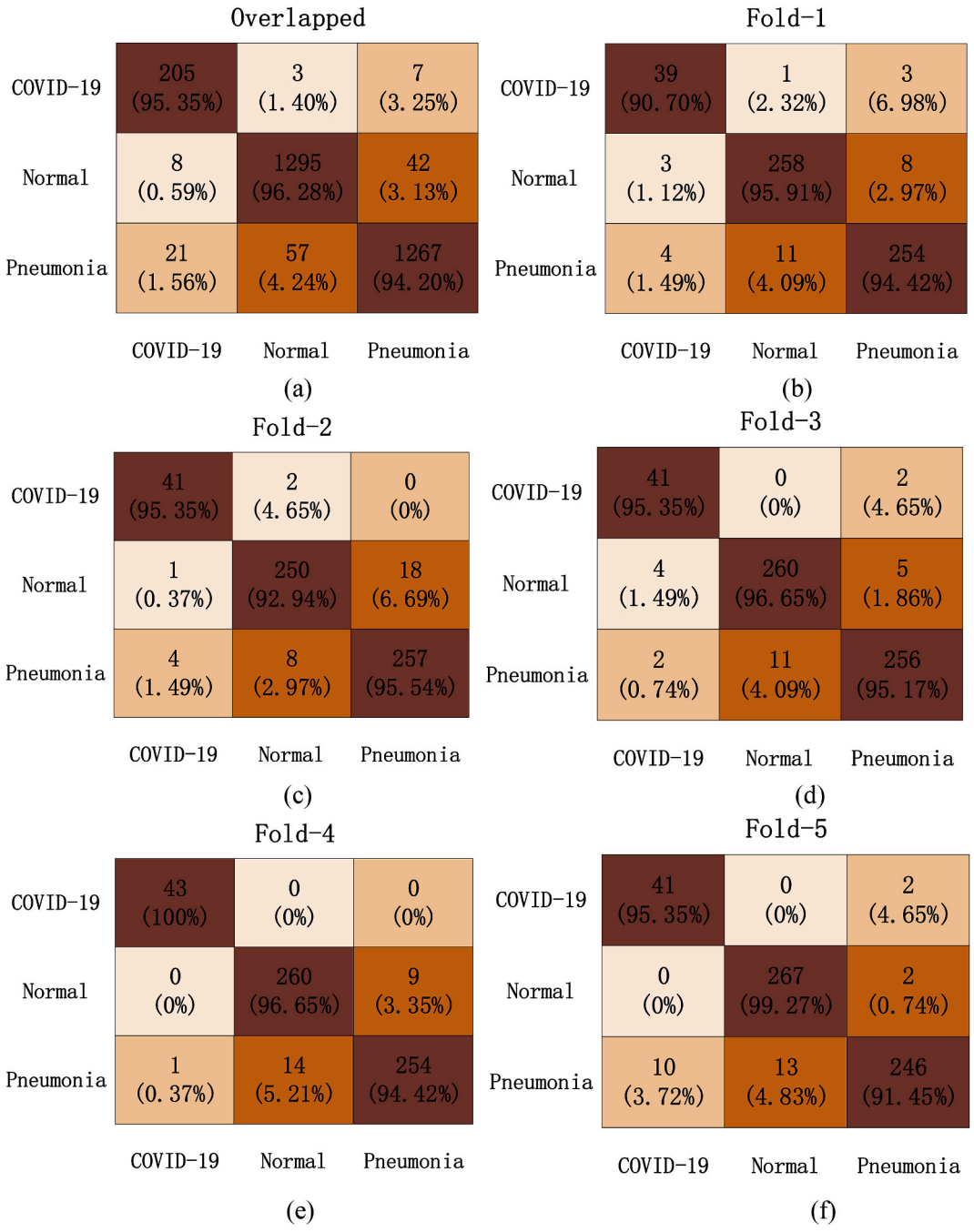

Fig. 7. Confusion matrix results of our ULNet for 3-class classification. (a) Overlapping CM, (b) Fold-1 CM, (c) Fold-2 CM, (d) Fold-3 CM, (e) Fold-4 CM, (f) Fold-5 CM.

Table 11
The experimental results in fold 1 for 3 classes of the proposed ULNet.

Class	Precision	Recall	F1-score	Accuracy
COVID-19	84.78%	90.70%	87.64%	94.84%
Normal	95.56%	95.91%	95.75%	
Pneumonia	95.85%	94.42%	93.65%	

Table 12
The experimental results in fold 2 for 3 classes of the proposed ULNet.

Class	Precision	Recall	F1-score	Accuracy
COVID-19	89.13%	95.35%	92.14%	94.32%
Normal	96.15%	92.94%	94.52%	
Pneumonia	93.45%	95.54%	94.48%	

added after each fully connected layer to prevent overfitting. In addition, the receptive field will be too large if a large filter is used so that most of the extracted information is irrelevant and the details will be lost, resulting in poor classification accuracy. Therefore, we use the small filter of size 3×3.

4. Experiment

In this section, we introduce the dataset used in this paper and the distribution of images used for training, verification and testing in detail. At the same time, we evaluate the performance of the proposed

Table 13
The experimental results in fold 3 for 3 classes of the proposed ULNet.

Class	Precision	Recall	F1-score	Accuracy
COVID-19	87.23%	95.35%	91.11%	95.87%
Normal	95.94%	96.65%	96.29%	
Pneumonia	97.34%	95.17%	96.24%	

Table 14
The experimental results in fold 4 for 3 classes of the proposed ULNet.

Class	Precision	Recall	F1-score	Accuracy
COVID-19	97.73%	100%	98.85%	95.87%
Normal	94.89%	96.65%	95.76%	
Pneumonia	96.58%	94.42%	95.49%	

Table 15
The experimental results in fold 5 for 3 classes of the proposed ULNet.

Class	Precision	Recall	F1-score	Accuracy
COVID-19	80.39%	95.35%	87.23%	95.35%
Normal	95.36%	99.26%	97.27%	
Pneumonia	98.40%	91.45%	94.80%	

Table 16
Performance of the proposed ULNet for 3 classes on Kaggle's COVID-19 Radiography Dataset.

Folds	Precision	Recall	F1-score	Accuracy
Fold-1	92.06%	93.68%	92.35%	94.84%
Fold-2	92.91%	94.61%	93.71%	94.32%
Fold-3	93.50%	95.72%	94.55%	95.87%
Fold-4	96.40%	97.02%	96.71%	95.87%
Fold-5	91.38%	95.35%	93.10%	95.35%
Average	93.25%	95.28%	94.09%	95.25%

Table 17
Performance of the proposed ULNet for 3 classes on the QaTa-COV19 dataset.

Folds	Precision	Recall	F1-score	Accuracy
Fold-1	95.29%	94.67%	94.73%	94.67%
Fold-2	94.82%	94.00%	93.96%	94.00%
Fold-3	99.02%	99.00%	98.99%	99.00%
Fold-4	97.18%	97.00%	96.99%	97.00%
Fold-5	98.36%	98.33%	98.33%	98.33%
Average	96.93%	96.60%	96.60%	96.60%

ULNet model by using different evaluation indicators. In addition, we use an external dataset (QaTa-COV19 Dataset) [31] to evaluate the model. The model is used for two classifications of normal people and COVID-19 patients, as well as three classifications of normal people, COVID-19 patients and viral pneumonia patients.

4.1. Dataset and experimental setup

In this paper, the chest X-ray image dataset we used is Kaggle's COVID-19 Radiography Dataset [32]. This dataset contains 2905 chest X-ray images of three different types, including 1341 normal images, 1345 viral pneumonia images and 219 COVID-19 images. To better observe the differences between different types of chest X-ray images, we show some images of Kaggle's COVID-19 Radiography Dataset [32] in Fig. 2. Kaggle's COVID-19 Radiography Dataset is an open dataset in which chest X-ray images come from three different datasets: chest X-ray images with COVID-19 are taken from the Italian Society of Medical and Interventional Radiology COVID-19 Dataset (SIRM) [33] and from the Novel Coronavirus 2019 Dataset, which was developed by Cohen et al. in GitHub [34], as well as from various recently published articles. Viral pneumonia and normal images were collected from Kaggle's Chest X-ray pneumonia dataset [35]. The metadata for this dataset can be found at https://www.kaggle.com/tawsifurrahman/covid19-radiography-database/version/1. Another dataset, the QaTa-COV19 dataset [31], was compiled by researchers from Qatar University and Tampere University. We randomly selected 300 chest X-ray images (including 100 COVID-19, 100 Normal and 100 Viral Pneumonia) to evaluate our model, and Fig. 3 shows some images of the QaTa-COV19 Dataset [31]. It is worth noting that these 300 images from the QaTa-COV19 Dataset [31] are not included in the COVID-19 Radiography Dataset [32].

For the setting of this experiment, all images are scaled to the size of 256×256 pixels, and the number of channels is 3. The ULNet architecture is implemented using Python and TensorFlow 2, which includes the Keras package. In addition, the experiment is performed on an Intel (R) Core (TM) i7–2.60 GHz processor and an NVIDIA GTX 1650 GPU with 4 GB of memory.

4.2. Training and evaluation standards

In this section, we conduct experiments to evaluate the performance of the ULNet model, which is used for the two classifications of COVID-19 and normal and the three classifications of COVID-19, viral pneumonia and normal. At the same time, a fivefold cross-validation method is used to evaluate the efficiency of the above classification problems. We divide Kaggle's COVID-19 Radiography Dataset into 5 independent and equal sets. Four of the five sets are used for training and verification, and the last set is used for testing. In other words, 70% of the dataset is used for training, 10% for verification and 20% for testing. The distribution diagram of the dataset is shown in Fig. 4. More specifically, for the binary classification problem, 219 COVID-19 chest X-ray images are used, including 158 images for training, 18 images for validation and 43 images for testing. Furthermore, 1341 normal chest X-ray images are used, including 965 images for training, 107 images for validation and 269 images for testing. The distribution of chest X-ray images for the two classifications is summarized in Table 2. For the three classification problems, the distribution of COVID-19 and normal images is the same as that of the binary classifications. In addition, 1345 viral pneumonia images are used, including 969 images for training, 107 images for validation and 269 images for testing. The distribution of chest X-ray images for the three classifications is summarized in Table 3. At the same time, for the two classifications problem, each fold of the trained model is tested on 200 chest X-ray images (including 100 COVID-19 and 100 Normal) on the QaTa-COV19 Dataset. For the three classification problems, each fold of the trained model is tested on 300 chest X-ray images (including 100 COVID-19, 100 normal and 100 viral pneumonia) on the QaTa-COV19 dataset. In addition, the proposed ULNet model is trained for 60 epochs, and the batch size is 4.

The performance of the ULNet model is measured for each fold using the confusion matrix (CM), receiver operating curve (ROC) and the following evaluation indices: precision, recall, F1-score and accuracy. The parameters used to compare performance are defined as follows: True Positive (TP) shows the number of correctly identified COVID-19; False Negative (FN) indicates the number of incorrectly identified COVID-19; Ture Negative (TN) shows the number of correctly identified non-COVID-19; False Positive (FP) indicates the number of incorrectly identified non-COVID-19.

$$\begin{aligned} &\text{Precision} = \text{TP}/(\text{TP}+\text{FP}) \\ &\text{Recall} = \text{TP}/(\text{TP}+\text{FN}) \\ &F1-\text{score} = 2\text{TP}/(2\text{TP}+\text{FP}+\text{FN}) \\ &\text{Accuracy} = (\text{TP}+\text{TN})/(\text{TP}+\text{FP}+\text{TN}+\text{FN}) \end{aligned} \tag{7}$$

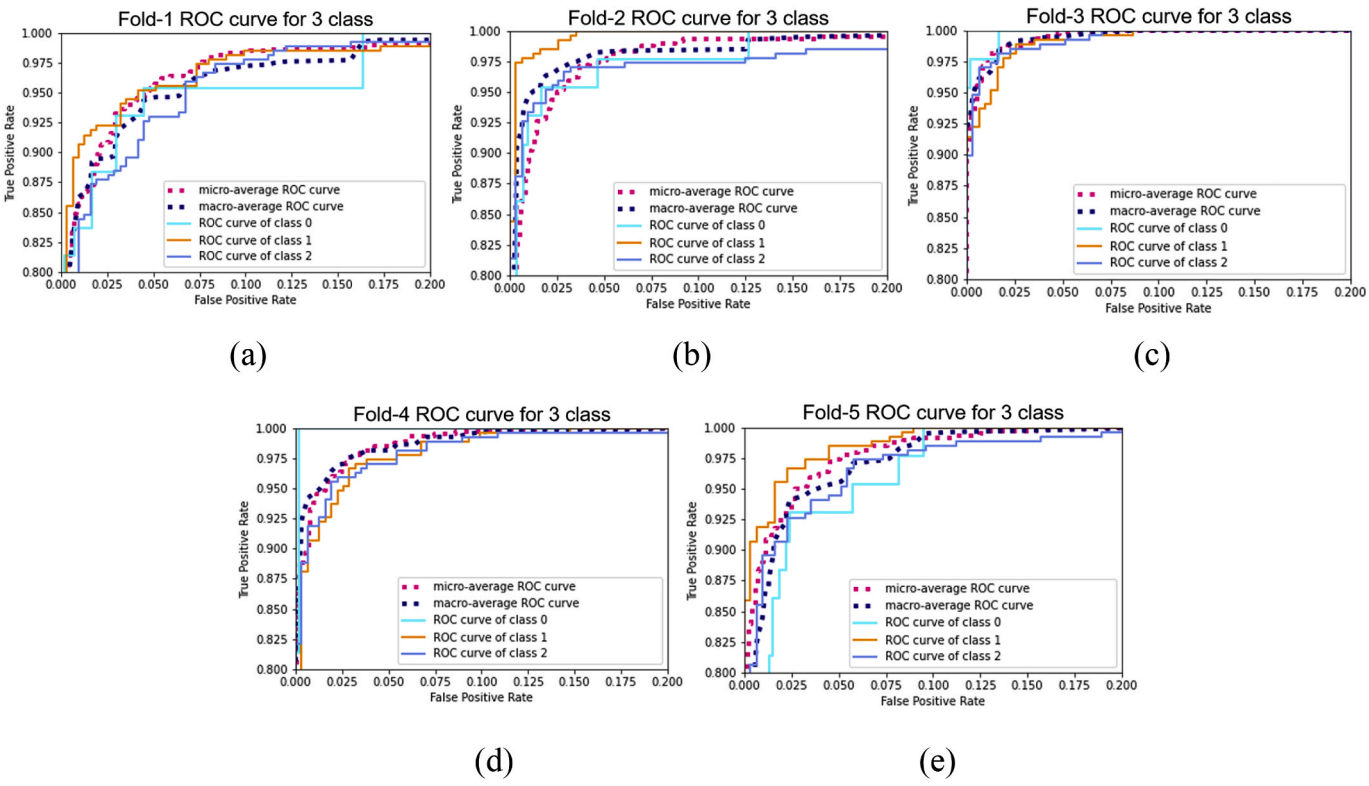

Fig. 8. Receiver operating curve of our ULNet for 3-class classification. (a) Fold-1 ROC, (b) Fold-2 ROC, (c) Fold-3 ROC, (d) Fold-4 ROC, (e) Fold-5 ROC. Class 0 is COVID-19, class 1 is normal, and class 2 is pneumonia.

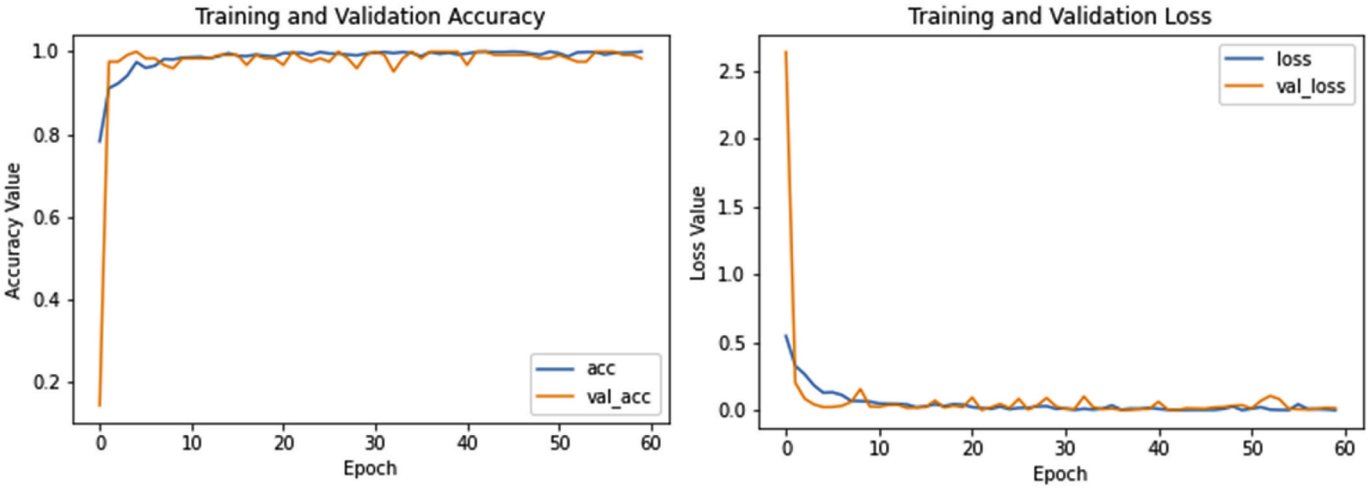

Fig. 9. Training and validation accuracy and loss for 2-class classification.

4.3. Results and discussion

4.3.1. Results of binary classifications

On Kaggle's COVID-19 Radiography Dataset, for two classifications, the overlapping and each fold individual CM are shown in Fig. 4, and each fold individual ROC curve is shown in Fig. 5. Moreover, precision, recall, F1-score and accuracy computed for each class (COVID-19 and Normal) as well as for each fold are presented in Tables 4–8. It can be seen from Table 9 that for Fold-1, precision, recall, F1-score and accuracy are 99.82%, 98.84%, 99.32% and 99.68%, respectively; for Fold-2 and Fold-3, precision, recall, F1-score and accuracy are 95.75%, 99.26%, 97.41% and 98.72%, respectively; and for Fold-4 and Fold-5, precision, recall, F1-score and accuracy are all 100%. Therefore, for the two classifications, the average precision, recall, F1-score, and accuracy are 98.26%, 99.47%, 98.83%, and 99.42%, respectively. In addition, on the QaTa-COV19 dataset, the performance of each fold of our ULNet model is presented in Table 10. The average precision, recall, F1-score, and accuracy are 99.31%, 99.30%, 99.30%, and 99.30%, respectively.

4.3.2. Results of three classifications

On Kaggle's COVID-19 Radiography Dataset, for three classifications, the overlapping and each fold individual CM are shown in Fig. 6, and each fold individual ROC curve is shown in Fig. 7. Moreover, precision, recall, F1-score and accuracy computed for each class (COVID-19, normal and viral pneumonia) as well as for each fold are presented in

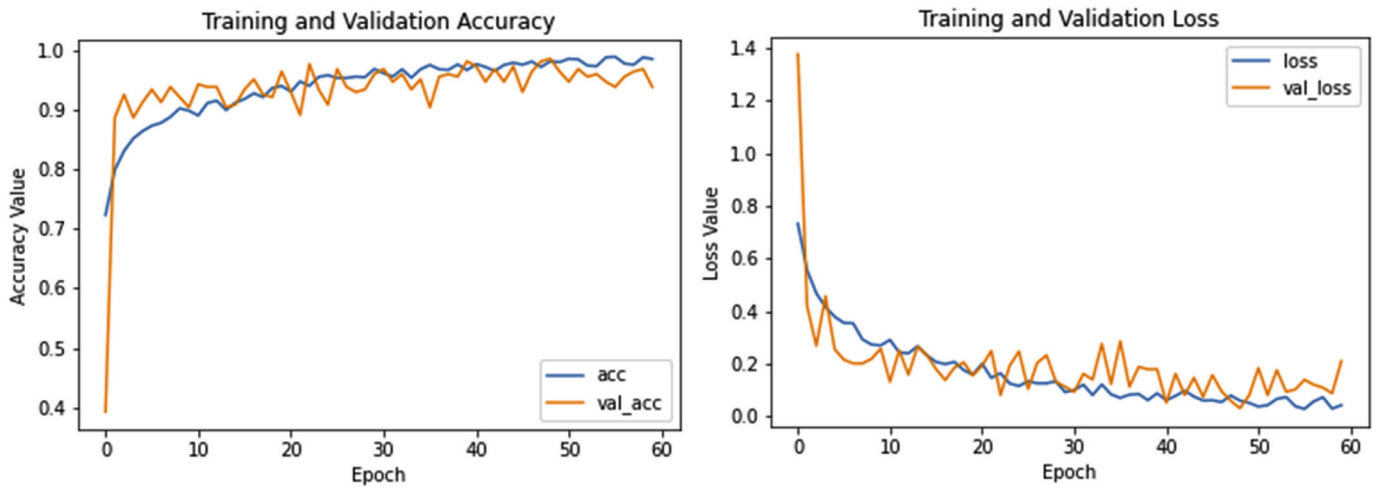

Fig. 10. Training and validation accuracy and loss for 3-class classification.

Table 18
Performance of ULNet for 2 classes in each fold.

Folds	Precision	Recall	F1-score	Accuracy
Fold-1	92.19%	75.21%	80.77%	92.63%
Fold-2	82.58%	95.73%	87.22%	92.63%
Fold-3	97.70%	84.89%	89.92%	95.83%
Fold-4	78.16%	88.19%	81.82%	89.74%
Fold-5	86.73%	92.94%	89.45%	94.55%
Average	87.48%	87.39%	85.84%	93.08%

Table 19
Performance of ULNet for 3 classes in each fold.

Folds	Precision	Recall	F1-score	Accuracy
Fold-1	94.50%	89.31%	91.55%	93.29%
Fold-2	91.77%	88.81%	90.16%	92.60%
Fold-3	83.01%	92.13%	85.88%	90.88%
Fold-4	86.67%	89.56%	87.97%	90.02%
Fold-5	84.00%	88.97%	85.97%	91.91%
Average	87.99%	89.76%	88.31%	91.74%

Table 20
Performance of the proposed ULNet without skip connections between U-shaped and L-shaped networks for 2 classes in each fold.

Folds	Precision	Recall	F1-score	Accuracy
Fold-1	90.74%	93.87%	92.21%	96.15%
Fold-2	91.35%	98.33%	94.41%	97.12%
Fold-3	95.95%	95.95%	95.95%	98.08%
Fold-4	93.72%	91.12%	92.36%	96.47%
Fold-5	97.07%	96.14%	96.60%	98.40%
Average	93.77%	95.08%	94.31%	97.24%

Table 21
Performance of the proposed ULNet without skip connections between U-shaped and L-shaped networks for 3 classes in each fold.

Folds	Precision	Recall	F1-score	Accuracy
Fold-1	92.80%	90.36%	91.50%	92.94%
Fold-2	91.03%	95.66%	93.02%	93.98%
Fold-3	91.52%	94.57%	92.91%	95.18%
Fold-4	87.87%	94.39%	90.41%	93.12%
Fold-5	91.19%	94.57%	92.64%	95.18%
Average	90.88%	93.91%	92.10%	94.08%

Tables 11–15. It can be seen from Table 16 that for Fold-1, precision, recall, F1-score and accuracy are 92.06%, 93.68%, 92.35% and 94.84%, respectively; for Fold-2, precision, recall, F1-score and accuracy are 92.91%, 94.61%, 93.71% and 94.32%, respectively; for Fold-3, precision, recall, F1-score and accuracy are 93.50%, 95.72%, 94.55% and 95.87%, respectively; for Fold-4, precision, recall, F1-score and accuracy are 96.40%, 97.02%, 96.71% and 95.87%, respectively; for Fold-5, precision, recall, F1-score and accuracy are 91.38%, 95.35%, 93.10% and 95.35%, respectively. Therefore, for the three classifications, the average precision, recall, F1-score and accuracy of the proposed ULNet are 93.25%, 95.28%, 94.09% and 95.24%, respectively. In addition, on the QaTa-COV19 dataset, the performance of each fold of our ULNet model is presented in Table 17. The average precision, recall, F1-score, and accuracy are 96.93%, 96.60%, 96.60%, 96.60%, and 96.60%, respectively.

Comparing the above experimental results with the data in Table 1, it can be seen that our ULNet model is better than the research in Table 1 in terms of precision and accuracy for the 2-class classification task (COVID-19 vs. Normal) and 3-class classification task (COVID-19 vs. Normal vs. Viral Pneumonia). The ULNet model successfully classifies COVID-19, and the accuracy of binary classifications is 99.42%, which is higher than the highest accuracy obtained by Apostolopoulos [14]; the accuracy of three classifications is 95.25%, which is higher than the highest accuracy obtained by Khan [17]. More importantly, the accuracy of positive COVID-19 in our model is 99.53% for two classifications, which is higher than 97% achieved by Harsh Panwar [26], and the accuracy of positive COVID-19 is 95.35% for three classifications. Moreover, our ULNet model also achieved good results on the QaTa-COV19 dataset. All of these results are very important and meaningful for the detection of COVID-19. In summary, our model performs well in detecting COVID-19 for two-class classification and three-class classification problems. Therefore, this model is expected to help doctors diagnose COVID-19. In addition, we performed data balancing using random up-and-down flipping of images, random left-right flipping of images, random changes in image brightness, and random changes in image contrast, but the accuracy of our results was essentially the same before and after applying data balancing solutions.

In addition, we showed the plots of accuracy and loss of training and validation with the epoch of training sessions for the first fold of 2-class and 3-class classifications, as shown in Fig. 8 and Fig. 9 (see Fig. 10).

4.4. Ablation experiment

In this section, we perform ablation experiments including UNet and ULNet without skip connections between U-shaped and L-shaped

networks for classification. The experimental results for UNet are shown in Table 18 and Table 19, and those for ULNet without skip connections between U-shaped and L-shaped networks are shown in Table 20 and Table 21.

The experimental results show that both an L-shaped network with increased network depth and skip connections incorporating different levels of features improve the classification performance.

5. Conclusion

In this paper, we propose the ULNet model for detecting COVID-19 cases by using chest X-ray images, which can accurately distinguish COVID-19 patients, normal people and viral pneumonia patients. This model was trained on a Kaggle dataset, and the classification ability of the ULNet model was examined. The experimental results show that ULNet achieves an average accuracy of 99.53% for detecting COVID-19 in 2-class classification and an average accuracy of 99.42% for 2-class (COVID-19 vs. Normal), and an average accuracy of 95.35% for detecting COVID-19 in 3-class classification and an average accuracy of 95.25% for 3-class (COVID-19 vs. Normal vs. Viral Pneumonia). The proposed model achieved 99.30% accuracy on the two classifications and 96.60% accuracy on the external test dataset. From these promising results, we can see that our ULNet model is a convolutional neural network with higher classification accuracy. At the same time, if possible, our ULNet is expected to achieve higher classification accuracy on a larger dataset. Therefore, our ULNet model can be an effective tool for doctors to quickly detect COVID-19.

In future research, we will verify the proposed ULNet model through more chest X-ray images from different hospitals. At the same time, we will use the model to address chest CT images for detecting COVID-19. We will also add salience maps to improve the interpretability of the model, which is also important for clinical practice. Moreover, we intend to use this model for the detection of COVID-19 caused by mutant virus and other types of pneumonia.

Declaration of competing interest

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

Acknowledgments

The authors thank the editor and reviewers for their comments and suggestions on the manuscript. This work is supported by the National Natural Science Foundation of China under Grants 11772081 and 11972106.

References

[1] World Health Organization, WHO Coronavirus (COVID-19) Dashboard, 2021. (Accessed August 2021). https://www.who.int/emergencies/diseases/novel-coronavirus-2019.
[2] E. Mahase, Coronavirus: covid-19 has killed more people than SARS and MERS combined, despite lower case fatality rate, BMJ 368 (2020) m641, https://doi.org/10.1136/bmj.m641.
[3] COVID-19 symptoms. https://www.who.int/health-topics/coronavirus#tab=tab_3. (Accessed August 2021).
[4] WHO, Coronavirus disease (covid-2019) r&d, Last accessed on Mar 2020, https://www.who.int/blueprint/priority-diseases/key-action/novel-coronavirus/en/.
[5] Gozes O, Frid-Adar M, Greenspan H, 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. arXiv preprint arXiv:2003.05037.
[6] D. Fanelli, F. Piazza, Analysis and forecast of covid-19 spreading in China, Italy and France, Chaos, Solit. Fractals 134 (2020) 109761.
[7] W. Wang, Y. Xu, R. Gao, et al., Detection of SARS-CoV-2 in different types of clinical specimens[J], JAMA The Journal of the American Medical Association 323 (18) (2020) 1843–1844.
[8] V.M. Corman, O. Landt, M. Kaiser, R. Molenkamp, A. Meijer, D.K. Chu, T. Bleicker, S. Brünink, J. Schneider, M.L. Schmidt, D.G. Mulders, Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR, Euro Surveill. 25 (3) (2020) 2000045.
[9] A. Bernheim, X. Mei, M. Huang, et al., Chest CT findings in coronavirus disease-19 (COVID-19): relationship to duration of infection[J], Radiology 295 (3) (2020) 200463.
[10] X. Xie, Z. Zhong, W. Zhao, et al., Chest CT for typical 2019-nCoV pneumonia: relationship to negative RT-PCR testing[J], Radiology 296 (2) (2020) 200343.
[11] Y. Fang, H. Zhang, J. Xie, et al., Sensitivity of chest CT for COVID-19: comparison to RT-PCR[J], Radiology 296 (2) (2020) 200432.
[12] F. Wu, S. Zhao, B. Yu, et al., A new coronavirus associated with human respiratory disease in China, Nature 579 (2020) 265–269, https://doi.org/10.1038/s41586-020-2008-3.
[13] Z.Y. Zu, M.D. Jiang, P.P. Xu, W. Chen, Q.Q. Ni, G.M. Lu, et al., Coronavirus disease 2019 (COVID-19): a perspective from China, Radiology 296 (2) (2020) E15–E25, https://doi.org/10.1148/radiol.2020200490.
[14] D. Shen, G. Wu, H.I. Suk, Deep learning in medical image analysis[J], Annu. Rev. Biomed. Eng. 19 (1) (2017) 221–248.
[15] I.D. Apostolopoulos, T. Bessiana, Covid-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks[J], Physical and Engineering Sciences in Medicine 43 (2020) 635–640.
[16] Hemdan E.E., Shouman M.A., Karar M.E. Covidx-net: a framework of deep learning classifiers to diagnose covid-19 in X-ray images. arXiv preprint arXiv:2003.11055.
[17] Tulin Ozturk, et al., Automated detection of COVID-19 cases using deep neural networks with X-ray images[J], Comput. Biol. Med. 121 (2020) 103792.
[18] 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[J], Comput. Methods Progr. Biomed. 196 (2020) 105581.
[19] P.K. Sethy, K. Santi, Behera, et al., Detection of coronavirus disease (COVID-19) based on deep features and support vector machine[J], International Journal of Mathematical, Engineering and Management Sciences 5 (4) (2020) 643–651.
[20] Suat Toraman, Talha Burak Alakus, Ibrahim Turkoglu, Convolutional CapsNet: a novel artificial neural network approach to detect COVID-19 disease from X-ray images using capsule networks[J], Chaos, Solit. Fractals 140 (2020) 110122.
[21] Zhang J, Xie Y, Liao Z, et al. Viral Pneumonia Screening on Chest X-Ray Images Using Confidence-Aware Anomaly Detection. arXiv preprint arXiv:2003.12338.
[22] Ghoshal B, Tucker A. Estimating Uncertainty and Interpretability in Deep Learning for Coronavirus (COVID-19) Detection. arXiv preprint arXiv:2003.10769.
[23] H. Panwar, P.K. Gupta, M.K. Siddiqui, et al., Application of deep learning for fast detection of COVID-19 in X-Rays using nCOVnet[J]. Chaos, Solitons & Fractals 138 (2020) 109944.
[24] Tawsifur Rahman, Amith Khandakar, Yazan Qiblawey, et al., Exploring the effect of image enhancement techniques on COVID-19 detection using chest X-ray images [J], Comput. Biol. Med. 132 (2021) 104319.
[25] Wang S., et al. A deep learning algorithm using CT images to screen for Corona Virus Disease (COVID-19). medRxiv, preprint DOI: https://doi.org/10.1101/2020.02.14.20023028.
[26] Zheng, C, Deng, X, Fu, Q, Zhou, Q, Feng, J, et al. Deep learning-based detection for COVID-19 from chest CT using weak label, medRxiv, preprint DOI: https://doi.org/10.1101/2020.0 3.12.20 027185.
[27] L. Li, L. Qin, Z. Xu, Y. Yin, X. Wang, B. Kong, et al., Artificial intelligence distinguishes COVID-19 from community acquired pneumonia on chest CT, Radiology 296 (2) (2020) 200905.
[28] Song Y, Zheng S, Li L, Zhang X, Zhang X, Huang, et al. Deep learning enables accurate diagnosis of novel Coronavirus (COVID-19) with CT images. medRxiv, preprint DOI: https://doi.org/10.1101/2020.02.23.20026930.
[29] O. Ronneberger, P. Fischer, T. Brox, U-net: convolutional networks for biomedical image segmentation, Medical Image Computing and Computer-Assisted Intervention – MICCAI 9351 (2015) 234–241, 2015.
[30] Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift. arXiv preprint arXiv:1502.03167.
[31] A. Degerli, M. Ahishali, M. Yamac, et al., COVID-19 infection map generation and detection from chest X-ray images, Health Inf. Sci. Syst. 9 (2021) 15, https://doi.org/10.1007/s13755-021- 00146-8.
[32] M.E.H. Chowdhury, T. Rahman, A. Khandakar, R. Mazhar, M.A. Kadir, Z. B. Mahbub, K.R. Islam, M.S. Khan, A. Iqbal, N. Al-Emadi, M.B.I. Reaz, M.T. Islam, Can AI help in screening Viral and COVID-19 pneumonia? IEEE Access 8 (2020) 132665–132676.
[33] Societa Italiana di Radiologia Medical Interventistica, COVID-19 database. [Online]. Available: https://www.sirm.org/category/senzacategoria/covid-19/, 2020. (Accessed November 2020).
[34] J.C. Monteral, COVID-chestxray database [Online] Available: https://github.com/ieee8023/covid-chestxray-dataset. (Accessed November 2020).
[35] P. Mooney, Chest X-ray images (pneumonia) [Online] Available: https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia, 2018. (Accessed November 2020).