

# ACRnet: Adaptive Cross-transfer Residual neural network for chest X-ray images discrimination

Wenyu Zhang (✉ [zhangwenyu8518@126.com](mailto:zhangwenyu8518@126.com))

---

## Research article

**Keywords:** Convolutional neural network, Chest X-ray image, Deep convolution generative adversarial networks, Cardiothoracic diseases

**Posted Date:** November 2nd, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-1015912/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

**Background:** Cardiothoracic diseases are a serious threat to human health and chest X-ray images have great reference value for the diagnosis and treatment. However, it is difficult for professional doctors to accurately diagnose cardiothoracic diseases through chest X-ray images sometimes and there will be different understanding based on human subjective differences, which will affect the judgment and treatment of diseases. Therefore, it is very necessary to develop a high-precision neural network to recognize chest X-ray images of cardiothoracic diseases.

**Methods:** In this work, we cross-transfer the information extracted by the residual block and by the adaptive structure to different levels, which avoids the reduction of the adaptive function by residual structure and improves the recognition performance of the model. To evaluate the recognition ability of ACRnet, VGG16, InceptionV2, ResNet101 and CliqueNet are used for comparison. In addition, we use the deep convolution generative adversarial network (DCGAN) to expand the original dataset.

**Result:** We find ACRnet has better recognition ability than other networks in identifying cardiomegaly, emphysema and normal. Besides, ACRnet's recognition ability has been greatly improved after data expansion.

**Conclusions:** The experimental result indicates that emphysema and cardiomegaly can be effectively identified by ACRnet. Using DCGAN technology to expand the dataset can further improve the recognition ability of the model.

## Background

Cardiothoracic diseases mainly include lung disease and heart disease. The lung which is connected with the outside world is the main organ of the respiratory system, and the blood of the whole body flows through it. The heart plays a major role in the body's circulatory system, pushing blood around the body and carrying away metabolic waste. Both lung and heart are important organs for us. According to the data from World Health Organization, lung disease is the third most lethal disease in the world, while heart disease is the first one. Cardiothoracic diseases cause more than ten million deaths worldwide every year.

Emphysema is characterized by increased lung volume, excessive bronchial dilation, decreased bronchial airway elasticity, and destruction of airway walls. The disease is usually caused by smoking, industrial dust, chronic bronchitis and pulmonary inflammation. Emphysema mainly includes: senile emphysema, compensatory emphysema, interstitial emphysema and obstructive emphysema. Emphysema is a common lung disease, which can lead to pulmonary insufficiency, pulmonary hypertension, pulmonary heart disease and so on. Cardiomegaly is an increase in heart load caused mainly by circulatory and respiratory diseases, it can be divided into physiological and pathological. The typical manifestations of the disease are chest tightness, palpitation, dyspnea and even heart failure. There are many reasons for

cardiomegaly, among which coronary heart disease, hypertension, pulmonary hypertension and cor pulmonale are the main ones.

Chest X-ray is the most commonly used medical technology for the diagnosis of cardiothoracic diseases, including emphysema and cardiomegaly. It is very effective in identifying and detecting cardiothoracic, pulmonary and interstitial diseases, and plays an important role in the treatment of cardiothoracic diseases [1]. Accurately analyzing patients' condition is a huge challenge for radiologists, and computer-aided diagnosis can reduce differences in radiologists' conclusions and provide reference for clinicians in abnormal cases. In recent years, we have been using deep learning in many fields which includes medical research and computer vision and made a significant progress [2–6]. related studies include the detection of pneumonia [7–9], pneumothorax [10], tuberculosis [11], pulmonary nodules[12, 13], lung cancer [14–16], and COVID-19 [17–19]. Here we summarize the current literature on deep learning technology for cardiothoracic disease detection.

Lacking data is a big problem for the application of deep learning in chest X-ray image diagnosis. Qasim et al. [18] used data expansion technology to increase the chest X-ray images of COVID-19 from 219 to 1000, which was used to train models to recognize chest X-ray images. By using these synthetic images to expand the existing data, the diversity of training data was increased and the classification performance was improved. Loey et al.[19] used Alexnet, GoogleNet and Resnet18 to detect chest X-ray images of COVID-19, normal, bacterial and viral pneumonia, in order to improve the recognition accuracy, they used the GAN network to expand the original dataset. GoogleNet reaches the highest detection accuracy of 99 % for the task of classifying normal chest and COVID-19. In the classification task with three specimens containing COVID-19, Alexnet reaches the highest detection accuracy, ie 85.2 % . Although using GAN for data enhancement helps to improve the detection accuracy of the model, it is difficult for shallow networks to correctly identify disease types from different chest X-ray images. Ferreira et al. [20] used the DenseNet model to detect cardiomegaly in chest X-ray images and they used a public dataset to train the neural network, and a private dataset network to verify the generalization ability of the model. DenseNet obtained an AUC value of 0.818 on the public dataset and 0.809 on the private dataset. Rajpurkar et al. [21]. developed the CheXNeXt network model by using Chest X-ray 8 data set, and compared the AUC values obtained by the model with those obtained by cardiothoracic radiologists. The AUC values of cardiomegaly and emphysema obtained by radiologists were 0.888 and 0.911 respectively, while those obtained by CheXNeXt were 0.831 and 0.704. CheXNeXt's performance in differentiating cardiomegaly and emphysema is far from that of radiologists. For the Chest X-ray 14 dataset, Angelica et al. proposed GraphXNET [22], and achieved good results on chest X-ray 14 dataset. Wang et al. proposed TieNet [23]. Compared with the previous classical neural networks, TieNeT has significantly improved the recognition ability of chest X-ray images (AUCs increase by 6 % on average). Zhao et al. [24] proposed AM \_ DenseNet for chest X-ray image classification. This model used a dense connection network and added a attention module after each dense block to optimize the ability of the model to extract features. They used the Focal Loss function to solve the data imbalance problem. The average AUC of AM \_ DenseNet in detecting the probability of chest diseases reached 0.8537. Ho et al.[25] proposed a knowledge distillation strategy for different kinds of diseases and a remarkable mapping technique for visualizing

abnormal features in X-ray images. As they mentioned, although the total number of the dataset is more than 100,000, a large number of chest X-ray images are in various diseases types. If the data are not screened and classified directly, it will cause great interference to the model and seriously affect the discriminant results. In this work, we select only cardiomegaly, emphysema and normal from the Chest X-ray 14 dataset for research. Our contributions are as follows :

1. We construct an adaptive cross-transfer residual neural network to identify three kinds of chest X-ray images. To identify the chest X-ray images better, we alternately transmit the feature information extracted by adaptive structure and residual block to different layers.
2. We use VGG16, InceptionV2, ResNet101 and CliqueNet to compare the recognition ability with ACRnet. The results show that the recognition performance of ACRnet is better than that of the four classical models. In addition, compared with state-of-the-art convolutional neural networks, ACRnet has better recognition ability too.
3. To further improve the recognition performance of ACRnet, we use DCGAN to expand the original data set. Compared with the previous, the recognition ability of ACRnet which trained by the expanded data is further improved.

The remainder of this paper is organized as follows: Section 2 describes methods in detail. Section 3 shows experimental result. Section 4 concludes the paper.

## Methods

### Design

This section describes the overall process of using the expanded dataset to train ACRnet to identify chest X-ray images. As shown in Fig. 1, during the whole process, we first use DCGAN to synthesize cardiomegaly and emphysema chest X-ray images, and then, send the expanded dataset into ACRnet. After several rounds of feature extraction, the input images will be finally recognized by the model and the category label will be output. We call module A as a block, and ACRnet is composed of multiple blocks in series.

### Data description

The images we use come from the Chest X-ray 14 public dataset, each with single or multiple pathological markers. Radiological reports show that the accuracy of pathological markers is more than 90 % [25]. We select images only suffering from cardiomegaly or emphysema from Chest X-ray 14 public dataset, and add normal images for research. We divide all data into 80 % for training and 20 % for test. Table 1 shows the number of chest X-ray images we screen from the Chest X-ray 14 dataset for training and test. Due to the small number of both cardiomegaly and emphysema ( less than 1000 ) in the original dataset, which is not conducive to training the network model, we use DCGAN to expand the original data set. We synthesize 2000 images of cardiomegaly and emphysema, since there are more than 60000 normal chest radiography images in the chest X-ray 14 dataset, we select 2000 normal images together

with the synthesized images to expand the original dataset. We use the same test set when we test the models trained with the original dataset and the ACRnet trained with the expanded dataset. In order to meet the requirements of the model for the input size, we resize both the original  $1024 \times 1024$  pixels image and the  $256 \times 256$  pixels image synthesized by DCGAN into  $224 \times 224$  pixels image.

**Table 1.** Number of the chest X-ray images in training and test.

Disease	Train set	Test set
cardiomegaly	840	210
emphysema	720	180
normal	800	200

## Deep convolution generative adversarial networks

Deep convolutional generative adversarial network is a special deep learning model which combines Convolutional neural network (CNN) and GAN[26]. DCGAN consists of two different types of networks, one is a generator network for synthesizing images, the other is a discriminator network for identifying real and synthetic images. In the process of training DCGAN, the discriminator and generator are trained simultaneously. The generator network we used consists of 7 transposed convolution layers, 6 ReLU layers, 3 batch normalization layers and 1 tanh layer on the end. The discriminator network consists of 7 convolution layers, 6 ReLU layers, 3 batch normalization layers and a Sigmoid layer on the end. All the convolutional and transposed convolutional layers have the same window size of  $4 \times 4$  pixel with 16 filters for each layer. DCGAN network can solve insufficient training caused by small dataset. DCGAN is trained by using all the chest radiographs of emphysema and cardiac hypertrophy from the original data. After 1000 iterations, we use DCGAN to synthesize 2000 chest X-ray images of emphysema and cardiomegaly to expand the training dataset. Table 2 shows the model structure and parameters of generator and discriminator in DCGAN.

## Overview of convolutional neural networks [CNNs]

CNNs training begins in a feedforward fashion. Input information is transmitted from the initial layer to the output layer, and errors are propagated forward from the last layer [27]. In this section, we will introduce 4 neural networks used in the experiments. VGG is a typical and effective neural network with 6 different configurations, of which VGG16 and VGG19 are commonly used [27]. The VGG model usually consists of 5 blocks, each containing several convolution layers and a pooling layer. VGG16 includes 13 convolution layers, 3 fully connected layers and 5 pooling layers. The convolution layer of VGG16 adopts  $3 \times 3$  convolution kernels, and the pooling layer adopts  $2 \times 2$  pooling kernels. InceptionV2 is an upgraded version of GoogleLeNet (InceptionV1) [28]. InceptionV1 is a 22-layer network structure with four channels,

using convolution kernels of  $1 \times 1$ ,  $3 \times 3$  and  $5 \times 5$  sizes. Since convolution kernels of different sizes have different receptive fields, InceptionV1 can better learn features of different scales.

In addition, in InceptionV1, visual information is aggregated at different scales, so that the subsequent network level can extract features from different scales. InceptionV2 first propose using batch normalization to accelerate network training and prevent gradient from disappearing [29]. InceptionV2 also replaces every convolution kernel of  $5 \times 5$  in InceptionV1 with two  $3 \times 3$  convolution kernels, which reduces the parameters and enhances the nonlinear properties while maintaining the same receptive field. ResNet is a widely used neural network with 5 structures of 18-layer, 32-layer, 50-layer, 101-layer and 152-layer [30]. The residual block in ResNet uses the jump connection structure to alleviate the gradient disappearance caused by the increase of network depth. Using residual structure not only improves the accuracy of the model, but also makes the model easier to optimize. With the deepening of the neural network, people find that improving the flow of information in the deep network model can make the neural network easier to be trained. In the block of CliqueNet [31], each layer is the input also the output of other layers, all these layers constitute a ring structure and update alternately. This structure makes high-level visual information be sent back to the previous level for feature reuse. In addition, the introduction of attention mechanism inhibits the irrelevant neurons representing background and noise, which further improves the recognition ability of the model.

**Table 2.** The model structure and parameters of generators and discriminator in DCGAN.

Generator			Discriminator		
Layer name	Input	output	Layer name	Input	output
ConvTranspose2d	100	16 x 32	Conv2d	3	16
BatchNorm2d			LeakyReLU		
ReLU			Conv2d	16	16 x 2
ConvTranspose2d	16 x 32	16 x 16	LeakyReLU		
BatchNorm2d			Conv2d	16 x 2	16 x 4
ReLU			LeakyReLU		
ConvTranspose2d	16 x 16	16 x 8	Conv2d	16 x 4	16 x 8
BatchNorm2d			BatchNorm2d		
ReLU			LeakyReLU		
ConvTranspose2d	16 x 8	16 x 4	Conv2d	16 x 8	16 x 16
ReLU			BatchNorm2d		
ConvTranspose2d	16 x 4	16 x 2	LeakyReLU		
ReLU			Conv2d	16 x 16	16 x 32
ConvTranspose2d	16 x 2	16	BatchNorm2d		
ReLU			LeakyReLU		
ConvTranspose2d	16	3	Conv2d	16 x 32	1
Tanh			Sigmoid		

## Constitution of ACRnet

This section introduces the structure of ACRnet, which is a 97-layer neural network. In ACRnet, we construct an adaptive cross-transfer residual structure to extract feature information to improve the recognition efficiency of the model.

To some extent, the feature information extracted by convolution kernel is related to the channel in convolution neural network. Adaptive learning is to obtain the importance of different features through autonomous learning, and then suppress secondary features and enhance main features according to importance[32]. In the adaptive module, the adaptive global average pooling layer ( Adaptive Avg Pool ) is used to compress the feature along the spatial dimension, and each two-dimensional feature is transformed into a figure which has global sensing ability and reflects the feature distribution. And then we use  $1 \times 1$  convolution kernel ( input channel is  $a$ , output channel is  $a / 4$  ) for dimension reduction, then

use ReLU activation function to increase nonlinearity, and finally use another  $1 \times 1$  convolution kernel ( input channel is  $a/4$ , output channel is  $a$  ) for dimension increase, so as to reduce the computational cost. In the last layer of the adaptive structure, the Sigmoid function is used to generate a figure between 0 and 1 as the output result. Passing adaptive figure to the hierarchy behind the network is equivalent to multiplying the eigenmatrix by a weight coefficient.

In order to extract image feature better, we combine adaptive and residual structure to construct a deep neural network model. In ACRnet, the residual structure transmits the features of layer 0 and layer 1 to the end of layer 3 and 4. The adaptive structure transmits the features of layer 0 and layer 2 to the end of layer 2 and layer 4. This structural design avoids the reduction effect of residual structure on adaptive function when residual and adaptive structure transmit the feature information extracted from the same layer to the same layer behind. The adaptive structure generates a characteristic adaptive coefficient between 0 and 1, and the residual structure generates a matrix. Adaptive and residual structures transfer feature information from the same layer to the same subsequent network layer, which is equivalent to adding the feature matrix generated by the residual structure to a matrix multiplied by the adaptive coefficient, resulting in the adaptive function weakened by the residual structure. Fig. 2 demonstrates ACRnet's adaptive cross-transfer residual block.

### Hyper-parameters tuning

In our work, we try to estimate the influence of various hyperparameters on the performance of the depth model. We divide all the parameters into external parameters and internal parameters and we fine-tune the parameters for each model. The external parameters of the model include activation functions, learning rate and batch size, which affect model performance a lot. We find that Adam is significantly better than Adadelta or SGD to select activation functions. Learning rate affects the convergence speed of the model, but excessive learning rate will reduce the accuracy of the model. We tried seven different learning rates  $\{ 1e-3, 1e-4, 1e-5, 3e-3, 3e-4, 3e-5, 3e-6 \}$  and find it's better in  $1e-4$  or  $3e-4$ . Experimenting repeatedly, we select 32 as batch size from  $\{ 16, 24, 32, 48, 64 \}$ . The internal parameters in the neural network include the size of convolution kernel, stride, the size of pooling layer window, channel and dropout. In the VGG16, InceptionV2, ResNet101 and CliqueNet, we choose the original internal parameters. Table 3 shows the structure and parameters of ACRnet model.

**Table 3.** Structure and parameters in ARnet.

layer	Output size	Kernel size	channel
Input (224x224x3)			
block x 4	112 x 112	3	60
block x 4	56 x 56	3	120
block x 4	28 x 28	3	240
block x 6	14 x 14	3	480
block x 6	7 x 7	3	960
AvgPool (kernel_size =7, stride=1)			
Linear (output=3)			

## Performance Evaluation

We choose Inception Score ( IS ) [33] and Fréchet Inception Distance ( FID ) [34] as indexes to evaluate the quality of synthetic images, both of them use InceptionV3 [37], pre-trained on ImageNet [36], to classify the generated images. The higher IS, the lower FID, and the better image quality. We introduce accuracy (ACC) [38], Precision, Recall and F1 score [39] as indicators to evaluate the performance of the detection model. The formulas for these indicators are as follows.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

Here TP, FP, TN and FN represent the number of true positive, false positive, true negative and false negative, respectively. In addition, in order to further evaluate the model performance [41, 42], we used the confusion matrix [40] and plotted the receiver operating characteristic (ROC) curve and the area under the curve (AUC). The higher the AUC, the better performance.

## Result

This section introduces the result of recognizing cardiomegaly, emphysema and normal by ACRnet and other four classical neural networks: VGG16, InceptionV2, ResNet101 and CliqueNet. In order to improve the recognition ability of ACRnet, we use DCGAN to expand the original dataset. Fig. 3 shows the original images from the dataset as well as the synthetic images generated by DCGAN, which are very similar. We calculate the IS scores of them respectively, and find that the scores are very close, indicating that the

quality of the synthetic images is good. This conclusion is proved after calculating the FID scores, because we sent them into the detection model simultaneously and obtain a relatively low FID value. Table 4 shows the IS and FID scores.

**Table 4.** IS and FID scores

	original images	synthetic images
IS	3.72	3.54
FID	77.6	

**Table 5.** The mean precision, recall, F1-score and AUC values obtained by ACRnet, VGG16, InceptionV2, ResNet101 and CliqueNet in detecting cardiomegaly, emphysema and normal using 5-fold cross validation.

Model	Disease	Precision	Recall	F1-score	AUC
ACRnet +	Cardiomegaly	<b>0.8532</b>	<b>0.8857</b>	<b>0.8691</b>	<b>0.95</b>
DCGAN	Emphysema	<b>0.8788</b>	<b>0.8056</b>	<b>0.8507</b>	<b>0.93</b>
	Normal	0.9372	<b>0.9700</b>	0.9533	0.99
ACRnet	Cardiomegaly	0.8045	0.8429	0.8233	0.92
	Emphysema	0.8137	0.7278	0.7684	0.90
	Normal	0.9282	<b>0.9700</b>	0.9486	0.99
VGG16	Cardiomegaly	0.7093	0.7667	0.7369	0.89
	Emphysema	0.6590	0.6333	0.6456	0.84
	Normal	<b>0.9789</b>	0.8857	0.9300	0.99
InceptionV2	Cardiomegaly	0.8143	0.6476	0.7214	0.90
	Emphysema	0.6473	<b>0.8056</b>	0.7178	0.87
	Normal	0.9548	0.9500	0.9524	<b>1.00</b>
ResNet101	Cardiomegaly	0.8086	0.6762	0.7365	0.88
	Emphysema	0.6872	0.7444	0.7147	0.83
	Normal	0.8995	0.9850	0.9437	<b>1.00</b>
CliqueNet	Cardiomegaly	0.6768	0.8476	0.7537	0.90
	Emphysema	0.7121	0.5222	0.6025	0.87
	Normal	0.9692	0.9450	<b>0.9569</b>	0.96

Table 5 shows the experimental results of different neural networks. Models are trained on all the available data in each training set, and global assessment is presented by 5-Fold cross validation on the whole dataset. When using four classical neural networks to identify cardiomegaly, the recall index obtained by CliqueNet is at a high level. The precision index obtained by InceptionV2 and ResNet101 are better than that obtained by VGG16 and CliqueNet. The F1-scores obtained by the four classical neural networks are close. The AUC index of InceptionV2 is lower than the other three networks. When using four classical neural networks to identify emphysema, CliqueNet achieves the optimal precision score. The recall and F1-score obtained by InceptionV2 and ResNet101 are significantly better than CliqueNet and VGG16. The recall index obtained by CliqueNet is the lowest. The AUC index obtained by InceptionV2 and CliqueNet are higher than the other two neural networks. When identifying normal, the F1-score and AUC index obtained by the four classical neural networks are relatively close. VGG16 achieves the best precision index and ResNet101 achieves better recall index than other networks. Overall, the existing four classical neural networks: VGG16, InceptionV2, ResNet101 and CliqueNet have their respective advantages in recognizing cardiomegaly, emphysema and normal chest X-ray images. But none of them are very good at identifying all types of chest X-ray images. ACRnet is a deep network model with cross residual structure and the adaptive mechanism enables ACRnet to extract image features better. Compared with the four classical network structures, ACRnet has higher precision, recall, F1-Score and AUC indexes in recognizing the three types of chest X-ray images. After data expansion by DCGAN, the recognition ability of ACRnet is further improved. Fig. 4 shows the detection result of all models for 3 types of chest X-ray images more intuitively.

Table 6 shows mean ACC and average-AUC index obtained by ACRnet and four classical neural networks using 5-fold cross validation. The result shows that the ACC and average-AUC values of four classical neural networks in identifying cardiomegaly, emphysema and normal are close. The overall recognition performance of ACRnet trained by the original dataset is better than that of other 4 classical neural networks. The recognition ability of ACRnet is further improved after using DCGAN to expand the original dataset.

**Table 6.** Mean ACC and average-AUC obtained by ACRnet trained from the expanded dataset and the networks trained from the original dataset.

Model	ACC	average-AUC
ACRnet+DCGAN	<b>0.8898</b>	<b>0.956</b>
ACRnet	0.8508	0.936
VGG16	0.7813	0.906
InceptionV2	0.7983	0.923
ResNet101	0.8017	0.903
CliqueNet	0.7814	0.910

**Table 7.** Comparison of the AUC between our method with 4 state-of-the-art convolutional neural networks.

methods	cardiomegaly	emphysema	normal
CheXNet	0.92	0.93	-
GraphXNet	0.88	0.84	-
TieNet	0.84	0.86	0.70
AM_DenseNet	0.91	<b>0.94</b>	-
ACRnet	0.91	0.89	<b>0.99</b>
ACRnet+DCGAN	<b>0.95</b>	0.93	<b>0.99</b>

Fig. 5 shows the confusion matrixes obtained by neural networks using 5-fold cross validation. By comparing the confusion matrixes, we find that VGG16 has better recognition ability for cardiomegaly than InceptionV2 and ResNet101, and worse recognition ability for cardiomegaly than InceptionV2 and ResNet101. Compared with the other three classical neural networks, CliqueNet has the best recognition ability for cardiomegaly, but the worst recognition ability for emphysema. Confusion matrix shows that ACRnet has a good and relatively balance recognition ability in identifying 3 types of chest X-ray images: cardiomegaly, emphysema and normal. The overall performance of ACRnet is better than that of the existing four classical neural networks. After data expansion by DCGAN, the recognition ability of ACRnet is further improved. Fig. 6 shows the ROC curves obtained by neural networks using 5-fold cross validation. The results show that the recognition ability of ACRnet for emphysema is better than the existing four classical neural networks, and the recognition ability for cardiomegaly and normal is also at a high level. After data expansion, ACRnet 's ability to recognize cardiomegaly and emphysema is significantly improved. To comprehensively evaluate the performance of our method for identifying cardiomegaly, emphysema and normal, we compare our method with several most advanced methods. Table 7 shows the AUC index of our method compared with 4 state-of-the-art convolutional neural networks: CheXNet [35], GraphXNet [22], TieNet [23] and AM\_DenseNet [24]. We can find that our method has excellent performance in identifying 3 types of chest X-ray images.

## Conclusions

Cardiothoracic diseases are a serious threat to human health, and chest radiography interpretation is very important for timely detection of cardiothoracic diseases. However, this time-consuming work is usually completed by radiologists, and subjective differences among different people will lead to missed diagnosis and misdiagnosis. In this way, using deep learning technology to build neural network model can reduce clinicians' subjective errors, which is of great significance for disease diagnosis and treatment. In this paper, we construct ACRnet to recognize cardiomegaly, emphysema and normal chest

X-ray images, and achieve good recognition effect. Using DCGAN technology to expand the existing data set also can further improve the accuracy of the model. In future research, we hope that our method can be used to distinguish more types of cardiothoracic diseases, so as to help clinicians diagnose cardiothoracic diseases more accurately.

## **Declarations**

### **Acknowledgements**

The authors thank the editor and anonymous reviewers for their constructive comments and suggestions, which greatly help us improve our manuscript.

### **Funding**

The work was supported by the Natural Science Foundation of Liaoning Province of China (No.20180551011), Foundation of Education Department of Liaoning Province (Grant No. LJKZ0280). The funding body did not play any roles in the study design; nor in the data collection, analysis and interpretation, or in the writing of the paper.

### **Ethics approval and consent to participate**

Not applicable.

### **Competing interests**

The authors declare that they have no competing interests.

### **Consent for publication**

Not applicable.

### **Authors' contributions**

Boyang Wang programmed the tool and drafted the manuscript. Wenyu Zhang and Qi Zhao tested the codes at different stages of development and critically reviewed the manuscript. All authors contributed critically to the development of the tool. All authors read and approved the final manuscript.

### **Availability of data and materials**

The dataset and full codes used in this study can be downloaded at Github:

<https://github.com/wby15281066101/ACRnet>

## **Abbreviations**

ACRnet: adaptive cross-transfer residual neural network; DCGAN: deep convolution generative adversarial network; CNN: Convolutional neural network; ACC: Accuracy; IS; Inception Score; FID; Fréchet Inception Distance; ROC: receiver operating characteristic; AUC: area under the curve.

## References

1. D. Brenner, J. McLaughlin, and R. Hung. Previous Lung Diseases and Lung Cancer Risk: A Systematic Review and Meta-Analysis. *PLOS ONE*. 2011; 6(3).
2. L. Zhang, P. Yang, H. Feng, Q. Zhao, and H. Liu. Using Network Distance Analysis to Predict lncRNA-miRNA Interactions, *Interdiscip Sci*. 2021;13(3);535–545.
3. M. Shorfuzzaman, and M. Masud. On the detection of covid-19 from chest x-ray images using cnn-based transfer learning. *Computers, Materials and Continua*. 2020;64(3); 1359–1381.
4. F. Carrillo-Perez, J. C. Morales, D. Castillo-Secilla, Y. Molina-Castro, A. Guillen *et al*. Non-small-cell lung cancer classification via RNA-Seq and histology imaging probability fusion. *BMC Bioinformatics*. 2021;22(1);454.
5. G. Liang, and L. Zheng. A transfer learning method with deep residual network for pediatric pneumonia diagnosis. *Computer Methods and Programs in Biomedicine*. 2020.
6. X. Wei, Y. Chen, and Z. Zhang. Comparative Experiment of Convolutional Neural Network (CNN) Models Based on Pneumonia X-ray Images Detection. *Proceedings - 2020 2nd International Conference on Machine Learning, Big Data and Business Intelligence, MLBDBI 2020*.
7. L. Rai, T. Popovi, S. aki, and S. andi. Pneumonia Detection Using Deep Learning Based on Convolutional Neural Network. 2021 *25th International Conference on Information Technology, IT 2021*.
8. A. G. Taylor, C. Mielke, and J. Mongan. Automated detection of moderate and large pneumothorax on frontal chest X-rays using deep convolutional neural networks: A retrospective study. *PLOS Medicine*. 2018;15(11).
9. T. K. K. Ho, J. Gwak, O. Prakash, J.-I. Song, and C. M. Park. Utilizing Pretrained Deep Learning Models for Automated Pulmonary Tuberculosis Detection Using Chest Radiography. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. 2019.
10. R. Zhang, M. Sun, S. Wang, and K. Chen. *Computed Tomography pulmonary nodule detection method based on deep learning*. US 10937157. L. Infervision Medical Technology Co. 2021.
11. C. Tong, B. Liang, Q. Su, M. Yu, J. Hu *et al*. Pulmonary Nodule Classification Based on Heterogeneous Features Learning. *IEEE Journal on Selected Areas in Communications*, 2021;39(2);574–581.
12. L. J. Hyuk, S. H. Young, P. Sunggyun, K. Hyungjin, H. E. Jin *et al*. Performance of a Deep Learning Algorithm Compared with Radiologic Interpretation for Lung Cancer Detection on Chest Radiographs in a Health Screening Population. *Radiology*. 2020;297(3).

13. A. Hosny, C. Parmar, T. P. Coroller, P. Grossmann, R. Zeleznik *et al.* Deep learning for lung cancer prognostication: A retrospective multi-cohort radiomics study. *PLOS Medicine*, 2018;15(11).
14. M. Masud, N. Sikder, A.-A. Nahid, A. K. Bairagi, and M. A. Alzain. A machine learning approach to diagnosing lung and colon cancer using a deep learningbased classification framework. *Sensors (Switzerland)*. 2021;21(3);1–21.
15. S. Roy, W. Menapace, S. Oei, B. Luijten, E. Fini *et al.* Deep Learning for Classification and Localization of COVID-19 Markers in Point-of-Care Lung Ultrasound. *IEEE Transactions on Medical Imaging*. 2020;39(8);2676–2687.
16. H. Tianqing, K. Mohammad, M. Mokhtar, P. GholamReza, T. K. S. H *et al.* Real-time COVID-19 diagnosis from X-Ray images using deep CNN and extreme learning machines stabilized by chimp optimization algorithm. *Biomedical signal processing and control*. 2021;68.
17. M. A. Khan, S. Kadry, Y.-D. Zhang, T. Akram, M. Sharif *et al.* Prediction of COVID-19 - Pneumonia based on Selected Deep Features and One Class Kernel Extreme Learning Machine. *Computers and Electrical Engineering*. 2021;90.
18. Y. Qasim, B. Ahmed, T. Alhadad, H. Al-Sameai, and O. Ali. The Impact of Data Augmentation on Accuracy of COVID-19 Detection Based on X-ray Images. *Lecture Notes on Data Engineering and Communications Technologies*. Springer Science and Business Media Deutschland GmbH. 2021.
19. M. Loey, F. Smarandache, and N. E. M. Khalifa. Within the Lack of Chest COVID-19 X-ray Dataset: A Novel Detection Model Based on GAN and Deep Transfer Learning. *Symmetry*. 2020;12(4).
20. J. R. Ferreira Junior, D. A. Cardona Cardenas, R. A. Moreno, M. D. F. De Sa Rebelo, J. E. Krieger *et al.* A general fully automated deep-learning method to detect cardiomegaly in chest x-rays. *Progress in Biomedical Optics and Imaging - Proceedings of SPIE*. The Society of Photo-Optical Instrumentation Engineers (SPIE).2021.
21. P. Rajpurkar, J. Irvin, R. L. Ball, K. Zhu, B. Yang *et al.* Deep learning for chest radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm to practicing radiologists. *PLoS Med*. 2018;15(11);e1002686.
22. A. I. Aviles-Rivero, N. Papadakis, R. Li, P. Sellars, Q. Fan *et al.* GraphX NET- Chest X-Ray Classification Under Extreme Minimal Supervision. 2019;504–512.
23. X. Wang, Y. Peng, L. Lu, Z. Lu, and R. M. Summers. TieNet: Text-Image Embedding Network for Common Thorax Disease Classification and Reporting in Chest X-Rays. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. 2018.
24. J. Zhao, M. Li, W. Shi, Y. Miao, Z. Jiang *et al.* A deep learning method for classification of chest X-ray images. *Journal of Physics: Conference Series*. 2021;1848(1).
25. T. K. K. Ho, and J. Gwak. Utilizing Knowledge Distillation in Deep Learning for Classification of Chest X-Ray Abnormalities. *IEEE Access*. 2020;8;160749–160761.
26. Y. Xiao, M. Lu, and Z. Fu. Covered Face Recognition Based on Deep Convolution Generative Adversarial Networks. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. 2020.

27. F. Ahmad, M. U. Ghani Khan, and K. Javed. Deep learning model for distinguishing novel coronavirus from other chest related infections in X-ray images. *Comput Biol Med.* 2021; 134;104401.
28. K. Simonyan, and A. Zisserman. Very deep convolutional networks for large-scale image recognition. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings.* 2015.
29. C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed *et al.* Going deeper with convolutions. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition.* 2015.
30. S. Ioffe, and C. Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *32nd International Conference on Machine Learning, ICML 2015.*
31. K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition.* 2017.
32. Y. Yang, Z. Zhong, T. Shen, and Z. Lin. Convolutional Neural Networks with Alternately Updated Clique. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition.* 2018.
33. J. Hu, L. Shen, S. Albanie, G. Sun, and E. Wu. Squeeze-and-Excitation Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 2020;42(8);2011-2023.
34. T. Salimans, I. Goodfellow, W. Zaremba, V. Cheung, A. Radford *et al.* Improved techniques for training GANs. *Advances in Neural Information Processing Systems.* 2016.
35. M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, and S. Hochreiter. GANs trained by a two time-scale update rule converge to a local Nash equilibrium. *Advances in Neural Information Processing Systems.* 2017.
36. C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna. Rethinking the Inception Architecture for Computer Vision. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition.* 2016.
37. L. Fei-Fei, J. Deng, and K. Li. ImageNet: Constructing a large-scale image database. *Journal of Vision.* 2009; 9(8).
38. N. L. Ramo, K. L. Troyer, and C. M. Puttlitz. Comparing Predictive Accuracy and Computational Costs for Viscoelastic Modeling of Spinal Cord Tissues. *Journal of Biomechanical Engineering.* 2019; 141(5).
39. P. D.M.W. EVALUATION: FROM PRECISION, RECALL AND F-MEASURE TO ROC, INFORMEDNESS, MARKEDNESS CORRELATION. *Journal of Machine Learning Technologies.* 2011; 2(1).
40. G. Zeng. On the confusion matrix in credit scoring and its analytical properties," *Communications in Statistics - Theory and Methods.* 2020;49(9);2080-2093.
41. T. Fawcett. An introduction to ROC analysis. *Pattern Recognition Letters.* 2006;27(8).
42. C. X. Ling, J. Huang, and H. Zhang. AUC: A better measure than accuracy in comparing learning algorithms. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics).* 2003.

# Figures

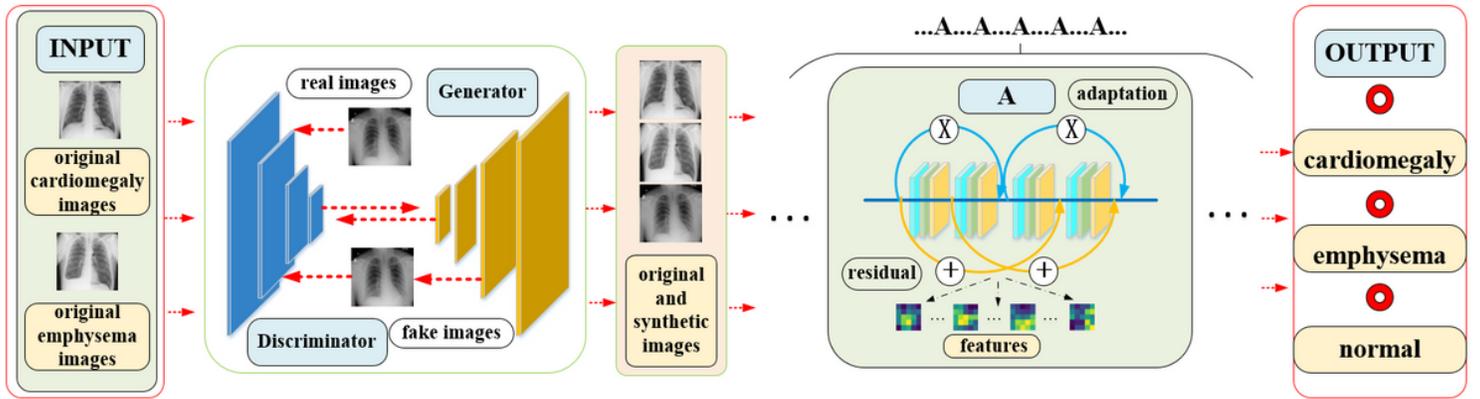


Figure 1

The flowchart of using DCGAN synthesis image and ACRnet recognition chest X-ray images

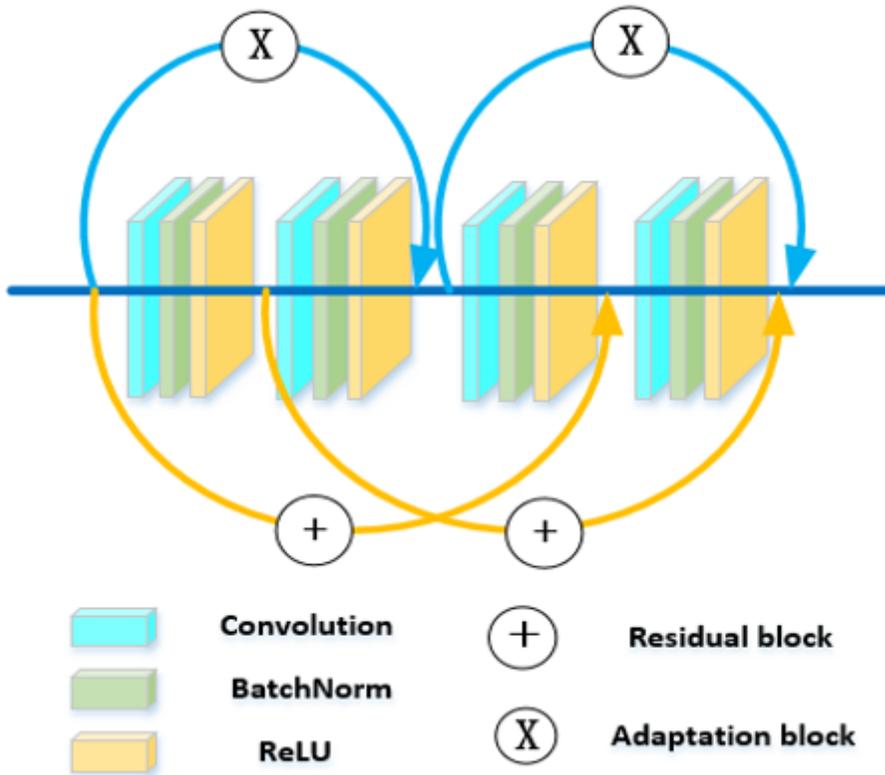
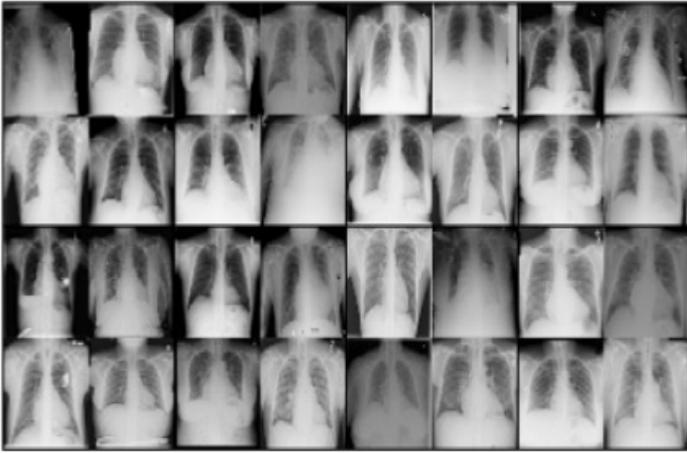
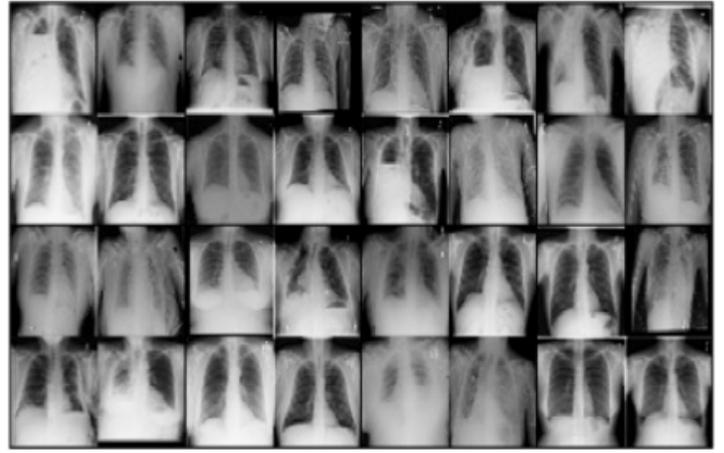


Figure 2

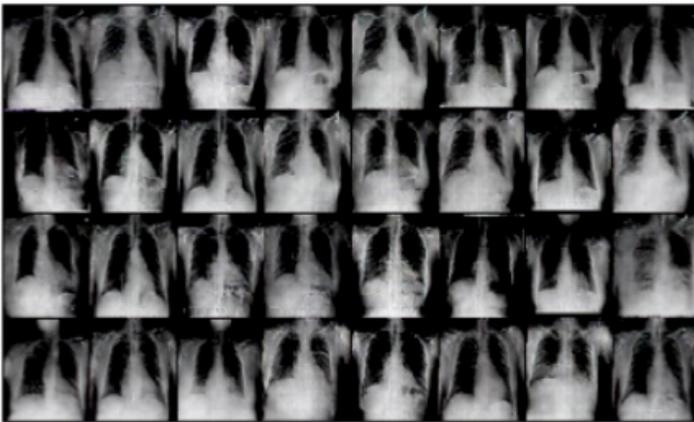
Adaptive cross-transfer residual block in ACRnet.



(a)



(b)



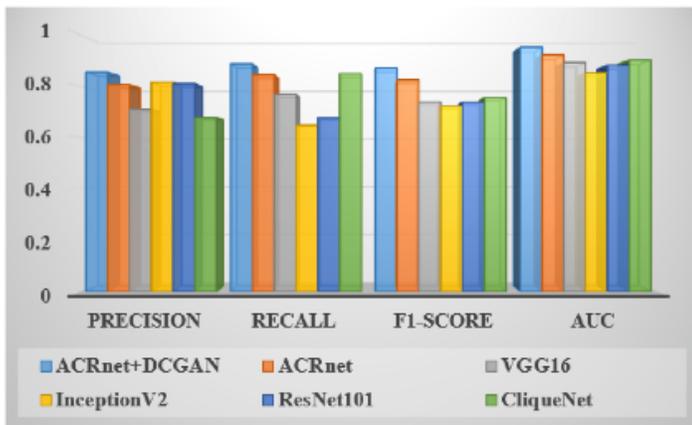
(c)



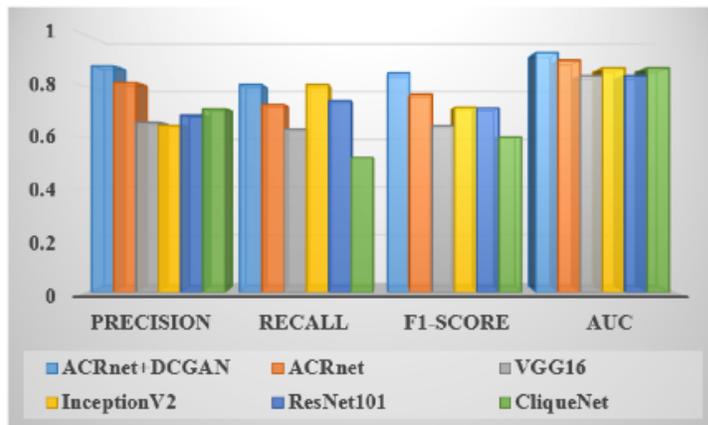
(d)

### Figure 3

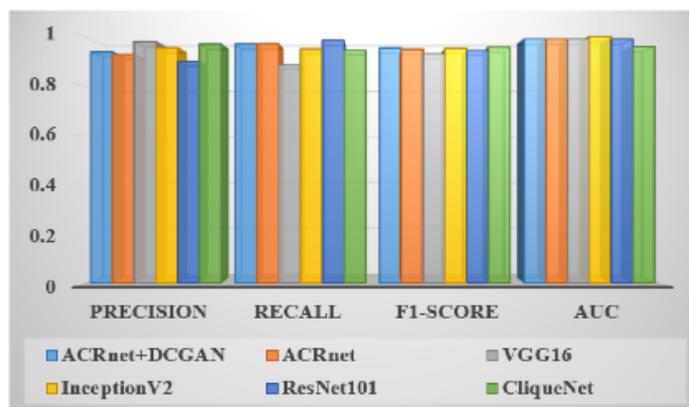
Original and the synthetic images. (a) original cardiomegaly chest X-ray images, (b) original emphysema chest X-ray images, (c) synthetic cardiomegaly chest X-ray images, (d) synthetic emphysema chest X-ray images.



( a )



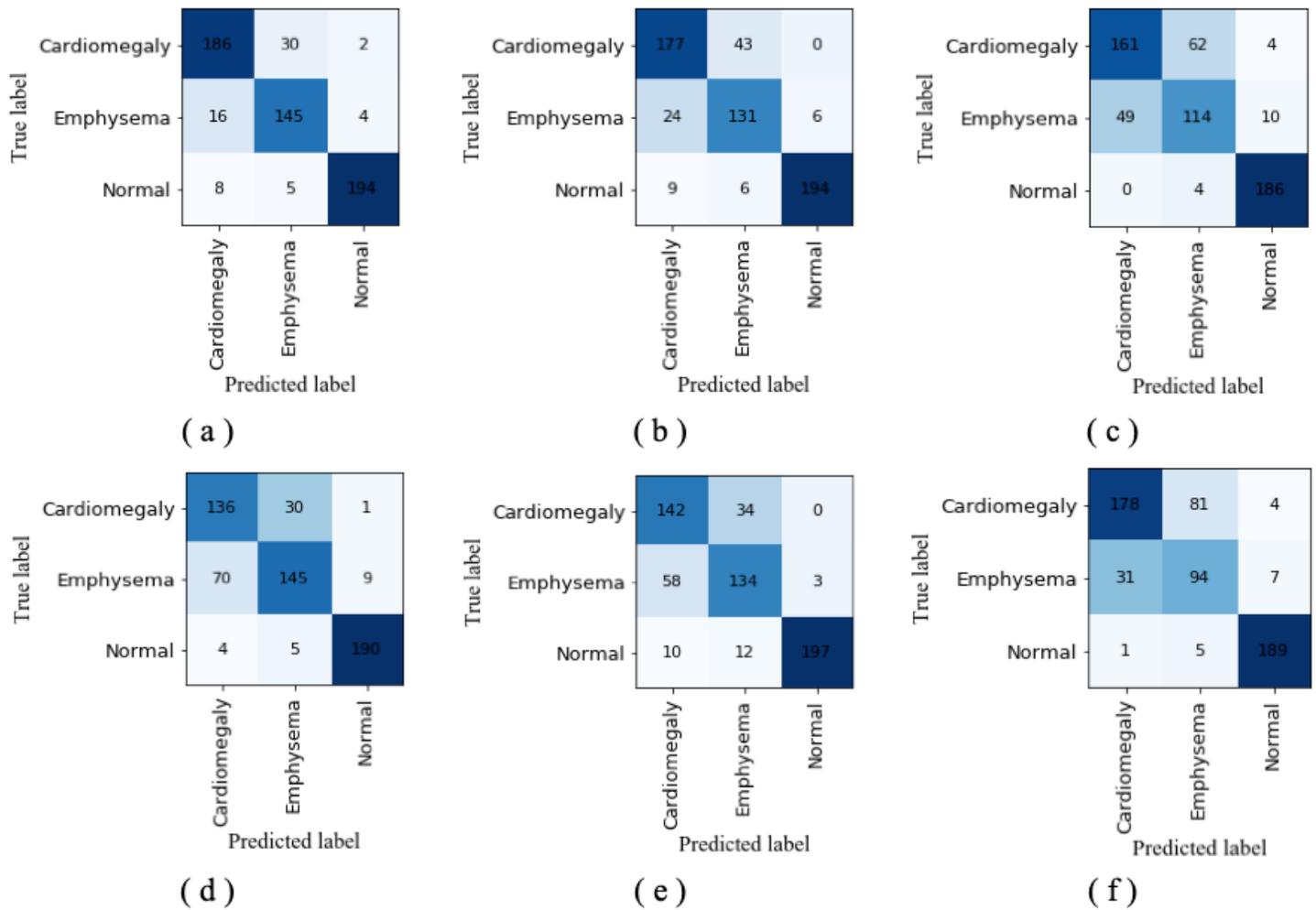
( b )



( c )

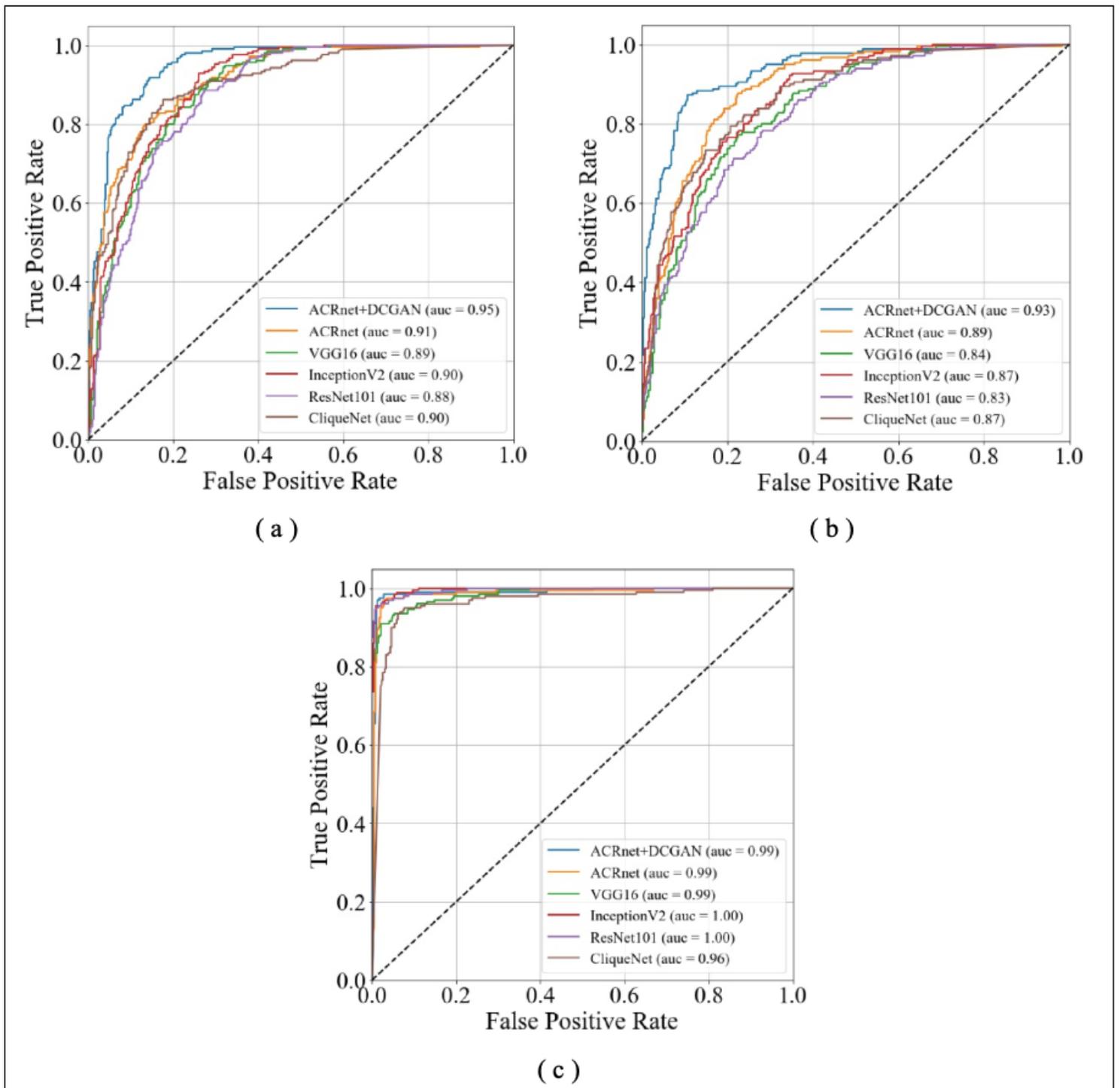
**Figure 4**

Mean precision, recall, F1 - score and AUC values obtained by different models in detecting 3 types of chest X-ray images using 5-fold cross validation. ( a ) Cardiomegaly, ( b ) emphysema, ( c ) normal.



**Figure 5**

Confusion matrixes obtained by neural networks. ( a ) ACRnet trained in dataset which expanded by DCGAN, ( b ) ACRnet trained in original dataset, ( c ) VGG16 trained in original dataset, ( d ) InceptionV2 trained in original dataset, ( e ) ResNet101 trained in original dataset, ( f ) CliqueNet trained in original dataset.



**Figure 6**

ROC curves obtained by neural networks. ( a ) cardiomegaly, ( b ) emphysema, ( c ) normal.