World Scientific Research Journal Volume 8 Issue 5, 2022
ISSN: 2472-3703 DOI: 10.6911/WSRJ.202205_8(5).0036

Semi-supervised X-Ray Image Classification Algorithm Based on Contrastive Learning and Generative Adversarial Networks

Xiancai Qin[1, a, *], Kun Liu[1, b]

[1]Collegel of Information Engineering, Shanghai Maritime University, Shanghai, China

[a]202030310189@stu.shmtu.edu.cn, [b]kunliu@shmtu.edu.cn

Abstract

A contrastive learning (CL) and generative adversarial network (GAN) -based method is proposed in this study to address the problem of the scarcity of labeled Chest X-Ray images. First, we adopted the Momentum Contrast (MoCo) combined with Convolutional Block Attention Module (CBAM) to extract the high-level semantic information from large amount of unlabeled data, and the pretrained CB-ResNet50 feature extraction network was obtained. Second, the artificial images which were similar to the real X-Ray images were generated by GAN, so as to expand the training data. Third, an external classifier (EC) based on the pretrained CB-ResNet50 was constructed by using transfer learning, which was attached to the GAN's generator(EC-GAN), rather than sharing an architecture with the discriminator. Finally, we used the generated fake data and limited real labeled data for semi-supervised training of EC, in order to obtain the final lung disease diagnosis network. Experimental results show that the proposed algorithm, hence named CLEC-GAN, can achieve the same classification performance as the fully supervised method with only a small number of labeled samples, and is better than the traditional semi-supervised classification methods.

Keywords

Image Classification; Momentum Contrast; Generative Adversarial Networks; Semi-supervised Learning.

1. INTRODUCTION

Computer Aided Diagnosis (CAD) [1] can be used to interpret medical images and retrieve valuable information to help doctors accurately diagnose and timely prevent different diseases. Chest X-Ray (CXR) is a commonly used medical imaging, which plays an important role in the diagnosis of lung diseases. In 2018, S.Rajaraman et al. [2] generated a heat map and superimposed it on the input of CNN model to classify pediatric CXRs. In 2019, A.K.Jaiswal et al.[3] used mask-RCN-based deep neural network to perform pixel-level segmentation, recognition and location of pneumonia in CXRs. In 2020, B. Elgen et al. [4] used mRMR feature dimension reduction method to extract and combine features, thus improving the model accuracy of decision tree, KNN and SVM classifier. In 2020, G.Liang et al. [5] proposed an automatic diagnosis algorithm for pediatric pneumonia using end-to-end learning, which initialized the model using transfer learning to prevent the influence of negative noises. Many scholars have carried out X-ray images analysis based on deep learning, which has high quality feature extraction capability and can process the original image. However, the performance of this method depends on a large amount of data to a great extent. The medical data collection process is complex and expensive, and medical images are rarely shared publicly due to privacy restrictions and ethical requirements. At the same time, medical images are quite difficult to label and require huge manpower. Therefore, it has become a research hotspot to improve the accuracy of classification and detection in the case of fewer labeled samples.

In the field of few-shot learning [6], the self-supervised learning (SSL) [7] method can be used to learn effective feature representations from unlabeled data without any manual label information. Contrastive learning [8] is an important branch in the field of self-supervision. In view of the inherent modal diversity and data complexity of CXRs, contrastive learning can learn high-quality features with high differentiation and high reusability from high-dimensional data, and improve the robustness and generalization ability of algorithms. At the same time, in the medical field, GAN can generate artificial images that are highly similar to real images at lower loss, and alleviate the under-sampling problem of X-ray datasets.

Based on this, this paper proposes a semi-supervised classification algorithm for X-ray images based on contrastive learning and ecternal classifier-generative adversarial network, hence named CLEC-GAN. Firstly, Momentum Contrast (MoCo) [9] based on Convolutional Block Attention Module (CBAM) [10] is used to learn the high-dimensional features of CXRs, and the pretrained feature extraction network are obtained. Secondly, the weight of the pretrained network is transferred to the classification task of real labeled CXRs and artificial images generated by GAN. Then, we utilize EC-GAN [11] to separate the tasks of discrimination and classification, and simultaneously leverage generated images to assist classification. Compared with traditional supervised learning and semi-supervised learning, the CLEC-GAN model can achieve better classification performance when the labeled samples are limited.

2. RELATED WORK

2.1. Contrastive Learning

Self-supervised learning needs to design a pretext task to learn feature representation containing advanced semantic information from unlabeled data and transfer the pretrained model to downstream tasks, such as semantic segmentation, object detection, image classification, etc. Its performance can match or even exceed that of fully supervised learning. Currently popular self-supervised learning methods mainly includes generative methods represented by autoencoders and its variants [12-14] and contrastive methods represented by MoCo [9], SimCLR [15] and PIRL [16]. Compared with the former, the contrastive method focuses on obtaining abstract semantic information from the original data, so it is more suitable for visual perception and understanding tasks with high professional requirements such as X-Ray image classification. Contrastive learning aims to learn a semantic embedding space by comparing a large number of positive and negative samples, so that similar samples are close to each other in this space, while dissimilar samples are far apart.

2.2. GAN

Generative Adversarial Networks (GANs) [17] can generate very vivid pictures, images, and even video, It consists of a generator network and a discriminator network. The process of GAN is shown in Figure 1. Generator generates samples similar to training images and confuses discriminator to the greatest extent. Discriminator can distinguish synthetic images from real images and stimulate generator to produce more realistic images. This architecture is equivalent to the minimum-maximum problem in game theory, where discriminator is discarded once the network converges.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202205_8(5).0036

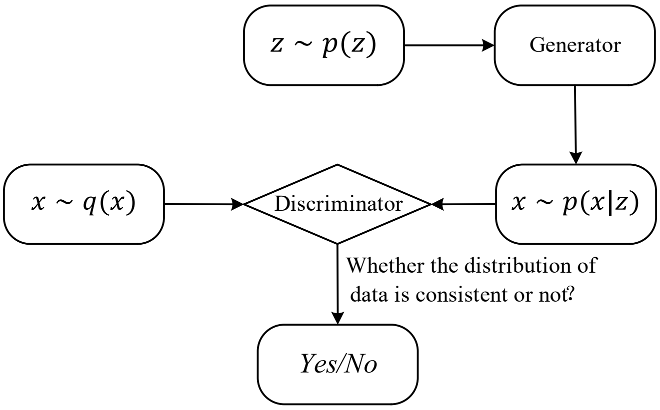


Figure 1. Flow chart of GAN

3. METHODS

The proposed CLEC-GAN consists of two basic modules: the contrastive self-supervised feature learning module and the EC-GAN semi-supervised image classification module. The contrastive learning module takes a large number of real unlabeled CXRs as input and minimizes the contrastive loss to obtain the pretrained feature extraction network. EC-GAN takes a small number of real labeled CXRs as input, generator outputs false images similar to real images, discriminator outputs the probability of belonging to real images, and external classifier (EC) outputs the normalized category probability to predict belonging to one of the 7 diseases and normal categories. Figure 2. shows the structure of the algorithm model. The classification performance of few-shot CXRs can be improved by combining momentum contrast feature learning based on a large amount of unlabeled data with semi-supervised learning based on generated data and a small amount of real labeled data.

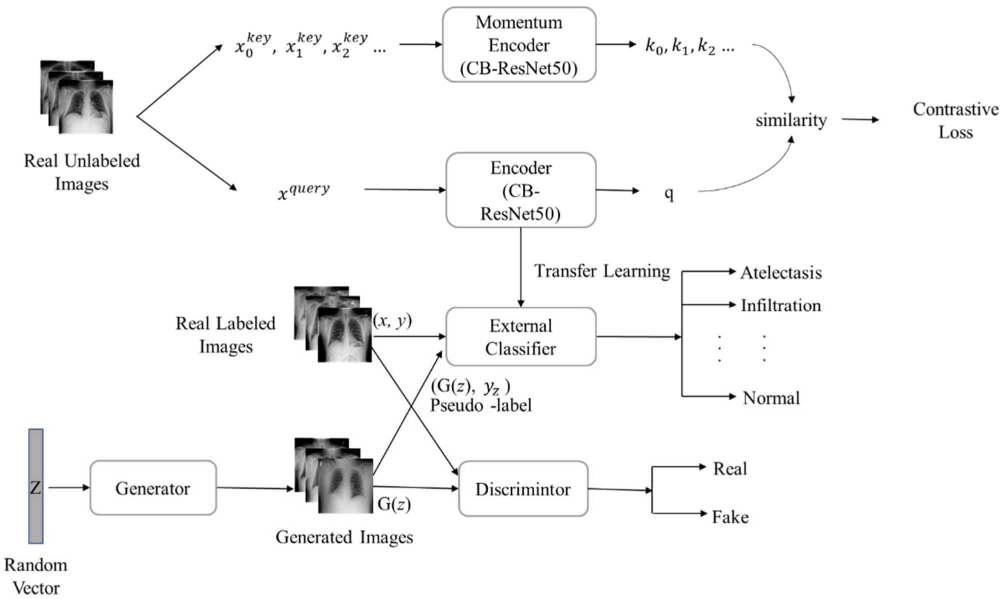


Figure 2. Model structure of CLEC-GAN

3.1. Lesion Feature Learning Based on MoCo

In this paper, MoCo proposed by He et al.[9] is used to train the feature extraction network on unlabeled X-ray images. MoCo regards contrastive learning as a dictionary query task and dynamically encodes keywords through the momentum encoder, expanding the dictionary

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202205_8(5).0036

capacity and maintaining the consistency of the dictionary. Random sampling of mini-batch images from unlabeled X-ray images is used as a dynamic dictionary. An image was randomly selected from the dictionary for random rotation (20°) and random horizontal inversion to obtain two enhanced images x^{query} and x_{+}^{key}. x^{query} is input into Encoder f_q as an anchor point to obtain query feature q. x_{+}^{key} is input into Momentum Encoder f_k as key to get feature k_{+}, x^{query} and x_{+}^{key} are matching positive samples, while other X-Ray images in the dictionary are input into Momentum Encoder f_k as key to form negative samples, resulting features are denoted as k_i. The object of contrastive learning is to find k_{+} that best matches q from the key queue $\{k_0, k_1, k_2 \dots\}$.

During the training process, InfoNCE is used as the contrastive loss to update the parameters of Encoder f_q, so as to maximize the similarity between q and a positive sample k_{+} and minimize the similarity between q and k other negative samples k_i, so that Encoder f_q has the ability to extract highly discriminative features from X-Ray images. The calculation formula of InfoNCE is as follows:

$$\mathcal{L}_q = -log\frac{exp(q*k_{+}/\tau)}{\sum_{i=0}^{K} exp(q*k_i/\tau)} \tag{1}$$

Where τ is a temperature hyper-parameter[18], q=$f_q(x^{query})$, k=$f_k(x^{key})$.

Instead of updating the parameters of Momentum Encoder f_k by back propagation, momentum update is adopted to maintain the consistency of dynamic dictionary. The updating formula is:

$$\theta_k \leftarrow m\theta_k + (1-m)\theta_q \tag{2}$$

where θ_k is the parameter of Momentum Encoder f_k, θ_q is the parameter of Encoder f_q, and $m \in [0,1)$ is the momentum coefficient. This momentum update way makes the change of f_k more stable, and maintains the consistency of the representation of each key in the queue, thus it is more conducive to the learning of focal features.

ResNet50[19], which is not pretrained on ImageNet, is used as the backbone network of Encoder and Moment Encoder in MoCo. ResNet50 contains two basic modules: Convolution Block and Identity Block. In order to enhance the long-range context-dependent representation of image data and further improve the feature extraction capability of the network, the Convolutional Block Attention Module (CBAM)[10] is added into the residual Block of ResNet50. Figure 3. shows the network structure. In the figure, ×2, ×3, and ×5 indicate stacking Identity blocks twice, three times, and five times respectively.

CBAM combines channel and spatial attention mechanism module. The feature graph output from the convolution layer is sent into a channel attention module to obtain the weighted result, and then through a spatial attention module to finally carry out adaptive feature optimization through weighting. CBAM is a lightweight general-purpose module, so its overhead can be almost ignored during training. At the early stage of training, it was found that the accuracy of the validation set was significantly different from that of the validation set, so a Dropout [20] layer with a ratio of 0.3 was added to the network to effectively alleviate overfitting.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202205_8(5).0036

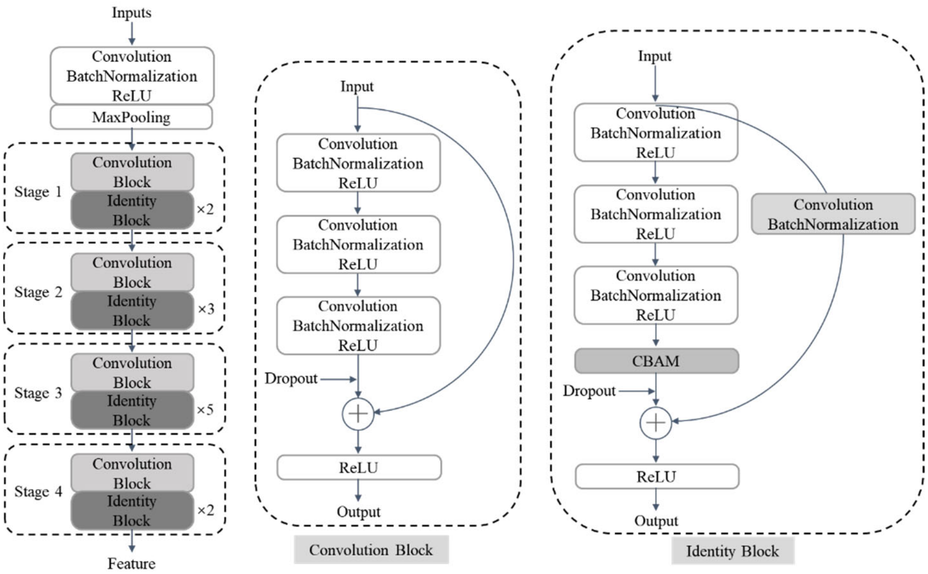


Figure 3. The Architecture of CB-ResNet50

3.2. X-ray Image Semi-supervised Classification Based on EC-GAN

In the popular GAN classification method, discriminator has to undertake two incompatible tasks during training, namely, identifying false samples and predicting sample labels, When the discriminator has two incompatible convergence points, the training process will fall into local optimum. Therefore, EC-GAN proposed by Haque [11] is adopted in this paper. Instead of using the shared discriminator model, an external classifier (EC) is attached to the generator, making the discriminator only undertake the task of distinguishing real and false images, while the external classifier trains the images generated by the generator and the real labeled images at the same time.

Discriminator and Generator in EC-GAN introduce DCGAN [21] which has strided convolution and deconvolution. Figure 4. shows the basic structure of DCGAN, Generator maps a 128-dimensional noise vector Z to a convolution layer containing 1024 feature maps, and converts Z into a 256pixel×256pixel CXR image through 6 fully connected deconvolution layers. ReLU activation function and batch normalization are applied at each fully connected layer, and tanh activation function is applied at the output layer. Discriminator contains six fully connected convolution layers, and LeakyReLU activation function and batch normalization are applied to each of the fully connected layers, with Sigmod activation function applied to the output layer.

Formula (3) gives the antagonistic components between Discriminator and Generator:

$$\min_{G}\max_{D}\mathcal{L}(D,G) = \mathop{\mathbb{E}}_{\boldsymbol{x}\sim\boldsymbol{p}_{data}}(\boldsymbol{x})[log\, D(\boldsymbol{x})] + \mathop{\mathbb{E}}_{\boldsymbol{z}\sim\boldsymbol{p}_{z}(\boldsymbol{z})}[log(1 - D(G(\boldsymbol{z})))] \tag{3}$$

Generator tries to minimize the above function, while Discriminator tries to maximize it. Where, $D(\mathbf{x})$ is the output of discriminator, and represents the probability that the input is a real sample. $G(\mathbf{z})$ is the output of the generator when noise Z is fed into it, $D(G(\mathbf{z}))$ is the output probability of the discriminator's ability to judge false image is true, and E is the expected value.

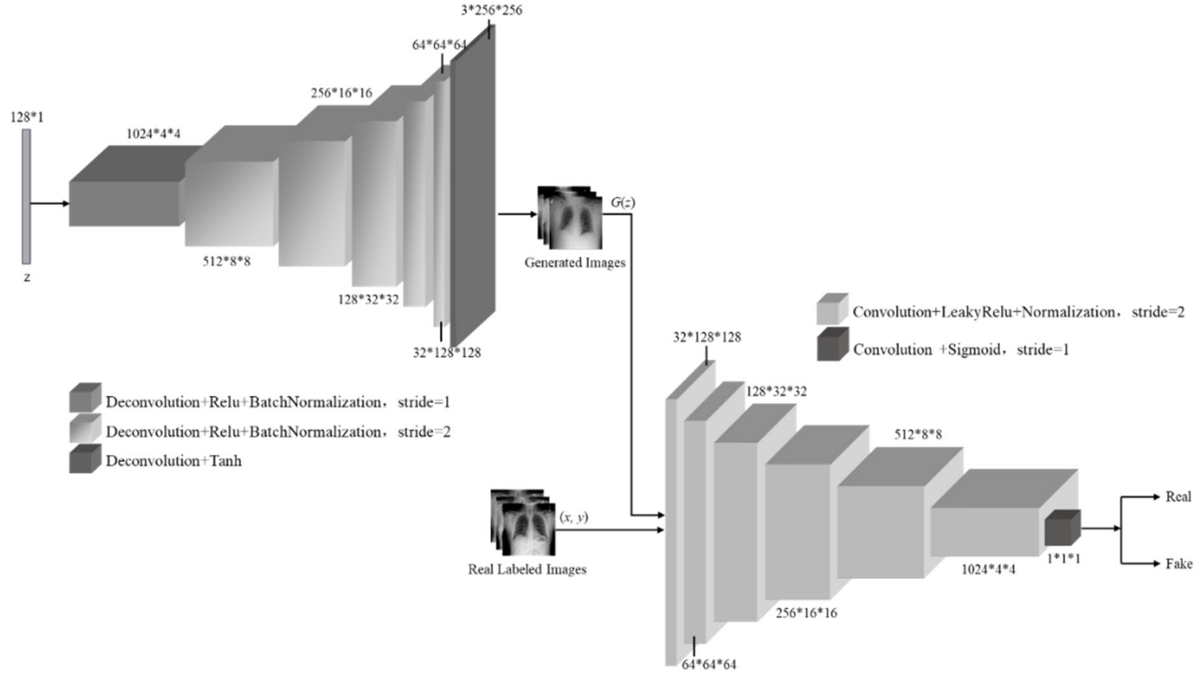


Figure 4. Basic structure of DCGAN

The external classifier of EC-GAN uses the CB-ResNet50 pretrained by MoCo in 3.1 for model initialization to improve image classification performance. The pseudo-label method based on confidence proposed by D.H.Lee[22] is adopted, which trains a small amount of real labeled data to create manual tags for the generated unlabelled images. Only when the model prediction data category with a high degree of confidence, or probability is higher than a certain threshold t, the generated image and label to be preserved.

The loss of external classifier is defined as:

$$L_C(x,y,z) = CE(C(x),y) + \lambda CE(C(G(z)), argmax(C(G(z))) > t) \tag{4}$$

The first part is the cross entropy loss of the input real labeled data, the second part is the cross entropy loss between the generated data and the corresponding hypothetical label, CE is the cross entropy loss, C is the classifier, λ is the unsupervised loss weight, and t is the pseudo-label threshold.

4. EXPERIMENTS

4.1. Data and Preprocessing

This study used the ChestX-Ray14 dataset published by Wang et al.[23], which contained 112,120 forward-looking X-ray images from 30,805 patients, as well as image labels for 14 categories of diseases mined from relevant radiological reports using Natural Language Processing (NLP). We select 39,700 Chest X-ray images as experimental data. There are 8 kinds of sample data, including five common pulmonary diseases (Atelectasis, Infiltration, Effusion, Nodule, Pneumothorax), two rare diseases (Fibrosis and Edema) and Normal. Secondly, in order to ensure the consistent input size of the image, the image is uniformly scaled to 256pixel×256pixel, and the pixels in the image are normalized between 0 and 1. Normalization can make the image converge quickly, and histogram equalization can be used to defog the image. Finally, the data is divided into training set, validation set and test set in a ratio of 7:1:2.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202205_8(5).0036

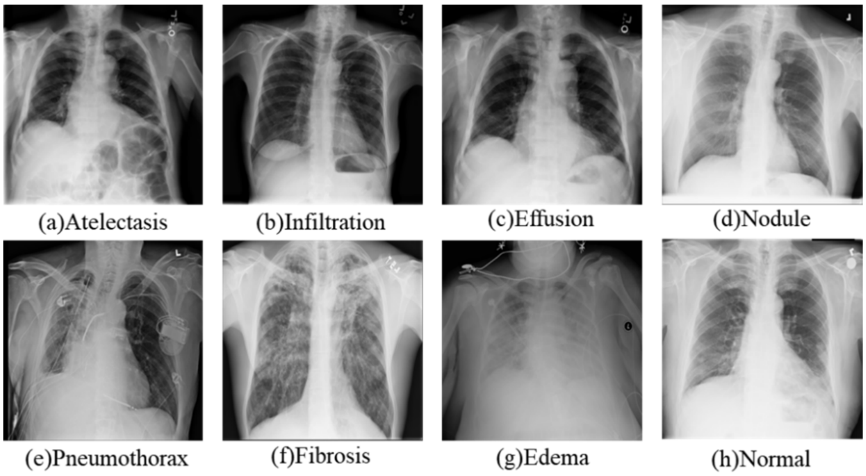


Figure 5. Training sample images

4.2. Algorithm Training Process

The flow of Chest X-ray image classification algorithm based on contrastive learning and external classifier-generative adversation network (CLEC-GAN) is divided into two stages:

The first stage is pretraining CB-ResNet50 based on MoCo, and minibatch is 16. Parameter update of Momentum Encoder f_k follows the momentum update method of Formula (1), where momentum coefficient m=0.999. Encoder f_q relies on InfoNCE of Formula (2) and uses stochastic gradient descent (SGD)[24] to update parameters. Weight attenuation of SGD is 0.0001, momentum of SGD is 0.9, and learning rate is 3×10-5. The pretrained CE-ResNet50 feature extraction network is obtained when the comparison loss L_q converges.

In the second stage, EC-GAN is trained, and the loss of external classifier is shown in Formula (4). The loss of Discriminator can be expressed as:

$$L_D(x,z) = BCE(D(x),1) + BCE(D(G(z)),0) \tag{5}$$

Accordingly, the loss of Generator is expressed as:

$$L_G(z) = BCE(D(G(z)),1) \tag{6}$$

Where, BCE is binary cross entropy, x is the input labeled data, z is a random vector, real label is "1", false label is "0". The training purpose of Discriminator is to maximize the probability of input correct classification, that is, to maximize $L_D(x,z)$. While Generator wants to minimize $L_G(z)$ to produce a better fake image.

Both GAN and External Classifier are trained during model training, and the training process is as follows:

i. Random vector z is sampled randomly in gaussian distribution and input z to Generator to obtain simulation image $G(z)$.

ii. The real labeled image x and simulation image $G(z)$ are input into identify Discriminator by batch, and the normalized probability values $D(x)$ and $D(G(z))$ are output by sigmoid.

iii. Fixed the parameters of Generator, using $D(G(z))$ as the loss function. The Adam gradient descent [25] with a learningrate of 0.0003 and momentum of 0.5 are used to adjust the parameters of Discriminator.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202205_8(5).0036

iv. Migrate CB-ResNet50 feature extraction network pretrained on a large amount of unlabeled data to External Classifier. Freeze the low-level network parameters, and add a fullly connection layer and a Softmax layer with 8 nodes at the end of the network. Send the generatored manual data and a small sample of labeled data into the fine-tuned network to predict 8 types of lung diseases including normal ones.

v. When the real labeled image x and the simulated image $G(z)$ are imported into Discriminator by batch, they are also imported into External Classifier. The real labeled images x are divided into 8 categories, and the unlabeled G(z) are labeled with pseudo-label threshold t and counterweight λ. The loss function of External Classifier is $L_C(x, y, z)$, and λ=0.1, t=0.7. Adam gradient descent is used to adjust the parameters of the External Classifier.

vi. The parameter of Discriminator is fixed, and the feedback information of Discriminator is used to modify the parameter of Generator.

vii. Repeat i-vi until iterations are complete.

4.3. Performance Evaluation of GAN

The training of GAN is a process of game between discriminator and generator. The final goal is to make the network reach Nash equilibrium by gradient descent. Figure 6. describes the evolution of Generator and Discriminator losses of EC-GAN in the training process. It can be seen that with the increase of training times of the model, the two networks contend with each other and restrict each other, so there is a large oscillation phenomenon in the figure. But on the whole the loss of Discriminator and Generator decreases and the amplitude of oscillation decreases.

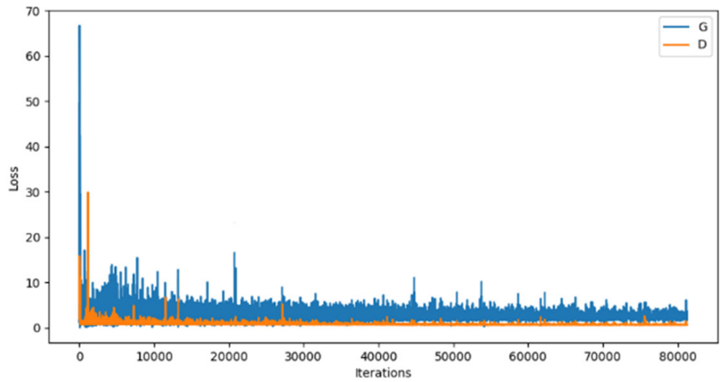
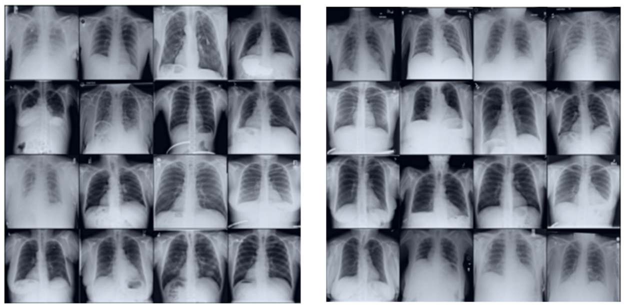


Figure 6. The losses of Generator and Discriminator in EC-GAN

The images generated by Generator are shown in Figure 7.(b), and compared with the original images shown in Figure 7. (a), it is obvious that the generated images have clear outline and reasonable structure.



(a)Original images (b)Generated images

Figure 7. Generated images of EC-GAN and original images

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202205_8(5).0036

Two indicators, Inception Score (IS)[26] and Fréchet Inception Distance (FID)[27], are introduced for quantitative evaluation. IS calculates the KL difference between conditional class distribution and marginal class distribution. Higher Inception score indicates better image quality. The FID calculates the Wasserstein-2 distance between the generated images and the real images, and the lower the FID score, the more similar the two groups of images are. EC-GAN was compared with variational autoencoders (VAE) and DCGAN. As shown in Table 1., the IS of the images generated by EC-GAN is closest to the original images, and its FID was the minimum, indicating that EC-GAN fitted the original dataset well, and generated images with good diversity and high resolution.

Table 1. FID scores and Inception scores for samples generated by various models

	Real Data	EC-GAN	VAE	DCGAN
FID	0.0	10.84	16.48	12.55
IS	4.43	3.31	2.76	3.18

4.4. Performance Evaluation of Classification

In order to evaluate the classification performance of CLEC-GAN when labeled sample scarcity, the GAN part was completely ignored, and the remaining external classifier (EC) was retained and used to distinguish between normal images and 7 different lung diseases. CLEC-GAN based on CB-ResNet50 is compared with CLEC-GAN based on original ResNet50, supervised learning convolutional neural network (CNN)[28], semi-supervised generative adversace network (SGAN)[29] and semi-supervised support vector machine (S3VM)[30] in experiments. The label fractions of training sets are 5%, 10, 15%, 20% and 25%. Table 2. shows the classification accuracy of five algorithms under different numbers of labeled samples.

Table 2. Classification accuracy under different number of labeled samples

Fraction of labeled data	Accuracy/%				
	CLEC-GAN (CB-ResNet50)	CLEC-GAN (ResNet50)	Supervied CNN	SGAN	S3VM
5%	75.73±4.1	74.15±3.7	63.21±4.2	68.39±3.4	66.79±4.3
10%	84.14±2.5	82.07±3.5	71.63±**5. 2**	75.20±4.6	70.23±2.1
15%	88.32±3.2	86.03±4.8	77.81±4.6	79.63±3.7	75.05±5.9
20%	91.35±2.9	89.41±2.6	81.14±3.8	83.16±2.4	79.34±3.2
25%	92.71±2.1	90.83±1.9	82.26±2.7	84.47±3.1	81.06±3.5

As shown in Figure 8. for the classification accuracy of each algorithm under different number of labeled samples trend chart, it can be seen that the less use of labeled data, the greater the performance improvement of semi-supervised learning, when only 5% labeled are used for each type, the accuracy of CLEC - GAN can reach 75.73%, and supervised learning CNN takes 15% labeled samples to achieve the same accuracy. With the increase of labeled data, the self-supervised performance gain gradually decreases, and CLEC-GAN is only 3% higher than SGAN when 25% labeled samples are used.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202205_8(5).0036

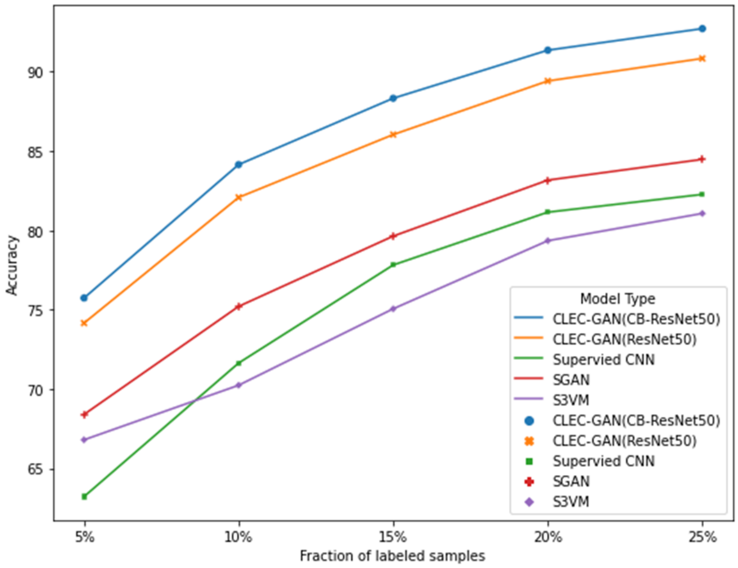


Figure 8. Trend of classification accuracy under different fractions of labeled samples

5. CONCLUSION

In this paper, we started from the lack of available labeled X-Ray images, and propsed the Chest X-Ray image classification algorithm based on contrastive learning and GAN. The main contributions are as follows: (1) CB-ResNet50 was pretrained with momentum contrast on a large number of unlabeled samples to extract high-level semantic information. (2) In order to further alleviate the problem of small sample dataset, the image synthesis algorithm based on GANs is used to generate X-Ray images that are highly similar to the original images in shape and texture. (3) The EC-GAN network structure was adopted to separate the identification and classification tasks, avoiding the degradation of the overall performance of the model because the discriminator shared two task systems. The results of ablation experiments and comparison experiments show that the proposed algorithm can still perform well under the condition of limited labeled samples. In the future, due to the difficulty in achieving Nash equilibrium in GANs and the uneven distribution of data categories, further research will be conducted to improve the accuracy of lung disease diagnosis under few labeled samples.

REFERENCES

[1] C.M. Chen, Y.H. Chou, N. Tagawa, et al. Computer-Aided Detection and Diagnosis in Medical Imaging, Computational & Mathematical Methods in Medicine, Vol.2013(2013).

[2] S. Rajaraman, S. CANDEMIR, I. KIM, et al. Visualization and Interpretation of ConVolutional Neural Network Predictions in Detecting Pneumonia in Pediatric Chest Radiographs, Applied Sciences, Vol.8, 10 (2018).

[3] A.K. Jaiswal, P. Tiwari, S. Kumar, et al. Identifying pneumonia in chest X-rays: A deep learning approach, Measurement, Vol. 145 (2019), p.511-518.

[4] M. Toaar, B. Ergen and Z. Cmert: A deep feature learning model for pneumonia detection applying a combination of mrmr feature selection and machine learning models, IRBM, Vol.41 (2020), p.212-222.

[5] G. Liang, L. Zheng: A transfer learning method with deep residual network for pediatric pneumonia diagnosis, Computer Methods and Programs in Biomedicine, Vol.187 (2020).

[6] Y. Wang, Q. Yao, J.T. Kwok, et al. Generalizing from a Few Examples: A Survey on Few-shot Learning, ACM Computing Surveys, Vol.53 (2020).

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202205_8(5).0036

[7] I. Misra, V. Laurens: Self-Supervised Learning of Pretext-Invariant Representations, 2020, p.6706-6716.

[8] R. Hadsell, S. Chopra and Y. LeCun: Dimensionality reduction by learning an invariant mapping, 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (June 17-22, 2006), Vol.2, p.1735-1742.

[9] K. He, H. Fan, Y. Wu, et al. Momentum contrast for unsupervised visual representation learning, 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (June 13-19, 2020), p.9726-9735.

[10] S. Woo, J. Park, J.Y. Lee, et al. CBAM: ConVolutional Block Attention Module, 15th European Conference (Munich, Germany, September 8–14, 2018), Vol.11211, p.3-19.

[11] A. Haque: EC-GAN: Low-Sample Classification using Semi-Supervised Algorithms and GANs, 2020. https://arxiv.org/abs/2012.15864v3

[12] C. Doersch, A. Gupta , and A. A. Efros: Unsupervised visual representation learning by context prediction, IEEE International Conference on Computer Vision (Santiago, December 7-13, 2015), p.1422-1430.

[13] M. Noroozi, A. Vinjimoor, P. Favaro, et al. Boosting Self-Supervised Learning via Knowledge Transfer, 2018 IEEE Conference on Computer Vision and Pattern Recognition (Salt Lake City, UT, USA, June 18-23, 2018), p. 9359-9367

[14] A. Dosovitskiy, P. Fischer, J. T. Springenberg, et al. Discriminative Unsupervised Feature Learning with Exemplar ConVolutional Neural Networks, IEEE Transactions on Pattern Analysis and Machine Intelligence (September 1, 2016), Vol.38 (2016) No.9, p.1734-1747.

[15] T. Chen, S. Kornblith, M. Norouzi, et al. A simple framework for contrastive learning of visual representations, arXiv preprint arXiv:2002.05709, 2020a.

[16] M. Raghu, C. Zhang, J. Kleinberg, et al. Transfusion: Understanding transfer learning for medical imaging. arXiv preprint arXiv:1902.07208, 2019.

[17] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, et al. Generative Adversarial Networks, Advances in Neural Information Processing Systems, 2014, No.3, p.2672-2680.

[18] Z. Wu, Y. Xiong, S. X. Yu, et al. Unsupervised Feature Learning via Non-parametric Instance Discrimination, 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (June 18-23, 2018), 2018, p.3733-3742,

[19] K. He, X. Zhang, S. Ren, et al. Deep Residual Learning for Image Recognition, 2016 IEEE Conference on Computer Vision and Pattern Recognition (June 27-30, 2016), 2016, p.770-778.

[20] N. Frazier-Logue, S. J. Hanson, Dropout is a special case of the stochastic delta rule: faster and more accurate deep learning, 2018. https://arxiv.org/abs/1808.03578

[21] A. Radford, L. Metz , and S. Chintala: Unsupervised Representation Learning with Deep ConVolutional Generative Adversarial Networks, Computer Science, 2015.

[22] D. H. Lee: Pseudo-Label: The Simple and Efficient SemiSupervised Learning Method for Deep Neural Networks, In ICML Workshop on Challenges in RepresentationLearning, 2013.

[23] X. Wang, Y. Peng, L. Lu, et al. ChestX-ray8: Hospital-scale Chest X-ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common ThoraxDiseases, 2017 IEEE Conference on Computer Vision and Pattern Recognition (Honolulu, HI, USA, 2017), 2017, p.3462-3471

[24] T. Lancewicki, S. Kopru: Automatic and Simultaneous Adjustment of Learning Rate and Momentum for St-ochastic Gradient Descent, 2019. https://arxiv.org/abs/1908.07607

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202205_8(5).0036

[25] D. Kingma, J. Ba: Adam: A Method for Stochastic Optimization. Computer Science, 2014.

[26] H. Zhang, I. Goodfellow, D. Metaxas, et al. Self-Attention Generative Adversarial Networks, 2018. https://arxiv.org/abs/1805.08318v2

[27] T. Chen, M. Lucic, N. Houlsby, et al. On Self Modulation for Generative Adversarial Networks, 2018. https://arxiv.org/abs/1810.01365v2

[28] K. Simonyan, A. Zisserman: Very Deep ConVolutional Networks for Large-Scale Image Recognition, Computer Science, 2014.

[29] N. H. Trinh, D. O'Brien, Semi-Supervised Learning with Generative Adversarial Networks for Pathological Speech Classification, 2020 31st Irish Signals and Systems Conference (June 11-12, 2020), 2020, p.1-5.

[30] X. Yan, Y. Bai, S. C. Fang, et al. A proximal quadratic surface support vector machine for semi-supervised binary classification, Soft Computing, Vol.22 (2018) No.20, p.6905–6919.