Automatic Diagnosis of COVID-19 Medical Images based on Graph Attention Network

Yingxin Lai
School of Software
Jiangxi Agricultural University
Nanchang, China

Wenjuan Zhao
Newsun Biotechnology Research Institute
Chengdu New Sun Crop Science Co., Ltd.
Chengdu, China

Tingzhuo Chen
Newsun Biotechnology Research Institute
Chengdu New Sun Crop Science Co., Ltd.
Chengdu, China

Hongyu Jiang
School of Software
Jiangxi Agricultural University
Nanchang, China

Wenlong Yi*
School of Software
Jiangxi Agricultural University
Nanchang, China
*Corresponding author:
yiwenlong@mail.ru

Ke Liu
Newsun Biotechnology Research Institute
Chengdu New Sun Crop Science Co., Ltd.
Chengdu, China

***Аннотация*. Учитывая пандемию COVID-19 и ее высокоинфекционные характеристики, традиционная диагностика, основанная на медицинской визуализации, хотя и способна обнаруживать поражение легких в организме человека, не всегда оказывается эффективной. Поэтому особенно важно разработать набор точных и автоматических методов диагностики пневмонии с помощью технологий искусственного интеллекта, чтобы диагностировать и лечить пневмонию у пациентов можно было на ранней стадии. Это исследование впервые вводит DenseNet в структуру сверточной нейронной сети (Convolutional Neural Network, CNN), чтобы улучшить обмен характеристической информацией изображения легких в сверточных слоях и, таким образом, получить более точные характеристики изображения. При этом характеристики пневмонии быстро распознаются с помощью Graph Attention Network. Авторы используют набор данных рентгеновских снимков в Радиологическом обществе Северной Америки (Radiological Society of North America, RSNA). Для обнаружения пневмонии, выпущен Kaggle для обучения и проверки сети. Согласно экспериментальным результатам, точность диагностики COVID-19 и F-Score достигают 98 %. Данный метод предоставляет врачам компьютерной томографии сквозную технологию глубокого обучения для диагностики пневмонии.**

Ключевые слова: COVID-19; глубокое обучение; графовая нейронная сеть; диагностика

I. Introduction

Coronavirus has spread throughout the world since it was identified at the end of 2019 year. Coronavirus Disease 2019 (COVID-19) is a contagious disease that has heavy damage on the respiratory system, which can easily cause people with background diseases to become critical patients, posing a great threat to people's health. To implement strategies of early detection and early treatment, quickly recognizing patients with common pneumonia and patients with COVID-19 is particularly important in conducting special medical interventions for patients with COVID-19 to prevent them from getting worse.

Traditionally, lung images are detected effectively and simply through radiologist's judgement. When many patients pack into the hospital, problems such as doctor shortage or lowered testing efficiency will be witnessed. Computer Aided Diagnosis (CAD) has been gradually introduced into clinical diagnosis along with the evolvement of digital image processing technology and computer, which can provide doctors with medical image-aided opinions. A variety of methods have been proposed by many scholars in order to address the problem of automatic recognition of medical images of COVID-19. For instance, decision tree, as a frequently-used machine learning classification algorithm, can achieve recognition and classification through a tree structure. Yoo *et al.* proposed to classify COVID-19 patients using decision tree, since the algorithm is simple in principle and easy to implement [1]. But overfitting might be found when the amount of data is not big enough in the algorithm. Also, unfavorable classification effect can be also observed in the face of complex data. Sethy *et al.* proposed to classify COVID-19 patients, pneumonia patients and healthy people using Support Vector Machines (SVM) and obtained the accuracy presenting a gap of about 5 % with that of the traditional manual work [2]. It can effectively solve the problem of inseparable linearity through high-dimensional mapping with

This study was supported by Jiangxi Students' Platform for innovation and entrepreneurship training program of China (Grant No. S202110410087).

its accuracy higher than that of manually extraction. Excessively large model calculation, however, might lower the model efficiency when SVM is adopted in testing large data volume. With the progress of deep learning, people can train and predict larger datasets with ease. Specifically, the deep learning algorithm represented by CNN have been thrust into the limelight. Aslan *et al.* proposed a COVID-19 image recognition algorithm based on CNN [3]. The algorithm can improve the recognition accuracy by about 3% compared to traditional Bayesian and SVM algorithms taking advantages of automatic learning of CNN. But a limited recognition rate can be still found in CNN. Keles *et al.* enhanced network performance through introducing Resnet to increase the number of network layers and amplifying the dataset [4]. But the underlying network might lead to large quantities of parameters and rising calculation amount. Salih *et al.* introduced AlexNet and transfer learning through constructing a feature fusion layer, optimizing relevant parameters of the feature fusion layer, reducing the number of parameters of the original network, and improving the algorithm efficiency [5]. However, the algorithm failed to focus on the important area of the detected target. Zhou *et al.* put forward a deep learning model based on spatial and channel attention mechanisms to classify images of pneumonia patients [6]. Although the algorithm can effectively extract important features of lung images, segmentation results cannot be provided by the network. Saeedizadeh *et al*. proposed a Unet-based semantic segmentation method to achieve the effect of automatically segmenting regions [7]. But it presents unsatisfactory classification accuracy with limitations in the application to small datasets. Shi *et al.* reviewed the application of artificial intelligence technology in diagnosing COVID-19 images systematically upon analyzing 87 papers that cover all the processes of medical imaging analysis technology involved in COVID-19, including image acquisition, segmentation, and diagnosis [8]. Although favorable classification performance can be obtained by deepening the number of network layers, the amount of deep network parameters will be also increased by tens of or even hundreds of times, and the memory requirements will also be increased geometrically. Such higher demands in computer performance may cause difficulties to the deployment of client side. Training results using only shallow networks, however, cannot satisfy the requirements of practical applications.

In response to the above problems, DenseNet network in GAT is introduced to effectively improve CNN ability to recognize pneumonia images, providing accurate aided diagnosis opinions for early screening of patients suffering from COVID-19.

II. Related Work

The attention mechanism has been extensively applied in the fields of image segmentation and machine translation, which can give different weights according to the importance of the targeted area. Hu *et al.* elevated the accuracy of image recognition by means of modeling the interdependence between feature channels [9]. Also, Sutskever *et al.* constructed an attention mechanism at the word level, enabling it to express the importance of different words [10]. Introducing an attention mechanism into the graph structure is a novel approach, which has received widespread concerns for its small parameter amount and easy promotion. Veličković *et al.* first introduced the attention mechanism in the convolutional layer of the graph neural network by defining the graph attention layer [11]. In this way, a deep learning model of arbitrary structure can be formed through superimposing various attention layers. Different levels of importance are assigned by GAT to different nodes within the neighborhood during the convolution process. In particular, larger weights are given to important neighbor nodes. Since each node is operated separately, parallel operation can be performed between each node to significantly improve the calculation efficiency. Meanwhile, the trained model can be easily extended to varied structure graphs without relying on prior graph structure information. Therefore, GAT can effectively cope with the problem of inductive and transductive node classification. The model obtained SOTA results from 4 data sets. As can be seen from Fig. 1, various attention mechanism weights are adopted by the structure to learn three different attention models, and finally summarized for expression. Of which, a_{ij} is the correlation coefficient between vertices *i* and *j* in the graph neural network; h_i is the new feature after the vertex *i* integrating into the vertex information of its domain.

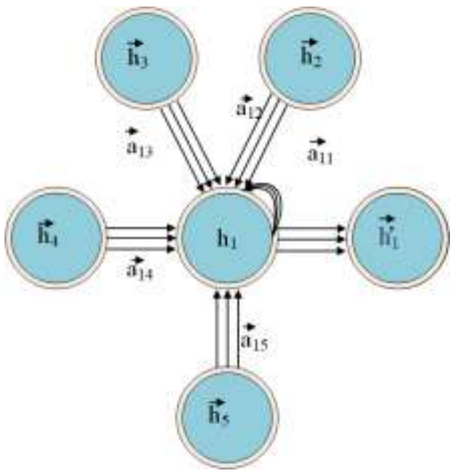

Fig. 1. Node Structure of the Graph Attention Network Mechanism

Adopting original convolutional neural network to process patient lung images might be greatly interfered by the accuracy of the entire network under the effect of local non-key regions although it can automatically classify pneumonia patients from COVID-19 patients. Burwinkel *et al.* proposed a CNNGAT model on the basis of introducing a GAT mechanism in the graph neural network to enhance the feature extracting capability of CNN [12]. As shown in Fig. 2, a DenseNet module is introduced into this model to improve the training efficiency of the model and reduce the occurrence of gradient disappearance and gradient explosion.

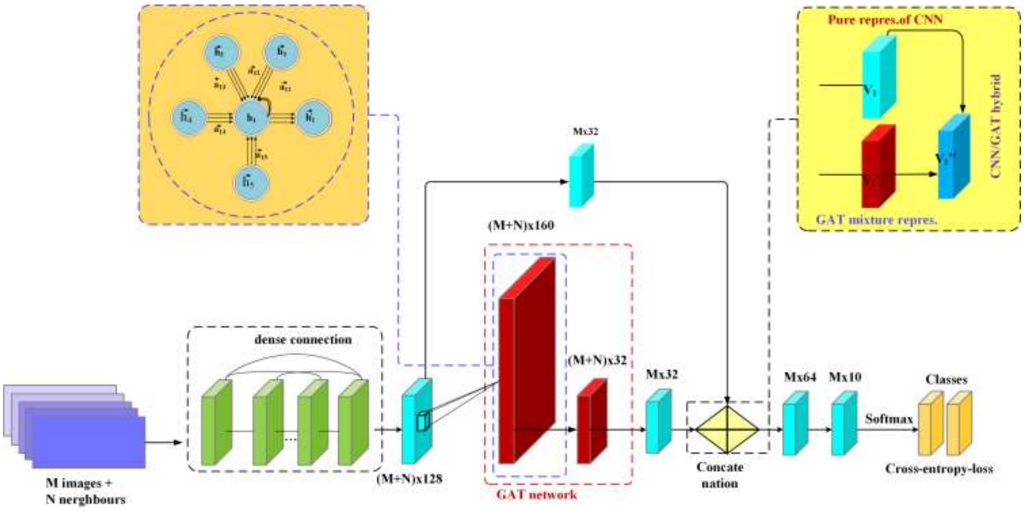

Fig. 2. Improved CNNGAT model framework

The objective function of the model is *F(X, G(V, E))*, where *X* is *M* input images; *G* is the graph structure of the network; and *V* and *E* are the vertices and edge of the graph, representing the characteristics of input image and the connection between various pixels, respectively. At first, the image feature matrix was extracted from the CNN module and assigned to the corresponding vertex V_i . Then, the weight information of different image regions was obtained through two GAT modules. Finally, classification prediction was performed on the image together with CNN information. Each node V_i determines the output value of each node processed by GAT in combination with domain characteristics and own characteristics. Different weight information should be given for various connections, which can represent the importance of different neighbor nodes. The weight of each node is calculated in (1).

$$a_{ij} = \frac{\exp\left(Leaky\,\mathrm{Re}\,LU\left(\vec{a}^{T}\left[W\,\vec{h}_i \,\|\, W\,\vec{h}_j \right]\right)\right)}{\sum_{k\in N_i} \exp\left(Leaky\,\mathrm{Re}\,LU\left(\vec{a}^{T}\left[W\,\vec{h}_i \,\|\, W\,\vec{h}_k \right]\right)\right)} \quad (1)$$

Where ‖ represents the connection operation, which is to splice the two vectors together; *T* is the vector transposition operation, each node will undergo a shared linear transformation, and *W* is the parameter weight matrix of the module, which can input low-order features. The parameters are converted to high-order, and the *LeakyReLU* activation function is used to filter the retained information of the node; and $\vec{h}_i$ are $\vec{h}_j$ the eigenvectors of nodes *i* and *j*; and $\vec{a}^T$ is the Attention Kernel. Results of multiple GATs are updated after calculating the weight of each node. The calculation process is shown in (2).

$$\vec{h}_i^{*} = \Big\|_{k=1}^{K} \sigma\left(\sum_{j\in N_i} a_{ij}^{k} W^{k}\,\vec{h}_j \right) \quad (2)$$

Among them, *K* is the number of attention networks for image feature extraction. After the network nodes a_{ij} are multiplied by the corresponding weights, the total information of each neighboring node is obtained through the accumulation operation. Finally, the activation function σ is used for nonlinear mapping to obtain the GAT module middle layer The output value. The output values obtained by GAT and CNN were superimposed as per the number of image channels, and fused into a new feature matrix with weight information. Next, the target image was classified using the convolutional layer and *Softmax* activation function. At the same time, the cross-entropy loss was utilized to train the network as its loss function.

III. Experiment and Discussion

A. Experimental environment

The hardware environment of the experiment is composed of GTX1080Ti graphics card, i5-9600k processor, and 16GBRAM memory, and the software environment involves windows10 64-bit operating system. Also, PyTorch deep learning framework was adopted to train and test the network, and the operating platform was Pycharm development tool. Also, the training set, validation set and test set are partitioned at the proportions of 80 %, 10 %, and 10 %, respectively. And the learning rate during network training is a fixed parameter, 0.0001, with 80 iterations. Adam optimizer is adopted with cross entropy loss as the loss function and Softmax function as the classification function.

B. Dataset and preprocessing

Data used in the experiment are obtained from the RSNA Pneumonia Detection Challenge published on the kaggle platform. The dataset is composed of medical images taken during routine clinical testing of patients at the American Medical Center [13]. A total of 5,956 chest X-ray photographs of human body in the image format of JPEG are included. The distribution of data is showed in Table 1.

TABLE I. DISTRIBUTION OF DATASET

Dataset type	Normal sample	COVID-19 samples	Bacterial pneumonia sample	Total amount
Training set	1349	1345	2538	5232
Test set	234	148	242	624
Total	1583	1493	2780	5956

X-ray images of the three categories are shown in the figure below. A normal chest X-ray image (on the left) shows that the lungs are clear in the image without obvious abnormal turbidity in the image. In the Bacterial Pneumonia image (in the middle), localized pneumonia consolidation can be found at the upper right lobe in this case, while Viral Pneumonia (COVID-19) image (on the right) presents interstitial characteristics at both lung lobes.

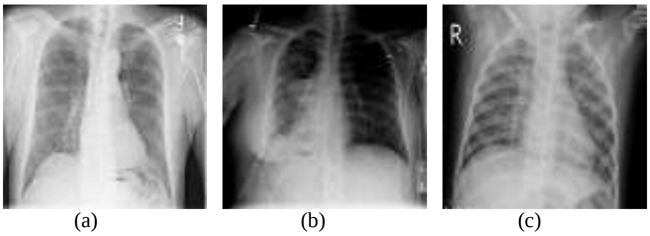

Fig. 3. Sample of chest X-ray photographs; a: Normal; b: Bacterial Pneumonia; (c) Viral Pneumonia

As chest X-ray photographs are important evidences for diagnosing pneumonia, the photograph quality has a great influence on the diagnosis results. Photographs in the dataset shall be firstly screened by experienced doctors, and then better-quality photographs obtained will be reviewed by experts for manually labelling the dataset. The dataset used in this paper was reviewed by medical experts from the American Medical Center, which is relatively reliable in data quality. X-ray photographs in the dataset were divided into training set and test set. The ratio of the number of the test set to the training set was about 1:9. Moreover, the original gray image was first adjusted to a 3-channel image to meet the input requirements of the CNN model. After that, the original image was cropped into five images, including top left corner *x1*, bottom left corner *x2*, bottom right corner *x3*, top right corner *x4*, and middle *x5*, and then unified the size as 224*224. A graph relationship is established according to the position relationship, as shown in Fig. 4.

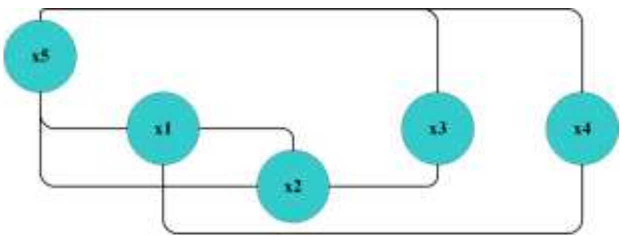

Fig. 4. Relationship after cropping

C. Evaluation criteria

In the field of medical image classification, accuracy and *F1*-score, in general, are adopted indicators for evaluating performance classification. More precisely, the accuracy is the ratio of the predicted situation of a COVID-19 patient to the actual result. And *F1* is the comprehensive ratio of accuracy and recall rate, both of which can express the predictive capability of the model. To be specific, the recall rate is the percentage of the predicted positive proportion in the total proportion, and the accuracy is the proportion of the positive samples in the classifier. *F1*-score is the harmonic mean that combines accuracy and recall with the minimum and maximum as 0 and 1, respectively. In other words, the higher the *F1* score, the lower the probability of missed diagnosis. Calculations of accuracy and F1 score are shown in (3)–(5), of which, *TP*, *FP*, *FN* and *TN* represent truth positive, false positive, false negative and truth negative, respectively.

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{3}$$

$$F1 = \frac{2 * precision * recall}{precision + recall} \tag{4}$$

$$recall = \frac{TP}{TP + FN} \tag{5}$$

D. Results and Analysis

Deep learning models of VGG16 [14], ResNet18 [15], and ResNet34 [15] were compared in the experiment in terms of the characteristics of image classification tasks. As can be witnessed in Table 2, this method has the performance in accuracy and *F1* score indicators.

TABLE II. RECOGNITION ACCURACY OF VARIOUS MODELS

Model name	Accuracy	F1 score
VGG16	87.680	86.541
ResNet18	88.312	87.723
ResNet34	90.021	91.656
CNNGAT	**98.164**	**98.032**

The change trend of the accuracy and *F1* score on the COVID-19 image test set using CNNGAT can be found in Fig. 5. Based on the figure, both indicators tend to be converged after reaching about 98 % before and after the 50th epoch.

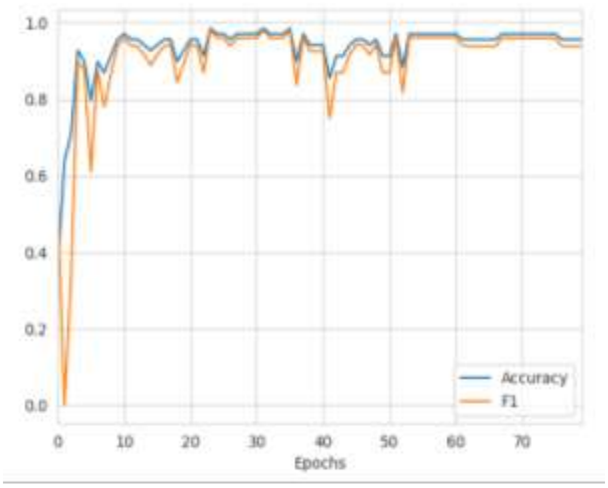

Fig. 5. Accuracy and *F1* score of verification sets

IV. Conclusion

Targeted at the classification characteristics of COVID-19 images, a graph neural network is adopted this paper to achieve automatic classification of COVID-19 images, which extracts image features with CNN and DenseNet as well as assign different weights to each pixel node of the image with the help of GAT network. The proposed method can enhance the feature extraction and classification capabilities of the network, while suppressing the gradient disappearance and gradient explosion of the network. According to the comparative test, the proposed method can provide an automated tool for determining medical images of pneumonia cases thanks to its favorable processing efficiency and accuracy, in comparison those common deep learning convolution methods.

References

[1] S. H. Yoo, H. Geng, T. L. Chiu, S. K. Yu, D. C. Cho, J. Heo, M. S. Choi, I. H. Choi, C. Cung Van, N. V. Nhung, B. J. Min, and H. Lee, “Deep Learning-Based Decision-Tree Classifier for COVID-19 Diagnosis From Chest X-ray Imaging,” // Frontiers in Medicine, vol. 7, Jul. 2020.

[2] P. K. Sethy, S. K. Behera, P. K. Ratha, and P. Biswas, “Detection of coronavirus Disease (COVID-19) based on Deep Features and Support Vector Machine,” // International Journal of Mathematical, Engineering and Management Sciences, vol. 5, no. 4, pp. 643–651, Aug. 2020.

[3] M. F. Aslan, M. F. Unlersen, K. Sabanci, and A. Durdu, “CNN-based transfer learning-BiLSTM network: A novel approach for COVID-19 infection detection,” // Applied Soft Computing, vol. 98, p. 106912, Jan. 2021.

[4] A. Keles, M. B. Keles, and A. Keles, “COV19-CNNet and COV19-ResNet: Diagnostic Inference Engines for Early Detection of COVID-19,” // Cognitive Computation, Jan. 2021.

[5] S. Q. Salih, H. K. Abdulla, Z. S. Ahmed, N. M. S. Surameery, and R. D. Rashid, “Modified AlexNet Convolution Neural Network For Covid-19 Detection Using Chest X-ray Images,” // Kurdistan Journal of Applied Research, pp. 119–130, Jun. 2020.

[6] T. Zhou, S. Canu, and S. Ruan, “Automatic COVID-19 CT segmentation using U-Net integrated spatial and channel attention mechanism,” // International Journal of Imaging Systems and Technology, vol. 31, no. 1, pp. 16-27, Nov. 2020.

[7] N. Saeedizadeh, S. Minaee, R. Kafieh, S. Yazdani, and M. Sonka, “COVID TV-Unet: Segmenting COVID-19 chest CT images using connectivity imposed Unet,” // Computer Methods and Programs in Biomedicine Update, vol. 1, p. 100007, 2021.

[8] F. Shi, J. Wang, J. Shi, Z. Wu, Q. Wang, Z. Tang, K. He, Y. Shi, and D. Shen, “Review of Artificial Intelligence Techniques in Imaging Data Acquisition, Segmentation, and Diagnosis for COVID-19,” // IEEE Reviews in Biomedical Engineering, vol. 14, pp. 4–15, 2021.

[9] J. Hu, L. Shen, and G. Sun, “Squeeze-and-Excitation Networks,” // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Jun. 2018.

[10] I. Sutskever, O. Vinyals, Q. V. Le, “Sequence to sequence learning with neural networks // Advances in neural information processing systems. 2014: 3104-3112.

[11] P. Veličković, et al, "Graph attention networks," arXiv preprint arXiv:1710.10903 (2017).

[12] H. Burwinkel, A. Kazi, G. Vivar, S. Albarqouni, G. Zahnd, N. Navab, and S.-A. Ahmadi, “Adaptive Image-Feature Learning for Disease Classification Using Inductive Graph Networks,” // Medical Image Computing and Computer Assisted Intervention – MICCAI 2019, pp. 640–648, 2019.

[13] T. Gabruseva, D. Poplavskiy, and A. Kalinin, “Deep Learning for Automatic Pneumonia Detection,” // 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Seattle, WA, USA, 2020, pp. 1436-1443. DOI: 10.1109/CVPRW50498.2020.00183

[14] K. Simonyan, A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” arXiv preprint arXiv:1409.1556, 2014.

[15] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Jun. 2016.