2024 16th International Conference on Advanced Computational Intelligence (ICACI) | 979-8-3503-0850-1/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICACI60820.2024.10537017

Detection of Idiopathic Pulmonary Fibrosis Lesion Area Based on Transfer Learning

JieLin Xue
School of Information and Communication Engineering University of Electronic Science and Technology of China
Chengdu, China
1731043850@qq.com

Tian Pu*
School of Information and Communication Engineering University of Electronic Science and Technology of China
Chengdu, China
* Corresponding author:
putian@uestc.edu.cn

Lu Guo*
Department of Pulmonary and Critical Care Medicine, Sichuan Provincial People's Hospital School of Medicine, University of Electronic Science and Technology of China
Chengdu, China
* Corresponding author:
guoluhx@126.com

ZhenMing Peng
School of Information and Communication Engineering UESTC
Chengdu, China
zmpeng@ uestc.edu.cn

Xing He
Department of Pulmonary and Critical Care Medicine, Sichuan Provincial People's Hospital School of Medicine, University of Electronic Science and Technology of China
Chengdu, China
hexinggh1993@126.com

Yicen Han
Pulmonary and critical care medicine Chengdu second people's hospital
Chengdu, China
469662925@qq.com

***Abstract*—Idiopathic pulmonary fibrosis (IPF) is a chronic lung disease that significantly impacts individuals' health. Early diagnosis of IPF can enhance patient survival rates. Considering the limited research on automatic detection of IPF in academic circles and the scarcity of labeled data, this paper proposes a transfer learning-based algorithm for detecting IPF lesion areas. The proposed method consists of two stages: firstly, construct an IPF-like dataset comprising natural texture images with similar visual features to IPF, and pre-train the U-net network using this dataset; secondly, fine-tune the network using both the pre-trained weights and actual IPF data, followed by utilizing the trained network for lesion area detection. Experimental results demonstrate substantial improvements in terms of accuracy and sensitivity compared to corner distribution-based methods for IPF detection.**

***Keywords*—idiopathic pulmonary fibrosis, transfer learning, CT image**

I. INTRODUCTION

Idiopathic pulmonary fibrosis (IPF) is a chronic and progressive interstitial lung disease characterized by the presence of common interstitial pneumonia (UIP) on both lung histology and high-resolution chest CT scans (HRCT), with an unknown etiology [1]. The incidence and mortality rates of IPF have been gradually increasing over the years. The average survival time following diagnosis is a mere 3-4 years, which is considerably shorter compared to many malignancies. Moreover, the insidious and gradual onset of IPF makes it challenging for physicians to make accurate assessments, further compounded by its resemblance to other lung diseases. The diagnosis of IPF necessitates a multidisciplinary consultation. Currently, physicians achieve an approximate diagnostic accuracy rate of 70% for IPF. HRCT imaging assumes a pivotal role in facilitating the diagnosis of fibrotic lung diseases. Accurate diagnosis of IPF holds immense significance for prognosis and treatment decisions, as there exist pharmacological interventions that can effectively inhibit disease progression and improve survival rates. Early detection of IPF can significantly extend survival and reduce exacerbations. At the same time, the automatic detection of IPF can reduce the workload of doctors and make full use of the large amount of data in the hospital.

Existing IPF detection methods can be broadly categorized into conventional approaches and deep learning-based techniques. Best et al. conducted an analysis of CT images from 144 IPF patients, revealing a moderate correlation between kurtosis on the histogram, skewness, average attenuation, and pulmonary physiological abnormalities. Notably, kurtosis exhibited the strongest association [2]. Best et al. demonstrated disease progression by employing continuous assessment of quantitative CT scans [3]. Matsuoka classified emphysema as areas with low attenuation below -950 Hounsfield Unit (HU), while IPF was identified as areas with high attenuation ranging from 0 to -700 HU. This study concluded that IPF played a pivotal role in the decline of lung function [4]. Kauczor et al. defined the ground glass opacity in various interstitial lung diseases as having a HU range between -750 and -300 [5]. However, this approach relies only on a single intensity threshold of a single pixel, ignoring any potential interpixel relationships and can only obtain rough classification results of the lesion area. By extracting the volume texture features of CT images, Depeursinge et al. can classify UIP patients into typical and atypical [6]. The local binary model (LBP) [7] has also been

This work is supported by Sichuan Science and Technology Program (2022YFG0178).

applied to classify lung tissue interstitial models, using the nearest neighbor classifier, with a classification accuracy of 95.2% for emphysema. Vo et al. used a multi-scale sparse representation method based on wavelet transform and contour wave transform to describe four modes of pulmonary diffuse diseases in HRCT [8]. Anthimopoulos et al. proposed to use local DCT features to classify pulmonary interstitial disease patterns [9], and to use filter banks composed of DCT basis functions to classify lung patterns, which could better classify lung patterns except for certain misclassification of honeycomb and reticular structures. Sun proposed a method to extract IPF lesion areas by using multiple mixed texture features [10]. Yu et al. proposed an IPF lesion area detection method based on the difference of corner point distribution (CPD) between healthy and diseased areas [11]. To the best of our knowledge, these are the only two reported methods of automated IPF detection in the literature. The challenge of IPF detection lies in the variety of lesion types observed at different stages, as well as the significant variations in lesion area size, among other factors. Furthermore, due to the absence of a comprehensive large scale IPF dataset, there have been no reports on deep learning (DL)-based IPF detection methods. The lack of data makes it difficult to train deep networks. A solution to overcome this problem is to use transfer learning. The method in [12] pre-trained the network from six publicly available texture databases and fine-tuned the lung tissue data. This method showed the potential of transfer learning in the analysis of lung tissue patterns.

This paper proposes an IPF lesion region detection method based on transfer learning. By analyzing the imaging features of IPF lesions, we identified similar natural texture datasets to construct the IPF-like dataset. Furthermore, we address the challenges related to IPF detection by a modified U-net network [13]. We use the IPF-like natural texture images to pre-train the network, then using the pre-trained weights and actual IPF data to conduct secondary training on the network. Finally, we use the trained network to detect the lesion area. The main contributions of this paper are as follows:

(1) We propose a transfer learning-based method for the detection of candidate regions associated with IPF lesions. We construct IPF-like datasets, the network is pre-trained by constructing IPF-like texture dataset to overcome the difficulty of deep network training due to lack of data.

(2) We use Vgg16 [19] and Resnet50 [17] as the backbone network of U-net. U-net can make more efficient use of valid labeling data by relying on data enhancement of a small number of training images. Vgg16 and Resnet50 can extract more intricate features, which can better reflect the complex structure of IPF in CT images and improve the detection ability of small lesions. This helps to improve the performance of IPF detection.

(3) We incorporate the convolutional attention module (CBAM) [16] into the skip connection, which employs a combined strategy of channel attention and spatial attention mechanisms to improve the ability to detect small lesion areas and enhance the detection rate. Simultaneously, we employ a generalized loss function based on Tversky index [18] to address data imbalance issues and achieve a more optimal balance between accuracy and recall rate.

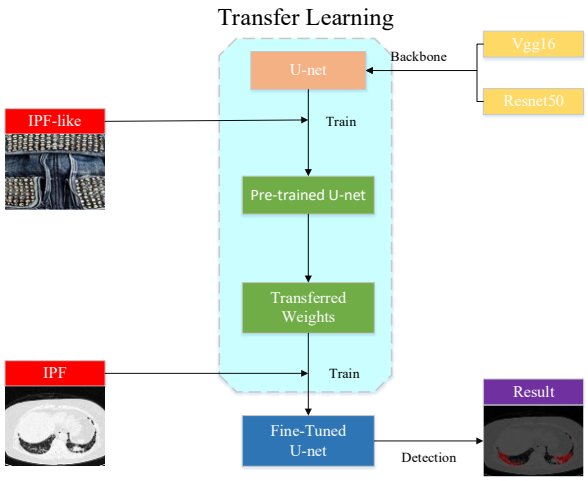

Fig. 1. Overall flow chart of the proposed method.

II. Methods

The flowchart of our method is shown in Fig. 1. Firstly, we analyze the imaging characteristics of the IPF lesion area to identify a natural texture dataset that closely resembles it, thereby constructing an IPF-like dataset. Subsequently, Vgg16 and Resnet50 are employed as the Backbone of U-net for network pre-training using the IPF-like dataset. Then, utilizing the trained weights and IPF data, U-net is further trained. Finally, the trained network is utilized for accurate detection of the lesion area.

A. Construction of IPF-like Datasets

Unlike numerous open datasets for natural image detection and recognition tasks, open medical image sets are not sufficient due to their high costs and expertise in acquisition and annotation. In contrast, ordinary people with simple training can annotate natural images. Given the difficulty in obtaining annotated IPF data, we propose to construct natural texture datasets similar to IPF lesion areas to improve detection quality through transfer learning.

As shown in Fig. 2, there are four imaging features of IPF on CT images: honeycomb, mesh shadow, ground glass opacity, tractive bronchiectasis and bronchiectasis [21]. Due to the lack of adequate ground glass opacity, tractive bronchiectasis and bronchiectasis CT images, we primarily focus on identifying honeycomb and mesh shadows. Honeycomb structures are characterized by a cluster of cystic cavities with thick walls and distinct profiles. Mesh shadows, on the other hand, manifest as an amalgamation of numerous small linear opaque objects that collectively form a net-like appearance. We have observed certain natural elements, such as bubbles and sponges, as shown in Fig. 3, which exhibit imaging features closely similar to their inherent textures. This study aims to construct IPF-like datasets based on selected natural texture datasets.

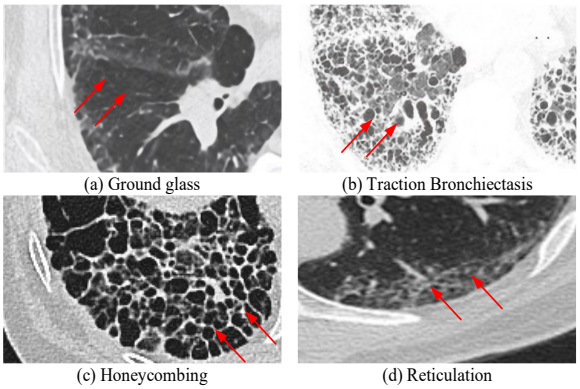

Fig. 2. IPF imaging features.

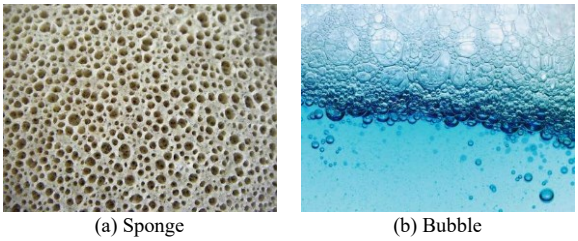

Fig. 3. Similar natural images.

This paper seeks suitable data from two open natural texture image datasets, including DTD [14] and A LOT [15]. ALOT comprises a collection of 250 color images featuring diverse textures, with each texture material represented by 100 images, resulting in a total dataset size of over 27,500 images. Each material exhibits variations in viewing angle, light angle, and light color. DTDS contains extensive texture datasets derived from natural images. Instead of labeling materials based on their names, this dataset employs 47 categories to describe the visual appearance of the images. Consequently, different materials may belong to the same category if they exhibit visually similar textures. This dataset can also serve as a valuable resource for material identification tasks. In this paper, we extract images with similar IPF texture characteristics from the two datasets and construct IPF-like datasets, such as bubbles and Mosaic. It is important to note that due to the lack of labeling in DTD and ALOT datasets, only RGB images are available. Therefore, it is necessary to annotate and convert the dataset into mask images, as shown in Fig. 4. Unlike medical images, this data can be labeled by individuals without expertise. In total, 99 sets of images (out of a total of 198) are selected to generate an IPF-like dataset.

B. *Deep Neural Networks*

The U-net has been widely adopted for semantic segmentation tasks across diverse applications. In this paper, we employ U-net as the fundamental network structure while leveraging Vgg16 and Resnet50 as its backbone networks. Additionally, CBAM is incorporated into the skip connection module to improve overall network performance, and the overall network structure is shown in Fig. 5.

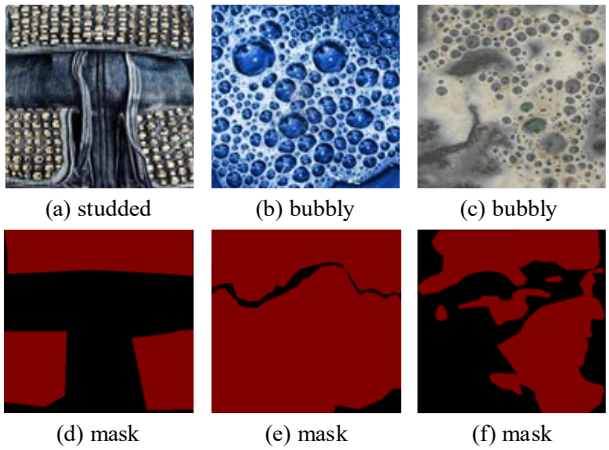

Fig. 4. Original image and mask image.

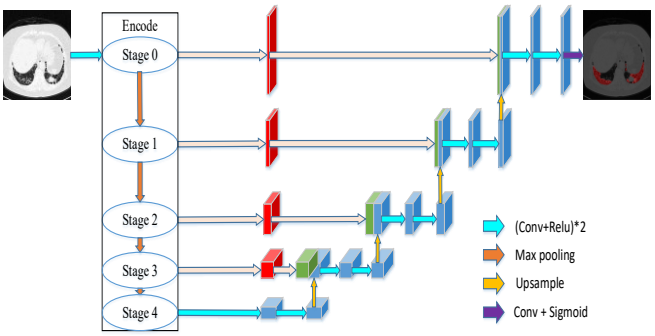

Fig. 5. Overview of U-net architecture.

a) *Encoder:* We modify U-net by replacing its encoder with Vgg16 and Resnet50 to improve the network's performance since. Vgg16 and Resnet50 can extract more intricate features, which is crucial for processing high-resolution images that encompass abundant detailed information. This facilitates a more accurate representation of the complex structure observed in HRCT images of IPF, thereby enhancing the detection capability for small lesions. Consequently, this advancement contributes to an overall improvement in IPF detection performance.

Experiments show that the performance of neural network can be improved to some extent by increasing the depth of neural network through VGGNet. So far, many researchers still choose VGGNet as the image feature extraction network. For IPF detection, the use of deeper network structure can improve the expressiveness and generalization ability of the model to better detect IPF lesion areas in CT images. VGGNet divides the network into five parts, each of which connects multiple 3×3 convolution cores, as shown in Fig. 6. Each segment is connected to a maximum pooled layer of size 2×2, followed by three fully connected layers and a softmax layer. The active units of all hidden layers use ReLU functions.

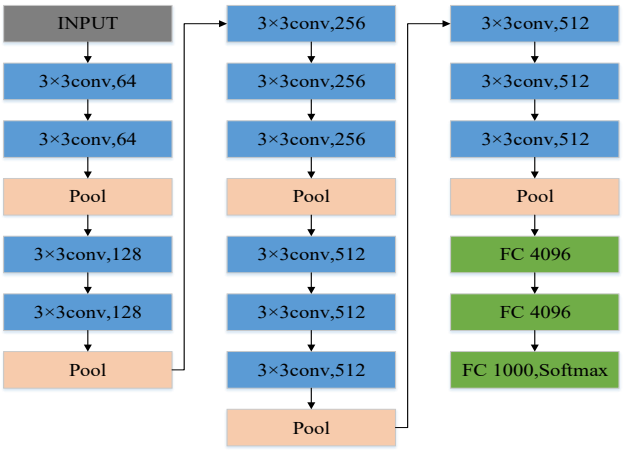

Fig. 6. The network structure of VGGNet.

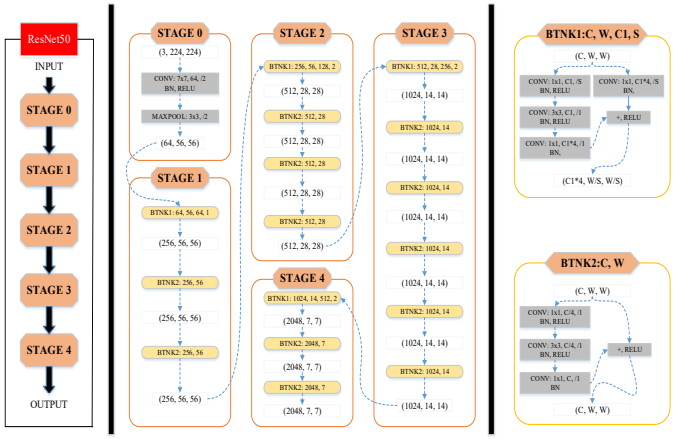

Fig. 7. The network structure of Resnet50.

In the development of convolutional neural networks, increasing the number of network layers can improve detection or classification performance. However, as the number of network layers increases, the issue of gradient vanishing arises within the neural network. To address this problem, residual networks have been introduced with a key focus on their residual units [17]. These residual units incorporate cross-layer connections that enable effective nonlinear transformations and facilitate better gradient propagation in deeper networks, thereby facilitating learning of more complex features. The overall structure of Resnet50 is divided into five phases, as shown in Fig. 7. The two bottlenecks correspond to two cases: the same number of input and output channels (BTNK2) and a different number of input and output channels (BTNK1).

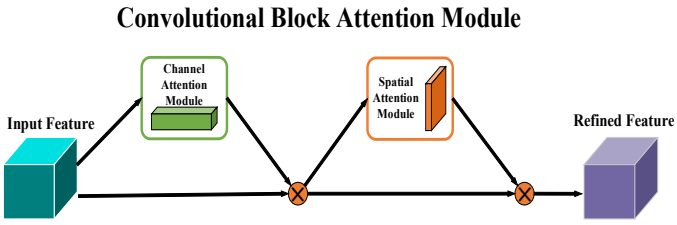

Fig. 8. The Overview of CBAM.

b) *Attention Mechanism Module:* The attention mechanism leverages the physiological mechanisms of human attention to enhance the flexibility of neural networks in learning specific features by learning different input weights during training. In this paper, we combine channel attention and spatial attention mechanisms to improve detection rates. Specifically, we employ CBAM which consists of a channel attention module (CAM) and a spatial attention module (SAM). The overall structure of CBAM is shown in Fig. 8. It first processes the input feature map through the channel attention module, then further refines it using the spatial attention module to obtain adjusted features. Using CBAM in U-Net's skip connections allows the network to pay more attention to features in the important up-sampling and down-sampling paths. At the same time, it helps the network better understand the context information of the input image. With the use of the low- and high-level features provided by skip connections, this module significantly improve image segmentation performance. Finally, the network has better adaptability for lesions of different sizes and shapes.

The input of CAM is the feature map of size $H \times W \times C$, where H represents the height of the feature map, W represents the width, and C represents the number of channels. The processing flow is as follows: 1) Firstly, global pooling and average pooling operations are performed on the input feature map in order to compress spatial dimensions and facilitate subsequent learning of channel features; 2) Secondly, both global and average pooling results are fed into a multi-layer perceptron (MLP) for learning purposes, taking into account MLP's ability to learn channel dimensions effectively along with each channel's importance; 3) Finally, an "addition" operation is applied to combine the output from MLP layers. By subsequently mapping this result through a Sigmoid function, we obtain the final value representing channel attention. The calculation formula can be expressed as follows:

$$M_c(F) = \sigma(MLP(AvgPool(F)) + MLP(MaxPool(F))) \quad (1)$$

The input of SAM is the feature map of the CAM output. The processing flow is as follows: 1) Firstly, the input feature map undergoes global pooling and average pooling operations, resulting in channel-wise compression and facilitating subsequent spatial feature learning; 2) Secondly, the results of global pooling and average pooling are concatenated along the channel dimension, yielding a feature map with dimensions $H \times W \times 2$; 3) Finally, a convolution operation is applied to the concatenated results, resulting in a feature map of size $H \times W \times 1$, which is then passed to an activation function. The calculation formula can be expressed as:

$$M_s(F) = \sigma(f^{7\times7}([AvgPool(F); MaxPool(F)])) \quad (2)$$

c) *Loss Function:* Data imbalance poses a significant challenge in medical image segmentation. In general, the number of diseased voxels is much lower than the number of non-diseased voxels, and training models with such data will produce high-precision results, but with low recall (sensitivity). We use a generalized loss function based on the Tversky index [18] to solve the problem of data imbalance. The TverskyLoss

enables the network to prioritize the lesion area and achieve a better trade-off between accuracy and recall. Moreover, false positives or missed positives can have detrimental consequences for medical images. By incorporating TverskyLoss into our model, we effectively mitigate the impact of false positives or false negatives and improve overall performance. The TverskyLoss is presented as:

$$T(A,B) = 1 - \frac{A \cap B}{A \cap B + \alpha \mid A - B \mid + \beta \mid B - A \mid}, \tag{3}$$

where B is the real mask, A is the predicted mask, $|A - B|$ is false positives, $|B - A|$ is false negatives, α and β can control the tradeoff between false positives and false negatives, and the more we focus on recall rates, the higher the β.

III. Result And Discussion

In this section, we present the experimental results. As far as we know, there are few reports about IPF automatic detection. A IPF detection methods based on difference of corner distribution (CPD) between healthy areas and diseased areas is reported in [11]. In this section, we compare the proposed method with the CPD method.

A. Experimental Data and Parameter Settings

We collect the IPF datasets from two sources: the Lung Interstitial Diseases Database (ILD-database) [20] and Sichuan Provincial People's Hospital (SPPH). A total of 371 fibrosis-labeled images are selected from the ILD database, while the SPPH dataset comprise 897 annotated CT images from 14 patients with IPF. In our experiment, we employ a training-to-test set ratio of 9:1 for IPF-like data and a ratio of 7:3 for IPF data.

The proposed method has been run on Ubuntu 18.04 with TITAN RTX at 24GB of RAM. The network structure is implemented through PyTorch. After downloading the vgg16 and resnet50 pre-training weights available online, U-net is instructed to use these weights and freeze the training (the backbone of the model is frozen at this time. The feature extraction network does not change and occupies a small amount of video memory. The network is only fine-tuned). We set freeze Epoch to 50, batch size to 2, and total training Epoch to 100. In addition, the maximum learning rate(LR) of the model is set to 1e-4, and the minimum LR is set to 0.01 times of the maximum LR. In addition, set the optimizer type to ADAM and the LR decline method to Cosine Annealing LR(so that the LR varies periodically according to the cosine function). TverskyLoss has an α set of 0.3 and a β set of 0.7. After training, U-net loads the newly generated pre-training weights and trains based on the IPF data set. The training parameters are the same as before, and the training to verification ratio is 7:3.

B. Evaluation Indicators

We use the following indexes to objectively evaluate the performance of the algorithm, including Dice coefficient, MIoU, sensitivity, accuracy, and F-score:

$$Dice = \frac{2 \times TP}{FN + 2 \times TP + FP}, \tag{4}$$

$$MIoU = \frac{TP}{FN + TP + FP}, \tag{5}$$

$$Sensitivity = \frac{TP}{TP + FN}, \tag{6}$$

$$Precision = \frac{TP}{TP + FP}, \tag{7}$$

$$F\text{-}score = \frac{2TP}{2TP + FN + FP}. \tag{8}$$

TP (True Positives) indicates that the actual category is lesion and the predicted category is lesion. FP (False Positives) indicates that the actual category is healthy and the predicted category is diseased. FN (False Negatives) indicates that the actual category is diseased and the predicted category is healthy. Since the CPD method employs its own defined metrics, we strive for consistency with these evaluation indicators for a fair comparison. Specifically, TP signifies that a lesion region has been correctly identified where at least 20% of the pixels within this region align with ground truth annotations; FP denotes regions falsely classified as lesions where less than 20% of their pixels correspond to ground truth annotations; FN represents regions erroneously considered healthy despite having at least 20 percent of their pixels affected by disease.

C. Experimental Results

TABLE I shows the ablation experiment results of the proposed method, demonstrating notable improvements across all indices following the incorporation of IPF-like data. Notably, with the integration of attention mechanism and Tversky-Loss, utilizing Resnet50 and Vgg16 as backbone networks, our approach has achieved remarkable outcomes in terms of Dice coefficient and Sensitivity.

Fig. 9 shows the experimental results when the attention mechanism is not added and TverskyLoss is not used. For the part of the loss function we use the common Dice-Loss in medical image segmentation. Column (a) represents CT images with lesions respectively. The red areas in columns (b) and (c) are the detection results obtained by Resnet50 and Vgg16 as Backbone, respectively. The red areas in columns (d) and (e) are the detection results obtained after adding the attention mechanism and Tversky-Loss with Resnet50 and Vgg16 as the Backbone, respectively. Column (f) is Ground truth. It is evident that without incorporating the attention mechanism and TverskyLoss, there are instances of missed detections and false positives within the network's output. However, these issues have been effectively addressed upon their inclusion, validating our proposed enhancements.

The results of the comparison between this method and CPD are presented in TABLE II, where evaluation indicators consistent with the CPD method were employed. It is evident that this method exhibits higher sensitivity and accuracy in segmenting the lesion area compared to CPD.

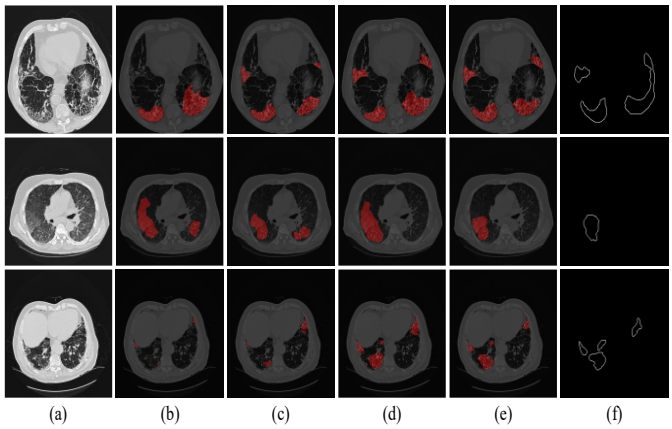

Fig. 9. (a) CT images. (b) The results obtained from the method of using resnet50 as the backbone. (c) The results obtained from the method of using vgg16 as the backbone. (d) The results obtained by using resnet50 as the backbone and incorporating CBAM and Tversky-Loss methods. (e) The results obtained by using vgg16 as the backbone and incorporating CBAM and Tversky-Loss methods. (f) ground truth.

TABLE I. ABLATION RESULTS

Change	Indicators				
	MIoU	*F-score*	*Dice*	*Precision*	*Sensitivity*
Resnet50	0.67	0.53	0.53	0.48	0.59
Resnet50+ IPF-like	0.69	0.57	0.57	0.43	0.60
Vgg16+ IPF-like	0.71	0.59	0.59	0.53	0.66
Resnet50+ attention	0.77	0.70	0.70	0.73	0.68
Vgg16+ attention	0.77	0.72	0.72	0.81	0.64
Resnet50+ TverskyLoss	0.79	0.73	0.73	0.68	0.80
Vgg16+ TverskyLoss	0.79	0.74	0.74	0.78	0.70

TABLE II. COMPARISON RESULTS OF DIFFERENT METHODS

methods	Indicators			
	Dice	*Precision*	*Sensitivity*	*F-score*
CPD	0.59	0.72	0.85	0.78
Ours(Resnet)	0.73	0.72	0.87	0.78
Ours(Vgg16)	0.74	0.83	0.87	0.85

Fig. 10 shows the detection results of IPF lesion area by our method. Column (a) represents CT images with lesions respectively. The red area in column (b) is the detection result obtained by Resnet50 as the Backbone. The red area in column (c) is the detection result obtained by Vgg16 as the Backbone. Column (d) is Ground truth. It can be observed that whether it is a relatively large lesion area or a relatively small lesion area, our method can accurately detect the IPF lesion area, and provide a more accurate segmentation of the lesion area.

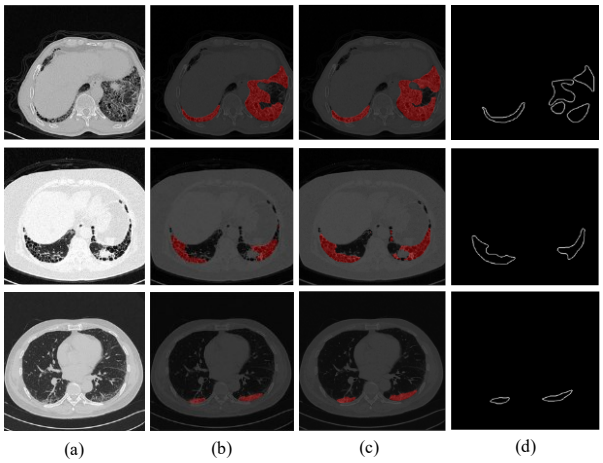

Fig. 10. (a) CT images. (b) The results obtained from the method of using resnet50 as the backbone. (c) The results obtained from the method of using vgg16 as the backbone. (d) ground truth.

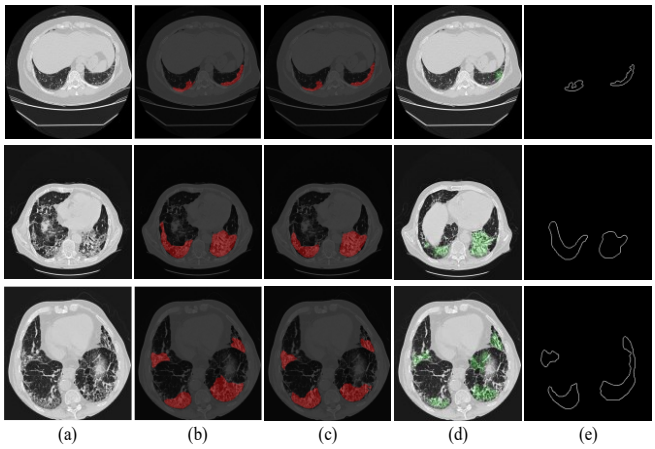

Fig. 11. (a) CT images. (b) The results obtained from the method of using resnet50 as the backbone. (c) The results obtained from the method of using vgg16 as the backbone. (d) CPD. (e) ground truth.

Fig. 11 shows a comparison of the proposed method with CPD. Column (a) represents CT images with lesions respectively. The red area in column (b) is the detection result obtained by Resnet50 as the Backbone. The red area in column (c) is the detection result obtained by Vgg16 as Backbone. The green area in column (d) is the detection result obtained by the CPD method. Column (e) is Ground truth. It can be observed that the detection result of CPD method for some diseased areas is smaller than the Ground truth, and there will be problems of missing detection. At the same time, some non-diseased areas produced false detection. The method we propose gives more accurate results.

IV. CONCLUSIONS

This paper presents an algorithm for the detection of IPF lesion areas based on CT images. To address the challenge of limited labeled data in IPF, we employ transfer learning. Firstly, we construct an IPF-like dataset using natural texture datasets and then build a U-net network with Resnet50 and Vgg16 as

backbones. Attention mechanisms are incorporated to enhance the network's capability in detecting small lesion areas, while TverskyLoss is utilized to improve sensitivity. We initially pre-train the network using the IPF-like dataset, followed by fine-tuning with parameters from pre-training on the IPF dataset. Finally, we utilize the trained network to obtain accurate detection results. Experimental results demonstrate that our method effectively detects and segments lesion areas, making it highly significant for automated detection of IPF lesions.

Most of the current research on IPF has been conducted on two-dimensional slices, which are detected by performing a series of treatments on the two-dimensional slices. However, medical images are usually 3-dimensional or 4-dimensional, and a large amount of spatial information contained in them will be ignored when detected on a two-dimensional plane. Coupled with the lack of IPF data, the accuracy rate of IPF detection has been very low and has not achieved good results. At the same time, many scholars have achieved good results in 3D detection in other medical fields, but they have not been applied to IPF detection. Therefore, the author believes that it is a future development trend to strengthen cooperation with hospitals to expand the amount of data and expand IPF detection to 3D.

ACKNOWLEDGMENT

This work is supported by Clinical Research and Sichuan Provincial People's Hospital.

REFERENCES

[1] T. M. Maher and M. E. Strek, "Antifibrotic therapy for idiopathic pulmonary fibrosis: time to treat," *Respiratory Research*, vol. 20, no. 1, Sep. 2019.

[2] A. C. Best, A. M. Lynch, C. M. Bozic, D. Miller, G. K. Grunwald, and D. A. Lynch, "Quantitative CT Indexes in Idiopathic Pulmonary Fibrosis: Relationship with Physiologic Impairment," *Radiology*, vol. 228, no. 2, pp. 407–414, Aug. 2003.

[3] A. C. Best, J. Meng, A. M. Lynch, C. M. Bozic, and D. Miller, "Idiopathic Pulmonary Fibrosis: Physiologic Tests, Quantitative CT Indexes, and CT Visual Scores as Predictors of Mortality," *Radiology*, vol. 246, no. 3, pp. 935–940, Mar. 2008.

[4] S. Matsuoka, T. Yamashiro, S. Matsushita, A. Kotoku, and A. Fujikawa, "Quantitative CT Evaluation in Patients with Combined Pulmonary Fibrosis and Emphysema," *Academic Radiology*, vol. 22, no. 5, pp. 626–631, May 2015.

[5] Hans-Ulrich Kauczor, K. Heitmann, Claus Peter Heussel, D. Marwede, and T. Uthmann, "Automatic Detection and Quantification of Ground-Glass Opacities on High-Resolution CT Using Multiple Neural Networks," *American Journal of Roentgenology*, vol. 175, no. 5, pp. 1329–1334, Nov. 2000.

[6] A. Depeursinge, A. Chin, A. N. Leung, D. Terrone, and M. R. Bristow, "Automated Classification of Usual Interstitial Pneumonia Using Regional Volumetric Texture Analysis in High-Resolution Computed Tomography," *Investigative Radiology*, vol. 50, no. 4, pp. 261–267, Apr. 2015.

[7] L. Srensen, S. B. Shaker, and M. de Bruijne, "Quantitative Analysis of Pulmonary Emphysema Using Local Binary Patterns," *IEEE Transactions on Medical Imaging*, vol. 29, no. 2, pp. 559–569, Feb. 2010.

[8] K. T. Vo and Arcot Sowmya, "Directional Multi-scale Modeling of High-Resolution Computed Tomography (HRCT) Lung Images for Diffuse Lung Disease Classification," *Computer Analysis of Images and Patterns*(CAIP), Berlin, Heidelberg, 2009, pp. 663–671.

[9] M. Anthimopoulos, S. Christodoulidis, A. Christe and S. Mougiakakou, "Classification of interstitial lung disease patterns using local DCT features and random forest," *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Chicago, IL, USA, 2014, pp. 6040-6043.

[10] Y. Sun, "Visual information modeling and lesion detection method for Idiopathic Pulmonary Fibrosis," unpublished.

[11] X. Yu, Y. Han, T. Pu, L. Guo, and Z. Peng, "Detection of idiopathic pulmonary fibrosis lesion regions based on corner point distribution," *2022 7th International Conference on Intelligent Computing and Signal Processing*, Xi'an, China, 2022, pp. 502-506.

[12] S. Christodoulidis, M. Anthimopoulos, L. Ebner, A. Christe, and S. Mougiakakou, "Multisource Transfer Learning With Convolutional Neural Networks for Lung Pattern Analysis," *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 1, pp. 76–84, Jan. 2017.

[13] O. Ronneberger, P. Fischer and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation", *Medical Image Computing and Computer-Assisted Intervention*, 2015, pp. 234-241.

[14] M. Cimpoi, S. Maji, I. Kokkinos, S. Mohamed, and A. Vedaldi, "Describing textures in the wild," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 3606–3613.

[15] G. J. Burghouts and J.-M. Geusebroek, "Material-specific adaptation of color invariant features," *Pattern Recognition Letters*, vol. 30, no. 3, pp. 306–313, Feb. 2009.

[16] S. Woo, J. Park, J.-Y. Lee and I. S. Kweon, "CBAM: Convolutional block attention module", *Proc. Eur. Conf. Comput. Vis*.(ECCV), 2018, pp. 3-19.

[17] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2016, pp. 770–778.

[18] S. S. M. Salehi, D. Erdogmus, and A. Gholipour, ''Tversky loss function for image segmentation using 3D fully convolutional deep networks,'' in *Proc. Int. Workshop Mach. Learn. Med. Imag.*, 2017, pp. 379–387.

[19] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *arXiv.org*, Apr. 10, 2015.

[20] A. Depeursinge, A. Vargas, A. Platon, A. Geissbuhler, and P.-A. Poletti, "Building a reference multimedia database for interstitial lung diseases," *Computerized Medical Imaging and Graphics*, vol. 36, no. 3, pp. 227–238, Apr. 2012.

[21] G. Raghu, M. Remy-Jardin, J. L. Myers, L. Richeldi, and C. J. Ryerson, "Diagnosis of Idiopathic Pulmonary Fibrosis. An Official ATS/ERS/JRS/ALAT Clinical Practice Guideline," *American Journal of Respiratory and Critical Care Medicine*, vol. 198, no. 5, pp. e44–e68, Sep. 2018.