Proceedings of the Fourth International Conference on Sentiment Analysis and Deep Learning (ICSADL-2025)
IEEE Xplore Part Number: CFP25UU5-ART; ISBN: 979-8-3315-2392-3

Unraveling Pulmonary Fibrosis Deep Learning Transfer Learning for Diagnosis

Rekha G
Dept.Computing Technologies
SRM Institute of Science and Technology
Kattankulathur,Chennai-603 203
rekhag1@srmist.edu.in

Pongiannan R K
Computing Technologies
SRM Institute of Science and Technology
Kattankulathur,Chennai-603 203
pongiank @srmist.edu.in

Venkata Sreeram Pachipulusu
Computing Technologies
SRM Institute of Science and Technology
Kattankulathur,Chennai-603 203
pp8129@srmist.edu.in

Sai Balaji Mogili
Dept.Computing Technologies
SRM Institute of Science and Technology
Kattankulathur,Chennai-603 203
mm6599@srmist.edu.in

Brindha R
Dept.Computing Technologies
SRM Institute of Science and Technology
Kattankulathur,Chennai-603 203
brindhar@srmist.edu.in

Telugu Maddileti,
Department of Electronics &Commn Engg,
Guru Nanak Institutions Technical Campus,
Ibrahimpatnam, Telangana-501506
madhu14283@gmail.com

***Abstract*— Pulmonary Fibrosis is a chronic lung disease demands prompt diagnosis for effective management. The proposed model develops an automated system using deep learning for feature extraction and machine learning for classification, reducing reliance on manual assessments. In the context, we use InceptionV3, VGG16, and MobileNetV2 for feature extraction as they are constantly successful and able to detect fine-grained patterns from medical images. The extracted features are further analyzed with popular machine learning techniques including but not limited to Random Forest and SVM. The proposed new hybrid model, leads to a faster diagnosis even if it is comparatively more accurate, than the previously used models. The system performance are evaluated using the accuracy, precision, recall, and F1-score. These are integrated into the Flask-based web application and it enables quick and accurate Pulmonary Fibrosis diagnosis, aiding timely treatment and better patient outcomes.**

Keywords*— *Pulmonary Fibrosis, Diagnosis, Deep Learning Machine Learning, CT Scan, Feature Extraction.

I. INTRODUCTION

Early diagnosis of pulmonary fibrosis is important so as to ensure that proper management is made and thus enhance prognosis of the illness. Unfortunately, early diagnosis of the disease is very difficult because the symptoms are vague and there is a lack of highly specific diagnostic methods. At the moment, diagnosis of pulmonary fibrosis employs a clinical approach, pulmonary function tests, and blood tests, imaging include (high-resolution computed tomography) and in some cases lung biopsy. Although patterns of fibrosis, in particular, can be determined utilizing HRCT scans, the analysis of the scans themselves may not be entirely easy and requires much time and skill.

The authors believe that the development of DL as a type of AI can potentially revolutionize the ways of diagnosing and treating pulmonary fibrosis. DL algorithms can learn and recognize a large amount of medical data, for examples, imaging data, and discover the existing patterns and features impossible to detect by clinicians. The use of deep learning (DL) may result in increased accuracy and efficiency in making diagnosis and subsequent treatment since patient profiles inside electronic health record systems will be known.

For the past few years, several studies have been talking about AI in pulmonary fibrosis. An example of such research is the one that assessed the performance of EfficientNet plus quantile regression when applied to the OSIC dataset; the result was that it could improve prediction. Fibrosis-Net was another research because it introduced a deep learning model focused on CT images in order to evaluate the disease progression. Our project takes these advances in work where Inception V3 was integrated with a Random Forest for CT fibrosis diagnosis and has also incorporated a Flask-based web application to ensure that healthcare professionals have a channel through which they may reach what promises to be an improved, accurate, and effective automated system for fibrosis detection.

II. LITERATURE REVIEW

Prognostic tests of idiopathic pulmonary fibrosis (IPF) determined the most accurate prognosis methods of predicting disease progression [1]. Ultrasonic scattering in diagnosing pulmonary fibrosis in rodents proposed its use as a noninvasive diagnostic tool [2]. Machine learning model based on EfficientNet and quantile regression, improving diagnosis for PF progression with the OSIC dataset [3]. Investigating early indication of lung function decrease in subjects enjoined an emphasis on timely intervention via predictive analytics [4]. Ultrasound imaging gave a 92% sensitivity and specificity in the diagnosis of pulmonary fibrosis [5]. Another technique which utilized multi-frequency analysis of B-lines gave a 100% sensitivity with a 90% specificity, indicating the promise of ultrasound in clinical applications [6]. Machine learning-based on CT scan applied models refine lung capacity estimation in PF patients, removing the constraints of standard spirometry [7]. It included a deep learning based model Fibrosis-Net made for CT-image analysis and prediction of the progress of the disease [9]. A multi-scale guided attention model, MGA-Net, contributed to automated IPF diagnosis by accurate lesion area detection in high-resolution CT images [10]. One study aimed at determining the role of supplemental oxygen on cardiovascular and respiratory dynamics in IPF patients, which may be reflected in the therapeutic strategies [11]. A newer two-stage methodology for IPF lesion detection using CT scans features any improvement over its earlier stage [12]. Domain generalization techniques to improve deep learning models for the detection of pulmonary fibrosis [13]. The predictive model combined EfficientNet and quantile regression, improving accuracy for PF progression prediction [14]. Explained vision transformers for medical image segmentation, advantages and applications in improving segmentation accuracy. These studies together improve diagnostic and prognostic tools for pulmonary fibrosis using machine learning, deep learning, and advanced imaging techniques.

Recent studies shown that AI can be potentially effective in the detection of pulmonary fibrosis, primarily through the use of deep learning models applied to CT images. Yet while it has been effective in improving predictive accuracy with these approaches. In most of the existing models, there is consideration for seamless integration into a clinical workflow, thereby rendering such models practically useless. Thus, the project involves developing an AI-based detection system that encompasses Inception V3 and that employs a Random Forest classifier for the analysis of CT scans. The web application is coupled and hosted by Flask that seeks to deliver a user-friendly interface that allows efficient and accessible automated fibrosis diagnosis.

III. METHODOLOGY

The detection approach in pulmonary fibrosis using CT images uses a two-step method that incorporates the best deep-learning feature extraction techniques along with machine-based classification. By utilizing InceptionV3, VGG16, MobileNetV2, Random Forest, and SVM classifiers, the performance on both accuracy and computational demands would be kept in balance. The current chapter briefly introduces the proposed system and identifies different parts. Factors/phases of the proposed system as presented in Figure 1.

A. Data Preprocessing

However, the preprocessing is highly important before feeding CT images to the deep learning models since the dataset must be cleaned and normalized first. Preprocessing includes:

Data Augmentation: There is the scarcity of labeled medical data hence; several data augmentation techniques are used in an attempt to overcome the current issue. Here we have rotation, zooming, flipping and shifting to make the model sample bigger and make the model ready to encounter different variations in the data.

Noise Reduction: In CT images, there is usually noise which influences the training of the model in learning some of the features to predict its output. Gaussian filtering or median filtering is used to reduce noises on the input data so as to enhance the quality of the input data.

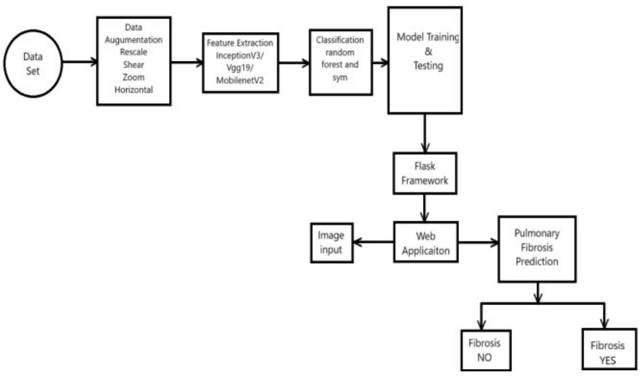

Fig:1 System Architecture

B. Feature Extraction Using Deep Learning Models

The three pre-trained CNNs used here for feature extraction are InceptionV3, VGG16, and MobileNetV2, all of which are good candidates for the following reasons: creating patterns in medical images and well-known in image classification tasks. Each of the models is fine-tuned on the Pulmonary Fibrosis dataset, and the following steps are performed:

InceptionV3: The present model accommodates complex features provided by its inception module and multi-scale features extracted from the input CT images. The architecture has both global and local features are captured within the network, making it quite eligible for detecting subtle patterns in lung fibrosis

VGG16: The VGG16 is a deeper network that uses 19 layers to capture hierarchical features. It is particularly effective in medical imaging tasks where spatial relationships between pixels are crucial. VGG16 extracts detailed features such as fibrosis patterns, honeycombing, and reticulation from the lung tissues.

MobileNetV2: The model is a lightweight model that remains very efficient and is targeted for mobile and resource-constrained environments. It uses depth wise separable convolutions, which reduces the computational complexity with almost the same performance. The MobileNetV2 model is best suited for the real-time applications that needs to be faster prediction and without compromising too much on accuracy. After the models have been pre-trained, these features are transmitted to classifiers trained using machine learning for further processing.

C. Feature Selection and Dimensionality Reduction

When the features have been successfully extracted by deep models, selection of features, which are adequate for classification purpose. Since the feature vectors often tend to have a high dimension in CNN, it is beneficial to reduce such dimensionality. For this particular reason, PCA, t-Distributed Stochastic Neighbor Embedding are applied. The present technique reduces the computational load and ensures that the classifiers focus on the most significant features, improving both accuracy and processing speed.

D. Classification Through Machine Learning Algorithms

These feature sets obtained from the operational perspective are used to train the machine learning classifiers. The proposed system uses two classifiers - the Random Forest and Support Vector Machines (SVM).

Random Forest: The Random Forest is an ensemble learning method that constructs multiple decision trees during training, with the mode from all of these trees classifying individual test trees. The advantage of Random Forest is its power to handle linear or non-linear data. By employing Random Forest, one can handle pulmonary fibrosis patterns, which can range in nature from shape to size to texture in different CT images.

SVM (Support Vector Machines): SVMs are excellent classifiers when dealing with high-dimensional data. SVM is great at separating large data that are not linearly separable. The SVM fits the input features into higher dimensions, and then it hunts down the best hyperplane on which to classify data in fibrosis and non fibrosis patterns. In the present study, non-linear classification used the radial basis function (RBF) kernel.

Apart from the above explanation, it must be noted in brief that both Random Forest and SVM have their strengths. Random Forest is effective on supporting complex data structures, while SVMs are remarkable for high-dimensional data.

E. Model Evaluation Metrics

In order to assess the performance of the proposed system, the following evaluation metrics are used:

Accuracy: Accuracy is the number of true positive classifications (true fibrosis) divided by all positive classifications (correct also, as well as wrong ones).

Recall: The calculation of recall involves dividing the number of true positive classifications (true fibrosis) in all population cases of the real positives in actual practice (real fibrotic cases).

F1 score: The F1 score is the harmonic mean of the precision and recall that produces a symmetric evaluation of the performance of the model.

These metrics will be calculated for each classifier, and the results are compared to identify which combination of the deep learning feature extractor and the machine learning classifier performs the best.

F. Integration with Flask-Based Web Application

Final component of the proposed system is a web based applications that enables easy and real-time access for medical professionals. Flask, a lightweight Python web framework, is used to create the web interface. The integration process involves:

Model Deployment: The trained deep learning models (InceptionV3, VGG16, and MobileNetV2) and machine learning classifiers (Random Forest and SVM) are saved and deployed within the Flask application.

User Interface: The Flask web interface provides a simple interface where medical professionals can upload CT images of the lungs. Upon submission, the images are preprocessed, and features are extracted using the trained deep learning models.

Real-Time Predictions: After feature extraction, the machine learning classifiers (Random Forest or SVM) predict whether the image shows signs of Pulmonary Fibrosis. The results, including confidence scores, are displayed in real-time on the web interface.

Backend Processing: The web application handles image uploads, preprocessing, model inference, and returns predictions within a few seconds, ensuring that the system is suitable for real-time medical diagnostics.

IV. Result Analysis

Random Forest Classifier in fig:2 & 3 outcomes of yours show the effectiveness of the model in identifying two classes (0, 1) in the dataset with state 1 being likely related to Pulmonary Fibrosis or other classifications in your system. Here's an explanation of the key metrics and what they reveal about the Inception-based feature extraction when combined with Random Forest for classification:

Accuracy is the degree to which the tools were able to classify the right instance out of those present in the dataset. Specifically, using the Inception model features, Random Forest Classifier is achieved an accuracy of approximately 75.33% of test-set classification.

The current performance is acceptable but it shows that there could be optimization in a way using a different model, over a different set of data or techniques like data augmentation.

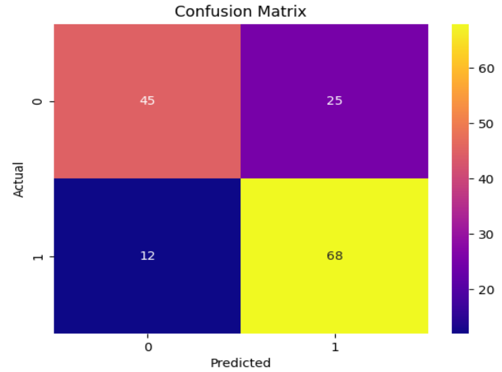

Fig:2 Confusion matrix of model1-(InceptionV3 and Random Forest)

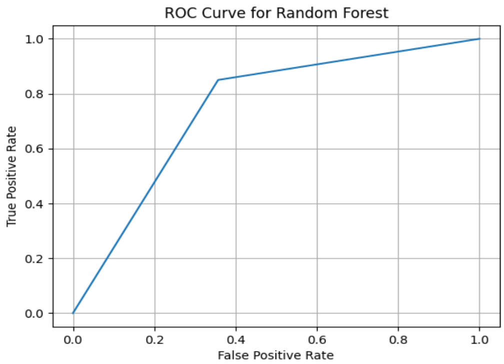

Fig:3 ROC of model1 (InceptionV3 and Random Forest)

Class 0 Performance (support = 70)

Precision (0.79): That is 3 out of all the instances that were classified under class 0 were precise, a precision of 79%. The current model indicates that it is actually quite capable of doing a reasonable job in not flagging many false positives for the current class.

Recall (0.64): As for class 0 instances, the model accurately classified 64 per cent of the actual class 0 samples. Fewer samples recalled to class 0 means the model is failing to predict some positive samples (the false-negative problem).

F1-Score (0.71): Precision and recall are combined into a single value by F1 score. While the score of 0.71 suggests fair performance for class 0 the model seems to be also predicting some instances of class 0.

support 80 = Class 1 Performance

Precision (0.73): If we consider class 1, it turns out that the model accurately classified actually 73 percent of the instances.

Recall (0.85): Class 1: 85% of actual class 1 instances were retrieved clearly showing improved identification of class 1 over class 0.

F1-Score (0.79): As for class 1, it also reaches 0.79 of F1, meaning that in fact we achieved better results than for class 0 because of increasing both precision and recall.

Macro and Weighted Averages

Macro Average (0.76): The Marco average is a three-point mean of precision, recall, and F1 for the two classes which are computed without weighting.

Weighted Average (0.76): The Weighted Average considers the number of instances that support each class and is generally a better measure of the performance of the model. These equal tendencies of both weighted and macro averages suggest that class imbalance is not a matter of concern here.

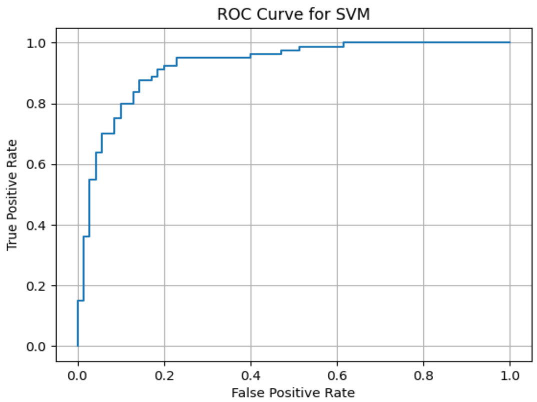

Fig:4 ROC of model2 (VGG16 and SVM)

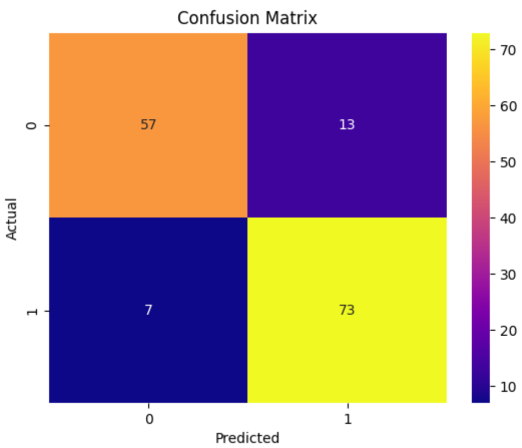

Fig:5 Confusion matrix of model2-(VGG16 and SVM)

The final work of the SVM classifier in fig:5 and 6 with the accuracy of 86.67% shows that the model is highly effective when using VGG16 for extracting features. Let's break down the results and how VGG16's architecture contributes to the performance:

The measurement of accuracy shows that the model identified slightly over 87 percent of the existent instances in the test data. Compared to the Random Forest classifier performance here, there is a slight increase and we can attribute to the fact that SVM is capable of creating more complex decision boundaries.

Class 0 Performance When support = 70

Precision (0.89): Among all the instances that the system classified as class 0, 89% were of the same classification. The current model implies that the model has a low tendency of misclassifying instances from class 0, which make it capture most instances from the current class appropriately.

Recall (0.81): As for the model accuracy in defining class 0, it was as follows: the model correctly determined 81% of actual cases. Nevertheless, the current recall is still rather elevated which hints at the fact that the required model might be missing up to 19% of class 0 instances or false negatives.

F1-Score (0.85): The F1-score combines both recall and precision and shows encouraging results: the accuracy of identifying class 0 is satisfactory while ensuring the minimum number of mistakes.

Class 1 Performance (support = 80)

Precision (0.85): "A precision of 85% was observed for class 1, indicating that 85% of instances predicted as class 1 were indeed accurate. However, the model exhibits slightly lower precision for class 1 than for class 0, raising the possibility of more false positives for class 1."

Recall (0.91): The high 91% of actual class 1 instances was accurately extracted by the model proving its good performance in terms of class 1 false negatives.

F1-Score (0.88): The balanced metric indicates that the classifier is nearly perfect on class 1 and it is pretty good on recall and precision tradeoff.

Macro and weighted averages

Macro Avg (0.87): The scores of two classes are relatively close, which means the proposed model is satisfactory on the whole.

Weighted Avg (0.87): The average represents relative strengths of the classes; the number of instances supporting or backing each class demonstrated that the classifier is sound and performs well across the board.

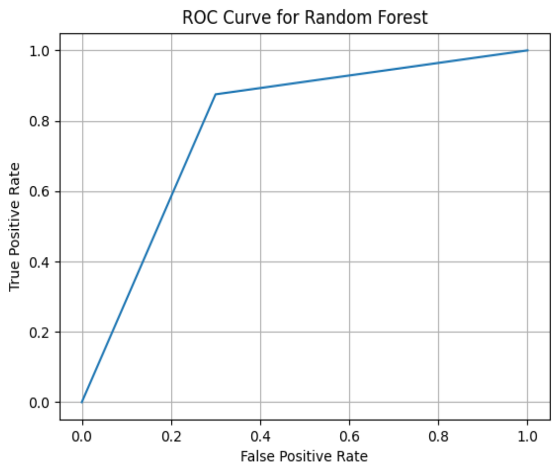

Fig:6 ROC of model3 (MobileNetV2 and Random Forest)

The outcome of the Random Forest Classifier with accuracy of 79.33% in fig 6 & 7 indicates good performance of the model when MobileNetV2 is used for feature extraction.

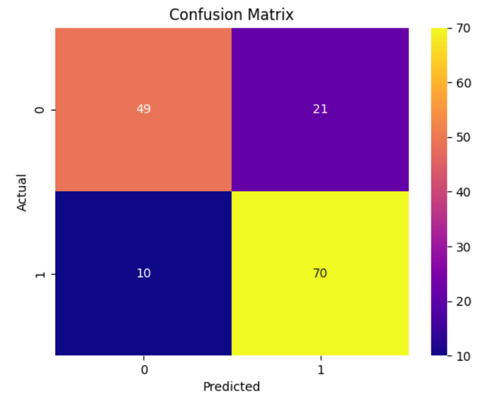

Fig:7 Confusion Matrix of model3 (MobileNetV2 and Random Forest)

The Random Forest Classifier performance was 79.33% accurate in classifying the entire range of instances. It is even better than in earlier cases when working with other models, which testifies to the fact that MobileNetV2 in the capacity of a feature extractor is efficient if used in the scenario.

Class 0 Performance (support = 70)

Precision (0.83): In class 0, the classifier got 83% of the instances right which had been tagged as class 0. Here it means that the classifier does a reasonably good job in recognizing class 0 even though an 83 percent precision indicate that there were some false positives.

Recall (0.70): The model was able to predict class 0 with good precision, where 70% of the actual class 0 samples were succeeded to be predicted as the model. The lower recall shows that model fails to rescue some instances from class 0, therefore resulting in false negatives.

F1-Score (0.76): In terms of precision and recall the F1-score shows that class 0 on the whole is correctly classified albeit at a lower recall value.

Support = 80 Class 1 Performance

Precision (0.77): In the class 1, the accuracy to the predictions was at 77 %. The specificity of the results points to the fact that while the main idea of the model is fairly successful, it is only hinting at false positives for class 1.

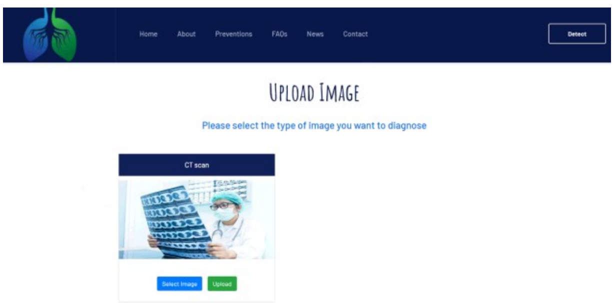

Fig:8 Web Interface to Upload the Image

Recall (0.88): In the aspects of recall, the model has 88% of recall rate to actual class 1 instance. Such a high recall also points to the strong ability of the classifier of identifying class 1 and preventing a large number of false negatives.

F1-Score (0.82): The score is quite fine because it will make a balance between precision and recall for class 1 which is quite important.

Fig:9 Web Framework Prediction

V. CONCLUSION

The new automatic system for recognizing Pulmonary Fibrosis from CT images should be seen as a major improvement to the diagnostic process which otherwise typically involves time-consuming and exhaustive interpretation of images by radiologists. Thanking to the selection of the InceptionV3, VGG16, and MobileNetV2 models for feature extraction, the identified pathology patterns are complex. The advantage of using the Random Forest and SVM classification models ensures that results are accurate and closer to actuality. The evaluation measures such as the accuracy, precision and F1 score proves that the proposed system works fairly well to identify pulmonary Fibrosis that makes it an effective diagnostic tool. Also, iincluding the system into a Flask-based web application is achieved in such a way that it provides easy interaction with the system as we can simply upload the CT images and get the result from the medical specialists. The model provides the current method through expanding the access to timely decision making, early interventional activities and positive patient outcomes and creating it as important tool in clinical contexts. The current hybrid model approach appears very promising to alter the detection and intervention of Pulmonary Fibrosis. The design is flexible and scalable to accommodate more patient cohorts and the framework can be utilized for other pulmonary diseases as well.

REFERENCES

[1] K. Mohanty et al., "Vivo Assessment of Pulmonary Fibrosis and Pulmonary Edema in Rodents Using Ultrasound Multiple Scattering," *IEEE Trans. on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 67, no. 11, pp. 2274-2280, Nov. 2020, doi: 10.1109/TUFFC.2020.3023611.

[2] R. A. Shehab, K. A. Apurba, M. Ahsanuzzaman, and T. Rahman, "Accurate Prediction of the Pulmonary Fibrosis Progression Using EfficientNet and Quantile Regression: Comprehensive mapping of key success factors for a high performing approach," *in Proc. IEEE-TENSYMP 2023*, Australia, pp. 1 - 6, 2023.

[3] S. Mandal, V. E. Balas, R. N. Shaw, and A. Ghosh, "Prediction Analysis of Idiopathic Pulmonary Fibrosis Progression from the OSIC Dataset," *in Proc. IEEE- GUCON 2020*, India, pp. 861-865, 2020, doi: 10.1109/GUCON48875.2020.9231239.

[4] W. Yu, H. Zhou, Y. Choi, J. G. Goldin, and G. H. J. Kim, "Mga-Net: Multi-Scale Guided Attention Models for an Automated Diagnosis of Idiopathic Pulmonary Fibrosis (IPF)," *in Proc. IEEE-ISBI'21*, FRANCE, pp. 1777-1780, 2021.

[5] X. Yu et al., “Detection of Idiopathic Pulmonary Fibrosis Lesion Regions Based on Corner Point Distribution,”*in Proc. IEEE-ICSP’22*, China, pp. 502-506, 2022, doi: 10.1109/ICSP54964.2022.9778477.

[6] Earle, M. E., & Al-Khassaweneh, M. , “Estimating Lung Capacity in Pulmonary Fibrosis Patients via Computerized Tomography (CT) Scan Data and Machine Learning,’ *in Proc. IEEE- ICIT*, pp.476-481,2021. doi:10.1109/ICIT52682.2021.9491120.

[7] Glotov, A. & Lyakhov, P., “Pulmonary Fibrosis Progression Prognosis Using Machine Learning,” *in Proc. IEEE-USBEREIT*, pp.327-329,2021. doi:10.1109/USBEREIT51232.2021.9455070.

[8] Mento, F., Soldati, G., Prediletto, R., Demi, M., & Demi, L, “Quantitative Lung Ultrasound Spectroscopy Applied to the Diagnosis of Pulmonary Fibrosis: The First Clinical Study,” *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 67(11), 2265-2273, pp.2020. doi:10.1109/TUFFC.2020.3012289.

[9] Mento, Francesco, Giuseppe Soldati, Renzo Prediletto et al, “Differentiation of Pulmonary Fibrosis by Means of Quantitative Lung Ultrasound Spectroscopy, Clinical Study in Humans,” *in Proc.IEEE-IUS’20*,pp.1-4,2020doi:10.1109/IUS46767.2020.9251740.

[10] Santiago-Fuentes, L. M., et al., “Cardiovascular and Respiratory Interactions in Idiopathic Pulmonary Fibrosis by Extended Partial Directed Coherence: Short-term Effects of Supplemental Oxygen.” *in Proc. IEEE- EMBS’21*, pp. 5446 – 5449, 2021.

[11] Carvalho, J. B. S., et al., “Domain Generalization for Diagnosis of Pulmonary Fibrosis Using Dose-Invariant Feature Selection,” *in Proc. IEEE-ISBI’23,* pp.1-5, 2023. doi:10.1109/ISBI53787.2023.10230504.

[12] R. A. Shehab, K. A. Apurba, M. Ahsanuzzaman, and T. Rahman, “Accurate Prediction of Pulmonary Fibrosis Progression Using EfficientNet and Quantile Regression: Comprehensive mapping of key success factors for a high performing approach,” *in Proc. IEEE-TENSYMP*, Canberra, Australia, pp. 1-6, 2023.

[13] A. Wong, J. Lu, A. Dorfman, P. McInnis, M. Famouri, D. Manary, J. R. H. Lee, and M. Lynch, “Fibrosis-Net: A tailored deep convolutional neural network design for prediction of pulmonary fibrosis progression from chest CT images,” *Fronties in AI*, pp.1-9, 2021, doi: 10.3389/frai.2021.764047.

[14] A. Khan, Z. Rauf, A. R. Khan, S. Rathore, et al, “A recent survey of vision transformers for medical image segmentation,” Electrical Engineering and Systems Science,Image and Video Processing, Cornell University, pp.1-59, 2023, doi:arXiv:2312.00634,.

[15] T. Moua, A. S. Lee, and J. H. Ryu, “Comparing effectiveness of prognostic tests in idiopathic pulmonary fibrosis," *Expert Rev. Respir. Med.*, vol. 13, no. 10, pp. 993-1004, Oct. 2019.