2025 3rd International Conference on Artificial Intelligence and Machine Learning Applications Theme: Healthcare and Internet of Things (AIMLA) | 979-8-3315-3896-5/25/$31.00 ©2025 IEEE | DOI: 10.1109/AIMLA63829.2025.11040221

Predicting Pulmonary Fibrosis Progression Using EfficientNet-LSTM Model: A Hybrid Approach

Akshara [a], Adith Sajeev [b], Sooraj Sunoj[c],Aswin C[d], Sunitha E.V[e]
Department of Computer Science and Engineering
Amrita School of Computing
Amrita Vishwa Vidyapeetham
Amritapuri,India

amenu4cse21007@am.students.amrita.edu[a], amenu4cse21071@am.students.amrita.edu[b], amenu4cse21055@am.students.amrita.edu[c], amenu4cse21014@am.students.amrita.edu[d] ,sunithaev@am.amrita.edu[e]

***Abstract*—The progressive lung condition known as idiopathic pulmonary fibrosis is defined by lung tissue scarring, which impairs lung function. Traditional models fail to accurately capture both structural change on CT scans and temporal trends in clinical variables and thus cannot be trusted for predictive accuracy. A good prediction method is needed to improve patient monitoring and treatment planning. This paper suggests a hybrid deep learning architecture that combines EfficientNet for the extraction of fibrosis features from DICOM CT scans with long short term memory networks to track the evolution of the disease over time. Compared to single CNN or LSTM models, the combination of spatial and temporal data increases the precision of FVC prediction. The models are trained and validated on the OSIC Pulmonary Fibrosis Progression dataset, which shows better performance and less prediction error. The results show how deep learning can be used to improve the prognosis of IPF and provide a more accurate method for monitoring the condition and planning the treatment.**

***Index Terms*—Pulmonary Fibrosis, EfficientNet, Long Short-Term Memory, Hybrid Model, DICOM Images, Forced Vital Capacity**

I. Introduction

Idiopathic Pulmonary Fibrosis is a chronic, progressive and invariably fatal interstitial lung disease where there is abnormal scarring of lung tissue and a resulting irreversible decline in pulmonary function. IPF runs a variable course and some individuals suffer a severe course while others remain relatively slow to progress. The Forced Vital Capacity, a prime spirometric parameter, is commonly used for the assessment of lung function and monitoring the progress of disease. Nonetheless, precise prediction of FVC decline continues to be an important clinical challenge. Conventional prediction approaches, such as statistical regression modeling and empirical clinical scoring systems, tend to disappoint by not being able to replicate the complex, nonlinear interaction between imaging biomarkers, physiological parameters, and patterns of disease progression. This deficiency underscores the necessity for sophisticated, data-driven methods with the potential to enhance predictive accuracy and assist medical professionals in making knowledgeable choices for early intervention. [1]

Recent advancements in deep learning and artificial intelligence (AI) have equipped us with the means to examine high dimensional medical data.Convolutional Neural Networks have proved to be incredibly successful in extracting useful features from radiological images, and Long Short-Term Memory networks are able to capture temporal dependencies in sequential data. Even though each of them is strong individually, most current methods are either image-centric or use only structured clinical data, and hence have poor predictive performance. There is a need for a more inclusive method to combine both imaging and physiological data for a balanced estimation of disease progress.

To address these limitations, this study offers a deep learning hybrid strategy involving EfficientNet, a highly optimized CNN model, to extract features from chest CT scans, combined with LSTM networks to learn time-varying clinical parameters. EfficientNet optimizes feature extraction by efficiently learning intricate patterns in high-resolution DICOM images, and LSTM networks process sequential clinical inputs to detect patterns in lung function decline [2]. By combining the two architectures, the proposed model is intended to improve FVC prediction accuracy and facilitate personalized estimates of disease progression.

This study is motivated by the need to improve predictive performance beyond standard methods, offering a more interpretable and reliable instrument for clinicians. The primary objectives are to design an EfficientNet LSTM hybrid model that combines multimodal medical information, perform a comparative evaluation relative to isolated deep learning models, and examine important imaging and physiological features that impact the development of pulmonary fibrosis. Using a multimodal learning approach, this study will aim to improve precision medicine for idiopathic pulmonary fibrosis by better early diagnosis, risk stratification, and patient management.By enabling more accurate and customized predictions of disease progression, the results of this study could improve clinical decision making, refine treatment strategies, and ultimately enhance patient outcomes

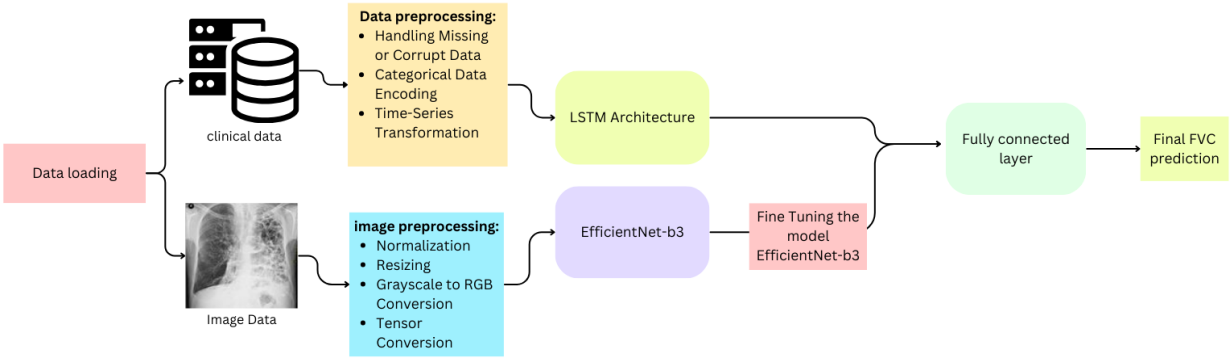

Fig. 1. Work Flow

II. Related work

The use of machine learning and deep learning ideas to forecast the course of idiopathic pulmonary fibrosis has advanced significantly in recent years. The rest of this chapter provides an overview of the contributions made to the subject, with an emphasis on models that have improved Forced Vital Capacity prediction by utilizing a range of regression techniques, neural networks, and hybrid approaches. By analyzing these methods, a general overview of the state-of-the-art can be given, which will help to identify areas that need further research and improvement.

Anh et al. introduce a hybrid methodology for forecasting forced vital capacity values in the evaluation of IPF progression using clinical data and CT scan images [3]This is based on novelty by using CNN-LSTM for feature extraction and LSTM-QRNN for improved prediction accuracy. The proposed model was found to be superior compared to other models with improved modified Laplace Log Likelihood scores on Kaggle dataset. Limitations are reliance on particular data that may not generalize very well to larger populations of patients.

Jain et al. suggested a method based on deep learning and machine learning to forecast FVC in order to diagnose pulmonary fibrosis (PF) early. The novelty of the paper lies in combining LSTM, Huber regression, and Support Vector Regression (SVR) models to predict FVC values from CT scans and FVC measurements. [4] SVR showed the best performance with an MSE of 0.029. This means that the number of patients, which is relatively small at 200, will limit the generalizability of the study.

Using the OSIC dataset, Mandal et al. examine and contrast several machine learning algorithms for predicting forced vital capacity in patients with idiopathic pulmonary fibrosis. [5] To increase prediction accuracy, several regression techniques including quantile regression, ridge regression, and elastic net regression are applied. This is where the uniqueness lies. The study shows the possibility of ML models in early diagnosis and resource-efficient healthcare management. However, it has limitations like the need for larger datasets and external validation for real-world deployment.

King et al, has highlighted its pathogenesis, clinical progression, and treatment options. The unique aspect, however, comes from identifying aberrant alveolar epithelial cell (AEC) activation as a cause of fibrosis as opposed to persistent inflammation. As a result, epithelial-mesenchymal transition, fibroblast proliferation, and extracellular matrix deposition may contribute to lung scarring. Further limitations include unknown mechanisms linking IPF to aging as well as limited available options for treatment. [6]

III. Proposed Methodology

This study applies a structured methodology for the evaluation and comparison of predictive performance in deep learning models. It involves a series of steps as illustarted in Figure 1, Accurate FVC prediction is critical in the monitoring of disease progression and thus improves patient management. The method is based on both image data from chest CT scans and clinical time series data to develop strong predictive models capable of predicting lung function accurately.

The methodology relies on the combination of various data modalities to capture the complex patterns and interactions for IPF progression shown in Figure 2. Advanced deep learning architectures will be used for the improvement of predictive accuracy and reliability of models. [7] The exhaustive evaluation includes the training and testing of multiple models including individual standalone models for each data type and hybrid models combining image and time series data. This would facilitate a proper comparison to identify which method is most effective for predicting FVC.

A. Data Collection

This study uses the OSIC Pulmonary Fibrosis Progression dataset, which is openly accessible on Kaggle. It contains high resolution computed tomography scans in DICOM format alongside structured clinical data in CSV format, providing a thorough summary of the course of the illness in patients with idiopathic pulmonary fibrosis. Age, gender, and smoking status are among the patient details covered by the clinical data along with primary medical indicators such as Forced Vital Capacity, an indicator of lung function, and percentage relative to expected Forced Vital Capacity based on patient types. The database also captures time information, tracing patient development across several weeks. The imaging data

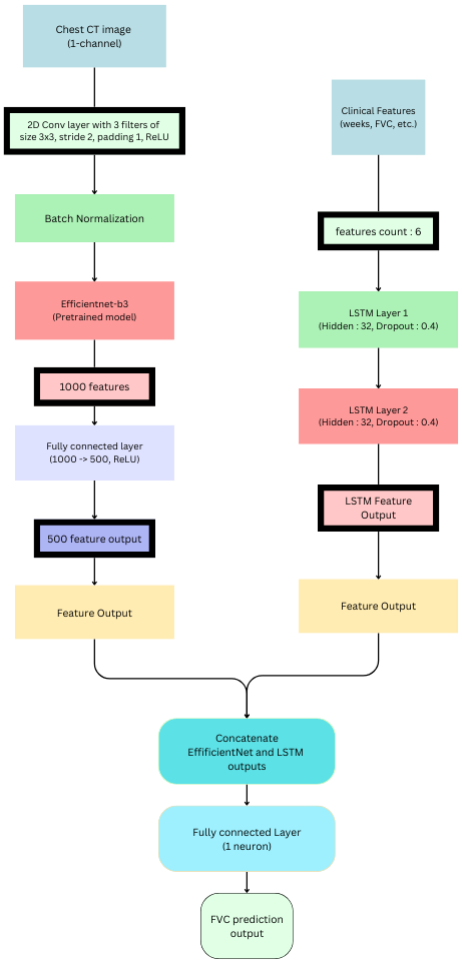

Fig. 2. System Architecture

are represented by a set of computed tomography scans reflecting structural lung impairment related to fibrosis and used as an important input for deep learning models to detect disease severity and patterns of progression. The combination of both clinical and imaging information offers the possibility of developing robust predictive models with numerical and visual features for improved prognostic accuracy. The data has been preprocessed to ensure consistency, including the removal of missing or redundant information, standardization of numerical variables, and encoding of categorical features to facilitate machine learning application. This structured and multi modal dataset is a strong foundation upon which to construct predictive models for the progression of idiopathic pulmonary fibrosis, facilitating more precise quantification of disease course.

B. Data Pre-Processing

Preprocessing of clinical data and both DICOM images was necessary in order to introduce consistency and reliability in training models. The pydicom library was used first to load DICOM images. It enabled retrieving pixel intensity values without losing necessary metadata. Considering raw pixel intensity values in CT scans can fluctuate based on acquisition parameters, normalization was introduced by using RescaleIntercept and RescaleSlope parameters to correct variations in intensity. This stage standardized pixel values prior to additional processing. [8] Subsequent to normalization, images were resized to a constant resolution of 224×224 pixels with OpenCV to preserve identical input sizes across the dataset.Because the CT scans were initially in grayscale form they were converted to RGB to be compatible with deep learning models that require colored images. For model compatibility, resized images were converted to PyTorch tensors to easily integrate into deep learning pipelines.

For preprocessing of clinical data, the dataset was initially imported with Pandas. Some patient records that generated an error upon opening DICOM files were found and removed subsequently to avoid discrepancies in the dataset. Since there were categorical features such as Sex and SmokingStatus, label encoding was applied to transform them into numeric formats in order to be machine learning model-compatible. Because of the nature of the data being time-series, patient records were converted into supervised learning structures by shifting past observations into a structured format that considers past values as predictors for future outcomes. [9] The transformation was achieved through a special function that restructured the data into lag-based representations, making the model learn temporal dependencies more effectively. Lastly, the 'Patient' column was also label-encoded to enable smooth processing while preserving patient-specific data. These preprocessing processes ensured that image and clinical data were properly structured and ready for idiopathic pulmonary fibrosis progression predictive modeling.

C. Model Architecture

The model built here combines EfficientNet-B3, utilized for image-focused feature extraction, with an LSTM network for sequential processing of clinical data. The hybrid model is designed specifically to maximize the prediction of the onset of pulmonary fibrosis. EfficientNet-B3, a leading convolutional neural network (CNN), is characterized by its efficient compound scaling, thus establishing a balance between depth, width, and resolution for improved feature extraction. The model is pre-trained with ImageNet weights first and then fine-tuned for the medical imaging application context. For the sake of facilitating the processing of grayscale computed tomography (CT) images, a convolutional layer with kernel size 3, stride 2, and padding 1 is utilized to convert single-channel images to three-channel formats conducive for EfficientNet utilization.Batch normalization is used to stabilize training and enhance generalization. EfficientNet-B3's classifier head is replaced and a fully connected layer with 500 neurons added to it followed by a ReLU activation function is added to include non-linearity. EfficientNet-B3's early convolutional layers are frozen to preserve pre-trained low-level feature extraction, and later layers are fine-tuned to learn fibrosis-specific patterns. This realignment makes the model more effective in disease-related visual feature extraction without

compromising computational efficiency. The LSTM module is trained on sequential clinical information and attains the ability to identify temporal dependency crucial in disease progression capture.

Six standardized clinical features are the input representing patient-specific parameters. There are two stacked LSTM layers with 32 hidden units each, which allow hierarchical learning of both short-term fluctuations and long-term trends in diseases. The dropout layer of 0.4 is placed in between the LSTM layers for avoiding overfitting and facilitating better generalization. Unlike usual dense networks, LSTM maintains long-range dependencies among data and is particularly well-positioned to determine pulmonary function shifts across time. The last state of the LSTM is taken out as a spatially compact description of the features of the time series. Extracted features using EfficientNet-B3 and LSTM are combined into a single, combined representation including spatial and temporal information. This combined feature vector is then fed into a fully connected output layer that yields one continuous value for the estimated Forced Vital Capacity (FVC). [10]

The Adam optimizer with a learning rate of 0.001 is applied to model training, and mean squared error (MSE) as Loss function is employed to reduce prediction differences and attain stability for regression operations. [11] With the use of deep spatial feature extraction via EfficientNet-B3 and sequential pattern learning via LSTM, the hybrid model enhances prediction accuracy, thus enabling more precise Follow-up on the evolution of pulmonary fibrosis.

D. Training Process

The training of the EfficientNet-LSTM joint model is designed to predict Forced Vital Capacity (FVC) evolution in Idiopathic Pulmonary Fibrosis (IPF) patients. The data is distributed in a way that 70% is assigned for training, and validation and testing each get 15%, to provide a balanced distribution that is favorable for the best generalization. The training subset is employed for model learning, the validation subset is for hyperparameter optimization and model performance evaluation [12], and the testing subset is the final assessment measure for determining the capacity of the model to generalize to new data.

The Mean Squared Error (MSE) loss function minimizes the squared differences between observed and predicted FVC values penalizing larger errors. It is mathematically represented as:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{1}$$

where y_i represents the true FVC value, $\hat{y}_i$ denotes the predicted value, and n stands for the overall number of samples

The Adam optimizer is used because of its learning rate adaptability and good weight update, using a learning rate of 0.001 and weight decay to avoid overfitting. The optimization update rules are given in the following equations: The update equations are:

$$v_t = \beta_2 v_{t-1} + (1-\beta_2) g_t^2 \tag{2}$$

$$\theta_t = \theta_{t-1} - \frac{\alpha}{\sqrt{v_t} + \epsilon} m_t \tag{3}$$

where g_t is the gradient at time t, and α is the learning rate. 30 epochs of training are performed with a batch size of 64. EfficientNet extracts high-level spatial features from CT scans, and LSTM processes sequential time-series data that capture FVC trends.

The training process goes through the following steps:

1) Forward Propagation: EfficientNet derives imaging features, and LSTM learns sequential dependencies.
2) Loss Calculation: MSE loss is computed between actual and predicted FVC values.
3) Backpropagation:Gradients are backpropagated and calculated to adjust the model weights.
4) Validation Check:The validation loss is tracked after each epoch.

To ensure generalizability, early stopping is utilized to terminate training if validation loss stalls in a specified number of epochs. Batch normalization and a 0.4 drop rate are also utilized to avoid overfitting and enhance model robustness. [13]

E. Data Handling

The data was loaded using custom PyTorch Dataset classes. These classes processed both image and time series data, which would ensure that each batch would contain both the image and corresponding time series data for each patient. Techniques such as random cropping, flipping, and rotation of the DICOM images were implemented to enhance the variety of data presented to the model in training and avoid overfitting. In order to improve the model performance when applied to non-sampled data, the data set was partitioned into training, validation, and test subsets. The partitioning was done in accordance with the stratified partitioning rules and thus ensured the patients were distributed evenly in the respective sets. It used PyTorch DataLoader to load the data in batches; this helps the models process efficiently and train faster on the GPU.

IV. Experiments and Results

The experimental setup used different computational resources and software tools required for model development and assessment.Table I captures the major hardware, software and tools employed during the experiments.

The models were assessed for their predictive validity using a test data set with emphasis on how well they could predict Forced Vital Capacity in patients with Idiopathic Pulmonary Fibrosis. The findings are presented in Table II, where the Root Mean Square Error (RMSE) and the Coefficient of Determination (R2) are two key measures of the degree of prediction accuracy. In addition, Figure 3 also presents a plot of predicted and actual FVC values, indicating the relationship between model predictions and actual values.

Category	Component	Details
Hardware	GPU Environment	Kaggle GPU (NVIDIA Tesla T4)
	RAM	16GB
Storage	Storage	Cloud-based storage for datasets and model artifacts
Software	Programming Language	Python 3.9
	Deep Learning Frameworks	PyTorch 1.9, CUDA
	Libraries	• **Data Handling:** NumPy, Pandas • **Image Processing:** OpenCV, PIL (Python Imaging Library) • **Optimization:** Torch • **Transform:** SciPy • **Visualization:** Matplotlib, Seaborn, Scikit-learn
	Development Environment	Jupyter Notebooks (Kaggle Kernels)
	Dataset Hosting	OSIC Pulmonary Fibrosis Progression Dataset (Kaggle)

TABLE I
EXPERIMENTAL SETUP

Models	RMSE	r^2
Standalone CNN	192.45	0.87
Standalone EfficientNet	180.50	0.88
Standalone LSTM	185.58	0.85
EfficientNet-LSTM	170.20	0.90

TABLE II
MODEL EVALUATION METRICS

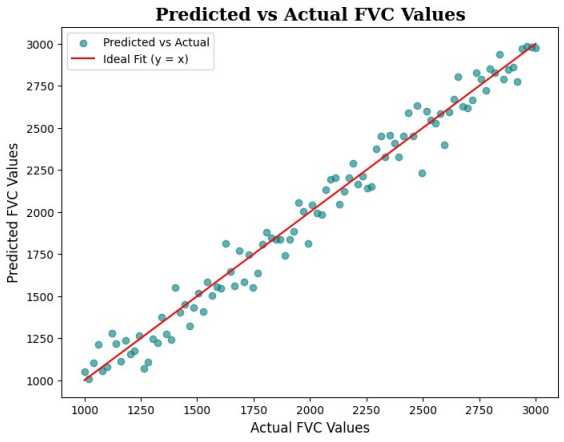

Fig. 3. Predicted vs Actual FVC

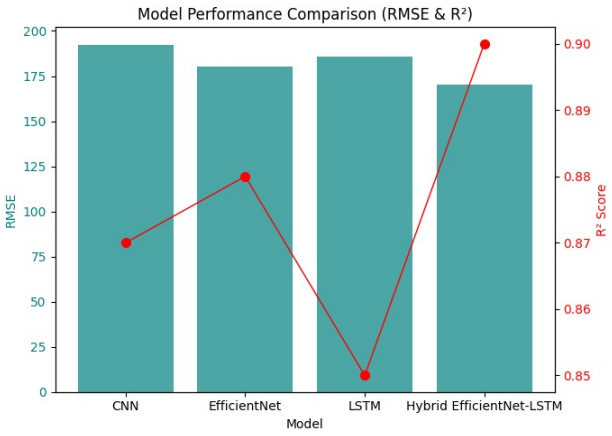

Fig. 4. Model Performance Comparison

These results imply that the EfficientNet-LSTM hybrid model outperforms the separate models, with the highest r^2 score (0.90) and the lowest RMSE (170.20mL)., implying better predictive accuracy. The CNN model, which captures spatial and texture-based features from CT scans, performs well with an r^2 score of 0.87. Likewise, the LSTM model, which handles sequential clinical data, performs with an r^2 a score of 0.85, very close to that of the CNN. Nonetheless, as these models handle different data modalities—CT scans for CNN and clinical time-series information for LSTM—a relative comparison among them is not entirely valid.

To provide fairness in assessment, the EfficientNet-LSTM hybrid model is used as a standard benchmark by combining both imaging and clinical features. The enhanced prediction accuracy seen in the hybrid method indicates that the integration of spatial and temporal information improves model robustness and generalization. [14] As indicated in Figure 3, the scatter plot of predicted vs. actual FVC values demonstrates the correspondence of model predictions with ground truth values. The red dashed line indicates the ideal y=x reference, with stronger clustering of points along this line showing improved predictive performance. The EfficientNet-LSTM model has the most compact distribution around the ideal line, indicating its high accuracy.

The findings further validate the importance of multimodal learning in IPF progression modeling Figure 4. Although individual models are successful in capturing modality-specific patterns, their failure to include complementary patient information constrains overall predictive performance. The EfficientNet-LSTM hybrid model, through the utilization of both imaging and clinical information, achieves better performance, highlighting the benefits of combining heterogeneous feature representations. [15]

These results indicate the promise of deep learning in improving IPF prediction. Future work can investigate other feature extraction methods, different architectures, and bigger datasets to more accurately improve predictive performance and clinical utility.

V. CONCLUSION

The intend of this study is to determine how efficiently different deep learning models predict Forced Vital Capacity (FVC) in individuals suffering from Idiopathic Pulmonary Fibrosis. The outcomes of the standalone models, Custom CNN,

EfficientNet, and LSTM, showed how crucial it is to take into account both image and time-series data in order to precisely forecast the course of a disease. All models were predictive, but the best one that used a hybrid of EfficientNet for image analysis and LSTM for time-series processing showed the highest performance among the other models with respect to the accuracy of RMSE and R^2.

Such an integration provides valuable information in increasing the predictivity performance for the diagnosis of IPF. Moreover, since it utilizes the information derived from structural details provided by chest CT scans as well as the temporal patterns available from clinical data, a deep learning method might improve the management of a patient and disease monitoring.

Future work includes further optimization of the architectures of the models used, addition of more clinical variables, and further refinement of model interpretability to be able to translate these results into clinical applications. Hybrid models, similar to what was proposed here, are very promising in advancing the prediction and tracking of IPF progression towards better patient outcomes.

REFERENCES

[1] L. Richeldi, H. R. Collard, and M. G. Jones, "Idiopathic pulmonary fibrosis," The Lancet, vol. 389, no. 10082, pp. 1941–1952, 2017. doi: 10.1016/S0140-6736(17)30866-8.

[2] Y. Shi, W. K. Wong, J. G. Goldin, M. S. Brown, and G. H. J. Kim, "Prediction of progression in idiopathic pulmonary fibrosis using CT scans at baseline: A quantum particle swarm optimization - Random forest approach," Artif. Intell. Med., vol. 100, p. 101709, 2019. doi: 10.1016/j.artmed.2019.101709.

[3] H. B. T. Anh, T. T. Dinh, L. T. Van, and H. Le Minh, "A combined CNN-LSTM and LSTM-QRNN model for prediction of idiopathic pulmonary fibrosis progression using CT scans and clinical data," in Proc. 2022 RIVF Int. Conf. Comput. Commun. Technol. (RIVF), Ho Chi Minh City, Vietnam, 2022, pp. 71–76. doi: 10.1109/RIVF55975.2022.10013925.

[4] D. Jain, P. Khurana, S. Yadav, and S. Sharma, "Prediction of pulmonary fibrosis disease," 2024 Control Instrumentation System Conference (CISCON), Manipal, India, 2024, pp. 1–6. doi: 10.1109/CISCON62171.2024.10696824.

[5] S. Mandal, V. E. Balas, R. N. Shaw, and A. Ghosh, "Prediction analysis of idiopathic pulmonary fibrosis progression from OSIC dataset," 2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON), Greater Noida, India, 2020, pp. 861–865. doi: 10.1109/GUCON48875.2020.9231239.

[6] T. E. King, A. Pardo, and M. Selman, "Idiopathic pulmonary fibrosis," The Lancet, vol. 378, no. 9807, pp. 1949–1961, 2011. doi: 10.1016/S0140-6736(11)60052-4.

[7] Anthimopoulos, M., Christodoulidis, S., Ebner, L., Christe, A., and Mougiakakou, S. (2016). Lung Pattern Classification for Interstitial Lung Diseases Using a Deep Convolutional Neural Network. IEEE Trans. Med. Imaging 35, 1207–1216. doi:10.1109/tmi.2016.2535865

[8] Taniguchi, H., Ebina, M., Kondoh, Y., Ogura, T., Azuma, A., Suga, M., et al. (2010). Pirfenidone in Idiopathic Pulmonary Fibrosis. Eur. Respir. J. 35, 821–829. doi:10.1183/09031936.00005209

[9] P. M. J, V. Bhushan, S. K. HR, S. K. G, and S. J, "Prediction of pulmonary fibrosis progression using CNN and regression," in Proc. 2021 3rd Int. Conf. Adv. Comput., Commun. Control Netw. (ICAC3N), Greater Noida, India, 2021, pp. 944–950. doi: 10.1109/ICAC3N53548.2021.9725730.

[10] S. Akarsh, S. Sriram, P. Poornachandran, V. K. Menon, and K. P. Soman, "Deep learning framework for domain generation algorithms prediction using long short-term memory," in Proc. 2019 5th Int. Conf. Adv. Comput. Commun. Syst. (ICACCS), Coimbatore, India, 2019, pp. 666–671. doi: 10.1109/ICACCS.2019.8728544.

[11] A. Tripathi, T. Singh, R. R. Nair, and P. Duraisamy, "Improving early detection and classification of lung diseases with innovative MobileNetV2 framework," IEEE Access, vol. 12, pp. 116202–116217, 2024. doi: 10.1109/ACCESS.2024.3440577.

[12] C. Dev, K. Kumar, A. Palathil, T. Anjali, and V. Panicker, "Machine learning-based approach for detection of lung cancer in DICOM CT image," in Ambient Communications and Computer Systems, Y.-C. Hu, S. Tiwari, K. K. Mishra, and M. C. Trivedi, Eds. Singapore: Springer, 2019, pp. 161–173. doi: 10.1007/978-981-13-5934-7_15.

[13] A. Iyer, H. Vyshnavi A. M., and K. Namboori P. K., "Deep convolution network-based prediction model for medical diagnosis of lung cancer - A deep pharmacogenomic approach: Deep diagnosis for lung cancer," in *2018 Second International Conference on Advances in Electronics, Computers and Communications (ICAECC)*, Bangalore, India, 2018, pp. 1–4, doi: 10.1109/ICAECC.2018.8479499.

[14] A. Tripathi et al., "A deep learning-oriented approach for lung CT augmentation: Leveraging U-Net and GAN architecture," in Proc. 2024 IEEE Recent Advances in Intelligent Computational Systems (RAICS), Kothamangalam, Kerala, India, 2024, pp. 1–7, doi: 10.1109/RAICS61201.2024.10690157.

[15] V. V. Kumar, D. Subham, G. Gopakumar, T. Singh, A. Tripathi, and P. Duraisamy, "Deep learning based pulmonary chest X-ray abnormalities identification," Kristu Jayanti Journal of Computational Sciences (KJCS), vol. 4, no. 1, pp. 1–13, 2025, doi: 10.59176/kjcs.v4i1.2429.