Rofiqul Alam Shehab, Kaysarul Anas Apurba, Md. Ahsanuzzaman, Dr. Tanzilur Rahman, IEEE

Prediction of Pulmonary Fibrosis Progression Using Machine Learning and Deep Learning Algorithms

***Abstract*—**Idiopathic pulmonary fibrosis is a progressive lung disease that usually worsens over time. Though this disease is incurable, early detection and proper diagnosis can help keep it under control. After the diagnosis of PF, patients usually have an average life expectancy of 3 to 5 years. Forced vital capacity (FVC) is one measure medical professionals can use to investigate this disease further. The deep learning models may assist in efficiently operating human resources and lower the costs associated with this fatal illness’s social and healthcare element. Here, we propose a model that can predict the progression of idiopathic pulmonary fibrosis. This is done in two steps where images and categorical data were used in the first step to create a model using transfer learning and the created model was subsequently used to generate a new data frame. The proposed model’s performance has been compared with various contemporary state-of-the-art deep learning models where it outperformed the existing literature for modified Laplace-Log-Likelihood (-6.64). Using this, medical professionals will be able to determine the severity of a patient's lung condition and the patients and their families will better understand the prognosis when they are first diagnosed with this disease.

***Keywords*—IPF, pulmonary fibrosis, neural network, computed tomography, chest CT, convolutional, pertained models, open source.**

# INTRODUCTION

"Predicting the progression of pulmonary fibrosis (PF) is crucial for the development of new treatments and for the management of patients with the disease. The existing methods of diagnosis and treatment of PF are often ineffective and the progression of the disease is difficult to predict. In this study, we propose to use EfficientNet, a state-of-the-art convolutional neural network (CNN) architecture and quantile regression (QR) to predict the progression of PF in patients, using the OSIC dataset available on Kaggle.

Pulmonary fibrosis is a chronic lung disease characterized by the formation of scar tissue in the lungs, which can lead to difficulty breathing and a reduced ability to oxygenate the blood. The progression of PF is difficult to predict, and current methods of diagnosis and treatment are often ineffective.

The progress of this disease is highly variable. The condition is sporadic and also runs in families. According to scientific research, various variables, including immunologic, environmental, and genetic, contribute to the emergence of the illness. The MUC5B gene variant, which increases mucus formation in the tiniest airways in the lung, is the most significant risk factor for developing IPF, accounting for 30% of the risk (respiratory bronchioles) [1]. Though IPF is incurable, it is treatable with some medicines (e.g., pirfenidone and nintedanib, and future therapeutic strategies are being investigated [2] [3]. Ultimately, some afflicted people would need a lung transplant.

The evaluation of lung function decrease is a crucial stage in treating and managing pulmonary fibrosis. Numerous methods for diagnosing pulmonary fibrosis follow the guidelines in an official ATS/ERS/JRS/ALAT clinical practice guideline journal [4], such as transbronchial lung biopsy, surgical lung biopsy, spirometry tests, etc. The transbronchial [5] and surgical biopsy [6] methods are used for diagnosing pulmonary fibrosis, and there are risks to the patient's health and lung function. Spirometry methods are well adapted to assess the decline in lung function after a diagnosis, as they are widely used to quantify the lung's FVC [7], a crucial marker of lung function. But, with only the spirometry test, it is challenging to understand the disease state, which severely limits its ability to anticipate the development of pulmonary fibrosis.

As medical technology has improved significantly over the last few decades, it has become helpful in making clinical decisions. Computed tomography (CT) imaging is one of the most valuable tools for determining the degree of lung damage caused by pulmonary fibrosis and the decline in lung function [8]. It is now clinically common practice to perform CT scans infrequently to better understand the underlying disease mechanisms and the course of the disease within the lungs. Radiologists have found and used many visual CT scan indicators to evaluate the reduction in lung function caused by pulmonary fibrosis. It might be challenging to predict how pulmonary fibrosis will develop because the pace of advancement in different patients can vary greatly. Therefore, developing novel techniques for enhancing presage precision is highly desirable when utilizing CT images to evaluate and forecast future lung function decline caused by pulmonary fibrosis.

Driven by recent improvements and the possibility of artificial intelligence techniques in the medical sector, machine learning, and deep learning can aid in this situation. On the other hand, the challenge [9] organized by Open Source Imaging Consortium (OSIC) in Kaggle has inspired us to implement a deep learning-based system. ​For example, Wong et al. [10] recently suggested Fibrosis-Net based on deep CNNs to predict the evolution of pulmonary fibrosis from chest CT scans. The FVC of a patient at a particular time-point in the future was indicated by Fibrosis-Net using clinical metadata, spirometry measurements, and chest CT images of a patient [10]. Following their work, Zabir Al Nazi et al. [11] have proposed another system named Fibro-CoSANet-based deep CNN and stacked attention layer [11]. We firmly feel there is still room for development in terms of overall correctness, even though the current CNN-based techniques have a more vital capacity to forecast the advancement of pulmonary fibrosis from chest CT images.

The authors of this paper, aim to predict a patient's severity of the decline in lung function based on a CT scan of their lungs, metadata, and baseline FVC as input. We want to predict the final three FVC measurements for each patient and a confidence value in our prediction. In light of this statement, we proposed an effective model for CT image processing that can distinguish human lungs with IPF and merge with the patient's existing metadata to produce FVC measures of the individual in the upcoming weeks. The rate of FVC decline, which is connected with the patient's survival, can be determined using this approach. The result of proposed model was compared with the existing state-of-the-art methods over the same metrics that shows promising outcomes.

The key contribution of this work are listed below:

1.

2.

3. We achieved a ……

As this system is not yet ready for production, we plan to release it as open source, which might encourage the general public and inspire future scholars to use and improve it.

# Related work

Predicting the progression of pulmonary fibrosis (PF) is crucial for the development of new treatments and for the management of patients with the disease. In recent years, there has been a growing interest in using machine learning and deep learning techniques to predict the progression of PF.

Yu Shi et al [12] was the first to show that it was possible to use only baseline HRCT scans to predict the progression of idiopathic pulmonary fibrosis using artificial intelligence. In their paper, they tried to develop a novel predictive model for the radiological progression pattern of idiopathic pulmonary fibrosis using only baseline HRCT scans. First, they implemented a study design and had an expert radiologist contour a region of interest (ROI) [13] at baseline scans, depending on the status of the ROI in follow-up visits. Then they integrated the feature selection with the prediction by developing an algorithm using a wrapper method that combines quantum particle swarm optimization to select a small number of features with random forest to classify early patterns of progression. They compared their result with other famous wrapper and non-wrapper methods, i.e., smoothly clipped absolute deviation (SCAD) [14], most minor absolute shrinkage and selection operator (LASSO), support vector machine (SVM), and neural network (NNET). Their proposed model yields an overall accuracy rate of 82.1%, which is superior to other feature selections and classification methods mentioned above.

The Kaggle Pulmonary Fibrosis Progression Challenge [9] recently demonstrated the potential and need for breakthroughs in artificial intelligence-driven solutions and computer-aided clinical decision support in pulmonary fibrosis. The Open Source Imaging Consortium (OSIC) [15] issued this challenge to encourage the scientific community to quicken the development of machine learning for pulmonary fibrosis evaluation. The patient information dataset compiled by OSIC for the challenge is the biggest publicly accessible dataset in the literature. A weighted ensemble of a deep convolutional neural network with a cutting-edge EfficientNet-B5 network architecture design and a multiple quantile regression was proposed in the first-place-winning solution [16]. The competition's winner used patients' CT scans, initial spirometry measurements, and clinical metadata to predict the lung function decline of a patient.

Authors of [17] performed prediction of pulmonary fibrosis on the same dataset as ours and their objective was to analyze and compare the performance of various machine learning models by predicting the final forced volume capacity measurements for each patient and a confidence value. Their main goal was to deploy the model on any computer to indicate a patient's severe condition regarding lung function, which is based on a CT scan of the lungs of the patients. The checked Lung function is based on a spirometer output that measures the lungs' forced vital capacity (FVC). They mainly followed three strategies to conduct the research: Multiple Quantile Regression, Ridge Regression, and ElasticNet Regression. They also discussed the results using those three strategies [17].

Wong et al. [10] constructed a customized network design tailored for predicting forced vital capacity (FVC). They conducted their experiment based on a patient's CT scan, initial spirometry measurement, and clinical metadata. They used machine-driven design exploration to identify a robust architectural design for CT lung analysis. Ultimately, they investigate the fibrosis-decision-making network's behavior and confirm that predictions are based on pertinent visual cues in CT images using an explain-ability-driven performance validation technique. They got the highest score from all previous experiments at that time when they conducted their investigation.

The Fibro-CoSANet model [11] combines the use of convolutional layers and self-attention mechanisms. Convolutional layers are commonly used in image analysis and can effectively identify essential features in medical images, such as CT scans. The self-attention mechanism allows the model to weigh different factors and identify essential patterns in the data, potentially leading to more accurate prognosis predictions. The Fibro-CoSANet model represents a significant advancement in pulmonary fibrosis prognosis prediction. The combination of convolutional and self-attention layers allows the model to effectively analyze patient data and make accurate predictions about disease progression. However, further research is needed to validate these results and test the model in more extensive and diverse patient populations to improve the model's generalizability.

Although a great deal of work has been completed, after doing extensive research it is found that there's still room for massive improvement. So, the goal was to provide a new and improved system that performed better than previous works. It is a system that might be extremely helpful for encouraging more clinical adoption of those solutions openly and reliably.

# Methodology

**Dataset*:*** The number of preprocessing steps used to prepare the training inputs and labels is discussed in this part. Kaggle has been used to collect data. OSIC Pulmonary Fibrosis Progression Challenge Benchmark Dataset, which includes CT scans, demographic information, and FVC measurements such as gender, age, and smoking status. Based on a patient's CT scan of their lungs, metadata, and baseline FVC as input. The csv file has 2270 rows and seven columns. The combined public and private test sets contain about 200 cases. 15 to 85 percent of this is public and private. This data set contains 176 distinct patient identifiers. The image folder contains 176 CT scan folders, with each file dedicated to a specific patient. Each folder contains a patient's entire CT scan history. Our goal is to estimate how severely their lung function will drop. For each patient, aim to forecast the last three FVC measures as well as a prediction confidence level.

**Preprocessing*:*** To generate pre-processed data, you must revisit the various image sizes. Data is divided into two size groups, with a few minor outliers. There were 132 patients who had underwent CT scans with 512 x 512-pixel images. There are 34 patients altogether in the photograph at this size, with the remaining ones being 1 or 2. Since the greatest piece of data is that size and have a total of 34000 images, each one must be resized to a fixed size of 512 x 512. So, producing the data will go more quickly. Attempted to base our conclusion on three characteristics, including age, sex, and smoking status.

The spectrum makeup of the x-rays relies on the measurement settings, such as acquisition parameters and tube voltage, so The CT-scan file must be converted to Hounsfield units.. [18] The images of various measures can be compared by normalizing to the values of water and air (water has HU 0 and air -1000). About 4000 of the grey values produced by a CT scan can't be seen by our eyes. This is the rationale for windowing. In this manner, the image is shown in a HU range that best suits the area of interest. The 3D pixel provided by a CT scan is known as a voxel. Our understanding is, it is spanned by the slice thickness in the z-direction and the 2d-plane of the pixel spacing property in the x- and y-direction. The dicom files' pixel spacing characteristic is a crucial one. It reveals the actual distance that one pixel covers. In the plane of a transversal slice, the x- and y-direction are described by only two values. This pixel spacing often remains constant across all slices for a single patient. However, according to individual or organizational preferences of doctors and the clinic, as well as depending on the type of scanner, the pixel spacing between patients can vary. Therefore, even if you compare two photographs of the same organ's size, the larger image does not necessarily reflect the actual size of the organ. Simply examining a slice's thickness, one may calculate how far a slice travels in the Z direction. Additionally, a specific area defined by row and column values is covered by the pixel array of raw values. More features can be seen in slices that are very thin. Thick slices, on the other hand, have less noise but are more susceptible to artifacts.

**System Architecture*:*** This project is finished in two stages. To create a model in the first section, which will combine some categorical data with image data. In the second step, utilize the model to generate a new data frame. With some

adjustments and calculations, create the final model to forecast the progression using that data set.

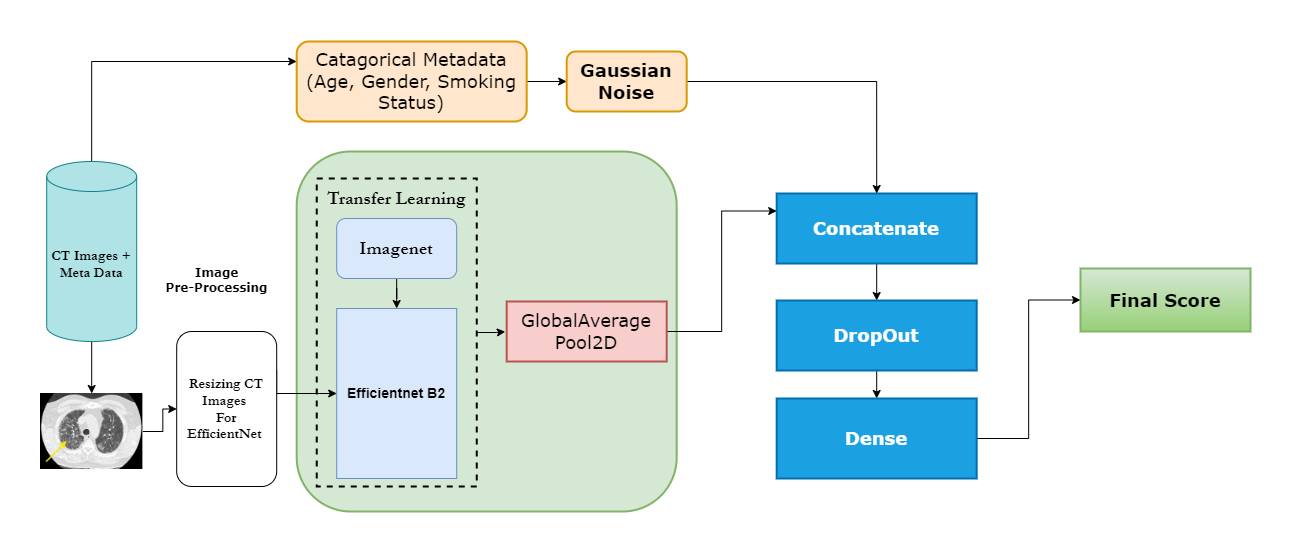


Figure 1- The Proposed System architecture for prediction of FVC. Given a patient’s CT scan, metadata (age, gender, smoking status) our system predicts the forced vital capacity of that patient at a specific week in future.

We gather every image and set it up for training in the first section. All of the image data should not be trained at once because doing so will generate an out of memory issue. In order to train a single piece of data at a time and avoid the error, we develop an image generator. After that, the model is built. In Figure 1, the model's architecture is displayed. Since all models' input layers are 224\*224, we start by getting a model and adding an input layer of size 512\*5121. Next, we set picture data as the input of the pre-trained model, and then we add a global average pooling 2D layer. Age, gender, and smoking status are categorical variables that we analyze by running them through a gaussian noise layer. [19]

We completed these two tasks simultaneously. Then, we combine the two, add a dropout layer, and feed the result through a thick layer. The training phase follows after that. During the training phase, early stopping is used. We employ it because an overfit model may arise from using too many epochs, whereas an underfit model may result from using too few. Early stopping is a technique that enables us to specify an arbitrary large number of training epochs and terminate training when the model's performance on a hold out validation dataset stops advancing. When a metric stops improving, we apply lowered learning rate on plateau to lower learning rate.

Once learning reaches a plateau, models frequently gain by decreasing the learning rate by a factor of 2–10. This callback keeps track of a quantity, and the learning rate is decreased if no progress is made after a specified number of epochs of "patience." We plot the training history once training is complete to show how well our model has learned.

In the second step of our proposed model, we've generated two variables, tr for train and chunk for test, for storing train and test sets. We have included the sample submission data into a variable called "sub" because we tend to forecast the fvc value with regard to week.

Table 1 Example Submission Table in OSIC Challenge

|  |  |  |
| --- | --- | --- |
| Patient\_week | FVC | Confidence |
| ID00419637202311204720264 -12 | 3111.484416 | 161.671271 |
| ID00421637202311550012437 -12 | 2857.890357 | 200.935648 |
| ID00422637202311677017371 -12 | 2007.297131 | 153.969624 |
| ID00423637202312137826377 -12 | 3420.942369 | 206.201271 |
| ID00426637202313170790466 -12 | 2974.510875 | 121.335843 |

Table 1 displays three features termed "Patient week," "fvc," and "confidence." Each entry in the "patient week" field is in concatenated form, as can be seen.

In order to save them, we divided them into "patient id" and "week" fields and added new features to the "sub" table. We have added a feature called "where" to the "tr," "chunk," and "sub" tables and initialized it with "train," "val," and "test," respectively.

Table 2 the table represents some patient's FVC value of First week.

|  |  |
| --- | --- |
| Patient\_ID | Start\_FVC |
| ID00007637202177411956430 | 2315 |
| ID00009637202177434476278 | 3660 |
| ID00010637202177584971671 | 3523 |
| ID00011637202177653955184 | 3326 |

Then combine them to form a single table called "data" from all of them. After that, added a new feature called "start week"

to the "data" table and assigned "nan" values (which stands for "not a number") to "start week" in accordance with the "test" rows. Following that, determined the arithmetic mean of the weeks for each patient and assigned a "start week" column to each patient based on that value. Then, to obtain the starting week fvc value, establish a new table named "base" in which only include the "start week" for each patient. This table is illustrated in (table 2). The initial weeks' fvc value for each table is contained in the "base" table. With regard to patient ID, merge them with the "data" table. Added a new feature called "base week" to the data table. For each entry, deduct "week" from "start week" and enter the result in "base week."

Table 3 Sample of Patient’s Meta data from OSIC Dataset

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Patient\_ID | Male | Female | Ex-smoker | Never smoked | Currently smokes |
| ID00419637202311204720264 | 1 | 0 | 1 | 0 | 0 |
| ID00421637202311550012437 | 1 | 0 | 0 | 0 | 1 |
| ID00422637202311677017371 | 0 | 1 | 0 | 1 | 0 |
| ID00423637202312137826377 | 1 | 0 | 0 | 1 | 0 |

Because the "base" table is no longer required, it can be deleted to free up RAM. Created those characteristics in the table and because "sex" and "smoking status" are categorical in nature, they were converted to boolean values. As you can see from (table 3) below, the corresponding entry is 1 if the patient is a male, otherwise 0, and the same is true for smoking status. The values for "age," "fvc," "week," and "percent" have now been normalized for each patient and added to the "feature" table. Conducted all of these analyses in a single dataframe to streamline the calculation. After that, divided them once more into train, validation, and test sets. Then, in order to train the new dataset displayed in (table 2**),** we had to design a new model. Used the categorical and numerical data that had been processed to form the model, and passed it through several dense layers. [20] After that, trained the model and used validation data to confirm its accuracy. The following section goes into detail on training, validation, and testing.

**Evaluation Metric: Laplace Log Likelihood**

To assess our proposed model, a modified version of the Laplace log likelihood will be used. It is helpful to assess a model's confidence in its conclusions when used in medical applications. [21] As a result, the measure is made to reflect both each prediction's accuracy and certainty. For each true FVC measurement, predict both an FVC and a confidence measure (which is standard deviation σ). The metric is determined by:

σclipped= max (σ,70),

Δ = min (|FVCtrue – FVCpredicted |, 1000),

Metric = - -ln (clipped) [17]

**Transfer Learning**

A model that has been trained for one task is repurposed for another activity that is related using the machine learning technique known as transfer learning [22]. When modeling the second task, it is an optimization that enables quick progress or enhanced performance. Transfer learning, however, is often used in deep learning due to the substantial resources needed to train deep learning models or the big and difficult datasets that deep learning models are trained on. Only general model features that were learned from the initial task can be used for transfer learning in deep learning. Inductive transfer is the name of the type of transfer learning that deep learning uses. This is where utilizing a model fit on a different but related job might help to reduce the range of potential models (model bias).

**VGG16**

The 16 layers that it has, as indicated in 4.1, are referred to as VGG16. It includes three completely connected layers and 13 convolutions. Input, convolution, pooling, fully linked layer, and softmax make up the total depth of 23, which is 23. ImageNet is the dataset that is used for testing, validation, and training. The photos are 256x256 in size and comprise 1.2 million training images, 50,000 validation images, and 150,000 testing images. VGG16 accepts images up to 224x224 in size. As a result, the photos were downscaled to 224x224, which is a fixed resolution. With 138.3 million parameters in all, this network is quite vast. In ImageNet, the model has a test accuracy of 92.7%. [23]

**ResNet-50**

The input size for ResNet is 224x224. About 25.6 million parameters make up the entire ResNet50 dataset. On the test image, ResNet50 had a 92.1% accuracy rate. Bottleneck layer is used in deeper networks, such ResNet50, to increase efficiency, maintains the same level of time complexity as the two layered convolutions. It enables us to both expand the number of layers and converge considerably more quickly. While VGG-16/19 nets have 15.3/19.6 billion FLOPS, 152-layer ResNet has 11.3 billion FLOPS. [24]

**Efficient Net**

The foundational EfficientNet-B0 network is built upon the MobileNetV2 inverted bottleneck residual blocks. In comparison to traditional scaling methods, this compound scaling strategy consistently increases model efficiency and accuracy when scaling up current models like ResNet (+0.7%) and MobileNet (+1.4% imagenet accuracy). comparing EfficientNets to other CNNs that are already running on ImageNet. In general, the EfficientNet models outperform previous CNNs in terms of accuracy and efficiency, reducing parameter size and FLOPS by a factor of two. For instance, EfficientNet-B7, which is 8.4 times smaller and 6.1 times quicker on CPU inference than the preceding Gpipe, achieves state-of-the-art 84.4% top-1 / 97.1% top-5 accuracy on ImageNet in the high-accuracy regime. EfficientNet-B4 consumes comparable FLOPS to the popular ResNet-50 while increasing top-1 accuracy from ResNet-50's 76.3% to 82.6% (+6.3%). [25]

**Result Analysis:**

By fully training our neural network, the process has begun. Create a dictionary of models that includes the following models in this case, including eight different types of EfficientNet models, ReNet50, and VGG16. Right now, train the various models from scratch without using any pre-trained weight, using simply their architecture. Since the input layer size for all models is 224\*224, obtain a model and add an input layer of size 512\* 512 \*1. Then, at the very end of the model design, included a pooling layer. Dropout layer is employed. In order to avoid overfitting, the Dropout layer randomly sets input units to 0 with a frequency of rate at each step during training. Make use of a 30 epoch size, a batch size of 8, and a learning rate of 0.003. The best model was saved using Model Checkpoint. The model with the lowest validation loss is the best one. When validation loss has stalled, also utilized Reduced on platow to slow down learning. If improvement is not made for five consecutive epochs, learning rate will be halved. Six distinct models, including four different types of EfficientNet (B0, B1, B2, and B4), have been trained. Moreover, ResNet50 and VGG16. The parameters and architecture of each of these models vary.

To enable comparison, trained each of these models using the identical set of hyper parameters. The lowest training loss is 3.6524 for EfficientNet B0. The lowest train loss for EfficientNet B1 were 4.2689. Lowest train loss for ResNet50 was 4.399539. The train loss for all 6 models is nearly the same, as shown in (Table 1). However, the training loss is a little bit smaller for ResNet50 and VGG16. However, discover that ResNet50 and VGG16 are overfitting when test the model. Since our dataset is not that vast and ResNet50 and VGG16 have complex architectures that necessitate a lot of training data, there is a chance that the data will be overfit. (Table 1) displays the training accuracy for each of the six models over a 100-epoch period. However, you cannot see any variances in training loss when looking at each graph in figure 5.3. This is due to the relatively minor difference between training loss. Have illustrated the training loss and validation loss of these models on smooth curves in figures 3, respectively, to help alleviate the issue.

On a modified version of the Laplace-Log-Likelihood mentioned in (Figure 3), have assessed the performance of our trained model. It is helpful to assess a model's confidence in its conclusions when used in medical applications. The confidence values are clipped at 70 ml to reflect the approximative measurement uncertainty in FVC, and the error is thresholded at 1000 ml to prevent huge mistakes from unfavorably affecting results. Averaging the statistic over all test set Patient Weeks yields the final score (three per patient). Keep in mind that metrics will have negative values, and higher is better. To understand the metric and use it with various instances, built the laplace log likelihood function. Find that when cross-validating models using train data, -8.023 is the default score to surpass. Any model that performs worse than this is useless.

**Table 4: Shows Comparison of training loss of All Models we tried to implement**

|  |  |  |
| --- | --- | --- |
| ***Model Name*** | ***Training Loss*** | ***Number of epoch*** |
| EfficientNet B0 | 3.6524 | 100 |
| EfficientNet B1 | 4.2689 | 150 |
| EfficientNet B2 | 3.8340 | 100 |
| EfficientNet B3 | 3.6728 | 100 |
| EfficientNet B4 | 4.1240 | 100 |
| ResNet 50 | 4.399539 | 100 |
| VGG16 | 4.187896 | 100 |

In this stage, the model has been trained across 100 epochs except B1 and has undergone five-fold cross validation. When it comes to validation, Attempt to get as close to the kaggle competition's scoring system as possible. They initially only assessed the last three forecasts (last three weeks for each patient), As a result, a framework for comparable validation has been developed. Only patients who were absent from the training set were included in the test set, and only the three most recent weeks were used for grading. The highest score received in the evaluation was -6.6468.

(Table 2) compares our results to those in the study, as well as to the public leaderboard on Kaggle. In order to determine whose score is higher, the comparison table has been arranged in ascending order. As you can see, our performed better than the competition and previously published work.

**Table 5: Comparison of Laplace-Log-Likelihood scores for all the existing works and this paper on the cohort from OSIC Challenge, Best Result is highlighted bold**

|  |  |
| --- | --- |
| ***Model Name*** | ***Scores*** |
| Our Model (EfficientNet B2 + Quantile Regression) | **-6.64** |
| Fibro-CoSANet ( [11]) | -6.68 |
| Elastic Net Regression ( [17]) | -6.73 |
| Fibrosis-Net( [10]) | -6.8188 |
| Kaggle 1st place [9] | -6.81 |
| Kaggle 2nd place [9] | -6.83 |
| Kaggle 3rd place [9] | -6.83 |
| Ridge Regression ( [17]) | -6.81 |
| Multiple Quantile Regression [17] | -6.92 |

  Comparison Table of All Scores

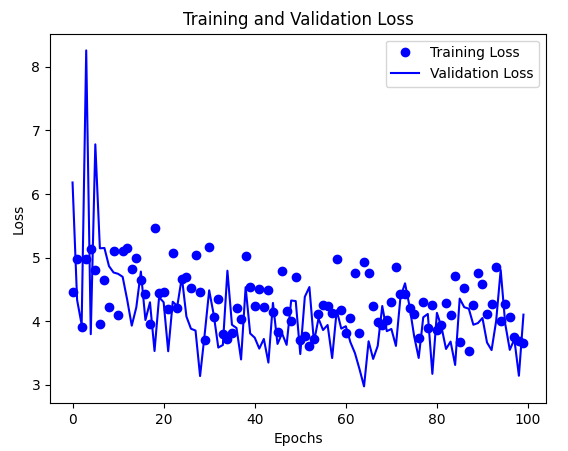


Figure 3- Showing the Training and validation loss of EffitientNet B0

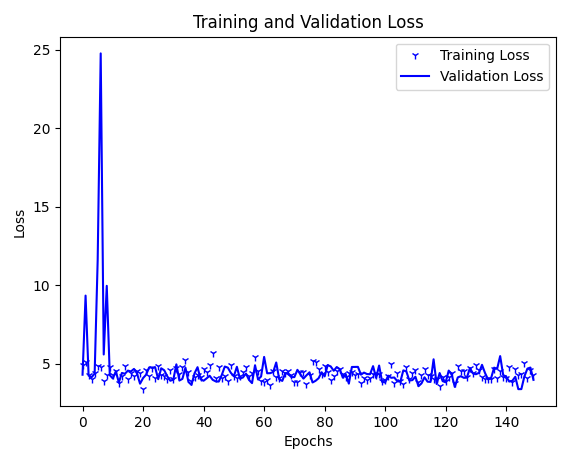


Figure 4-- Showing the Training and validation loss of EffitientNet B1

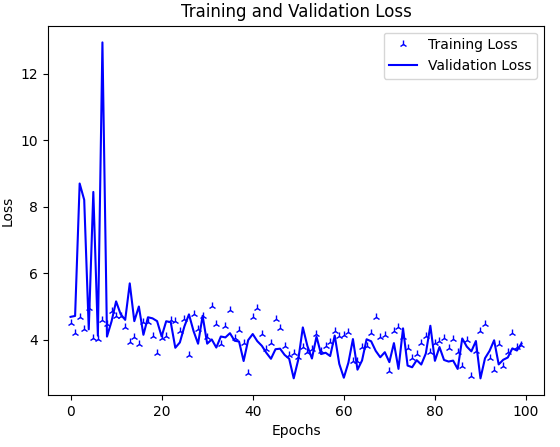


Figure 5-- Showing the Training and validation loss of EffitientNet B2

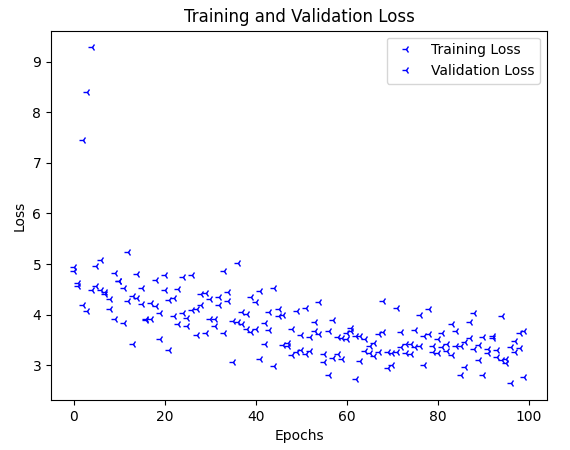


Figure 6-- Showing the Training and validation loss of EffitientNet B3

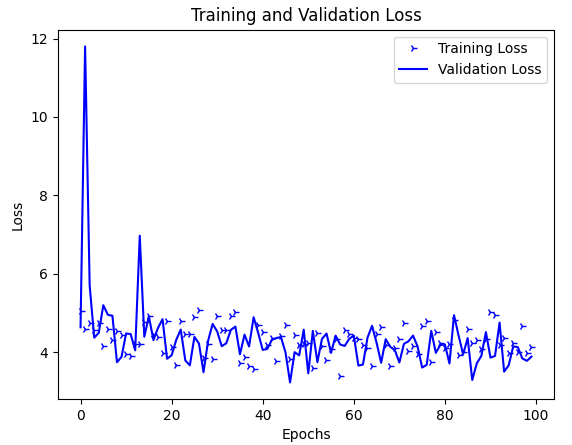


Figure 7-- Showing the Training and validation loss of EffitientNet B4

From the (Figure 4-7) we are showing training and validation loss of our trained models.

# Conclusion

Our proposed work has used metadata and CT-scan images to predict FCV and measured its performance using an evaluation matrix called Laplace-Log-Likelihood (Figure 3). Our achieved score is -6.64, which is the best of all other existing models' performances. From our experiment, we can state that high-resolution CT can be used with our proposed deep learning techniques and can provide a low-cost, efficient, fast, and suitable way to find the decline in the lung function of a patient suffering from idiopathic pulmonary fibrosis. This will help not only the patient but also the clinical specialist to make better decisions faster manner.

# References

|  |  |
| --- | --- |
| [1] | M. David A. Schwartz, "Idiopathic Pulmonary Fibrosis," Department of Medicine and Department of Immunology, University of Colorado School of Medicine,, 1987, 1989, 1999, 2000, 2004, 2007, 2014, 2017. [Online]. Available: https://rarediseases.org/rare-diseases/idiopathic-pulmonary-fibrosis/. |
| [2] | M. E. Y. K. T. O. A. A. M. S. Y. T. H. T. K. N. A. S. M. T. G. R. S. K. T. N. H. Taniguchi, "Pirfenidone in idiopathic pulmonary fibrosis," *Eur. Respir. J. 35, ,* p. 821–829., 2022. |
| [3] | L. d. B. R. M. R. G. A. A. B. K. K. C. U. e. a. Richeldi, "Efficacy and safety of nintedanib in idiopathic pulmonary fibrosis," *N. Engl. J. Med. ,* pp. 370, 2071–2082, 2014. |
| [4] | G. R.-J. M. M. J. L. R. L. R. C. J. L. D. J. Raghu, "Diagnosis of Idiopathic Pulmonary Fibrosis. An Official Ats/ers/jrs/alat Clinical Practice Guideline," *Am. J. Respir. Crit. Care Med,* no. doi:10.1164/rccm.201807-1255st, pp. 198, e44–e68., 2018. |
| [5] | L. C. H. R. a. J. M. G. Richeldi, "Idiopathic Pulmonary Fibrosis.," *The Lancet,* no. doi:10.1016/s0140-6736(17)30866-8, pp. 389, 1941–1952, 2017. |
| [6] | S. W. A. U. C. U. C. A. C. T. V. R. G. Tomassetti, "Bronchoscopic Lung Cryobiopsy Increases Diagnostic Confidence in the Multidisciplinary Diagnosis of Idiopathic Pulmonary Fibrosis," *Am. J. Respir. Crit. Care Med.,* no. doi:10.1164/rccm.201504-0711oc, pp. 193, 745–752, 2016. |
| [7] | L. C. K. T. E. S. M. I. W. J. A. S. R. E. a. C. R. M. Watters, "A Clinical, Radiographic, and Physiologic Scoring System for the Longitudinal Assessment of Patients with Idiopathic Pulmonary Fibrosis.," *Am. Rev. Respir. Dis,* no. doi:10.1164/arrd.1986.133.1.97, pp. 133, 97–103, 1985. |
| [8] | M. S. R. M. R. A. K. T. M. M. S. R. P. F. M. M. P. A. D. a. R. A. R. P. Brian J Bartholmai, "Quantitative CT Imaging of Interstitial Lung Diseases," *J Thorac Imaging.,* no. doi: 10.1097/RTI.0b013e3182a21969. , pp. 298-307, 2013. |
| [9] | O. S. I. C. (OSIC), "OSIC Pulmonary Fibrosis Progression-Predict lung function decline," 2019. [Online]. Available: https://www.kaggle.com/c/osic-pulmonary-fibrosis-progression. |
| [10] | J. L. A. D. P. M. M. F. D. M. J. R. H. Alexander Wong, "Fibrosis-Net: A Tailored Deep Convolutional Neural Network Design for Prediction," *Frontiers in Artificial Intelligence,* vol. 4, no. DOI=10.3389/frai.2021.764047, 2021. |
| [11] | F. R. M. ,. M. A. I. ,. S. S. Zabir Al Nazi, "Fibro-CoSANet: Pulmonary Fibrosis Prognosis Prediction using a Convolutional Self Attention Network," *Physics and Engineering in Medicine,* no. DOI: 10.1088/1361-6560/ac36a2, 2021. |
| [12] | W. K. W. 1. J. G. G. 2. M. S. B. 2. G. H. J. K. Yu Shi 1, "Prediction of progression in idiopathic pulmonary fibrosis using CT scans at baseline: A quantum particle swarm optimization - Random forest approach," *In: Artificial intelligence in medicine ,* no. DOI: 10.1016/j.artmed.2019.101709, p. p. 101709, 2019. |
| [13] | S. Y. Y. W. W. W. Kim GHJ, "A study design for statistical learning technique to predict radiological progression with an application of idiopathic pulmonary fibrosis using chest CT images.," *Contemp Clin Trials,* no. doi: 10.1016/j.cct.2021.106333., 2021 . |
| [14] | H. S. R. A. M. R. A. Abolfazl Mehranian, "Smoothly Clipped Absolute Deviation (SCAD) regularization for compressed sensing MRI Using an augmented Lagrangian scheme, Habib Zaidi," *Magnetic Resonance Imaging,* vol. 31, no. 8, pp. 1299-1411, 2013. |
| [15] | OSIC, "Open Source Imaging Consortium (OSIC)," [Online]. Available: https://www.osicild.org/. |
| [16] | A. Kulakov, "How I achieved the 1st place in Kaggle OSIC Pulmonary Fibrosis Progression competition," Medium, 7 Oct 2020. [Online]. Available: https://towardsdatascience.com/how-i-achieved-the-1st-place-in-kaggle-osic-pulmonary-fibrosis-progression-competition-e410962c4edc. [Accessed 19 January 2022]. |
| [17] | S. Mandal, V. E. Balas, R. N. Shaw and A. Ghosh, "Prediction Analysis of Idiopathic Pulmonary Fibrosis Progression from OSIC Dataset," in *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON),*, Greater Noida, India, 2020. |
| [18] | M. N. a. F. A. G. Tahmineh Razi, "Relationship between Hounsfield Unit in CT Scan and Gray Scale in CBCT," *J Dent Res Dent Clin Dent Prospects,* p. 107–110, 2014 Jun 11. |
| [19] | J. Brownlee, "How to Improve Deep Learning Model Robustness by Adding Noise," Machine Learning Mastery, 14 December 2018. [Online]. Available: https://machinelearningmastery.com/how-to-improve-deep-learning-model-robustness-by-adding-noise/. [Accessed 20 january 2023]. |
| [20] | D. Kaplan, "Dense Layer: The Building Block To Neural Networks," EML, 7 June 2022. [Online]. Available: https://enjoymachinelearning.com/blog/dense-layer/. [Accessed 20 January 2023]. |
| [21] | F. &. O. C. Agu, "Modified Laplace Distribution, Its Statistical Properties and Applications," *Asian Journal of Probability and Statistics,* vol. 4, pp. 1-14, 10 June 2019. |
| [22] | S. J. P. a. Q. Yang, "A survey on transfer learning," *IEEE Transactions on Knowledge and Data Engineering,* vol. 22, no. 10, pp. 1345 - 1359, 16 October 2009. |
| [23] | K. S. a. A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv,* 2014. |
| [24] | K. He, X. Zhang, S. Ren and J. Su, "Deep Residual Learning for Image Recognition," *Proceedings of,* p. 770–778, 2016. |
| [25] | M. T. a. Q. V. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," *arXiv,* 2019. |