Improving Idiopathic Pulmonary Fibrosis Damage Prediction with Segmented Images in a Deep Learning Model

1st Sheila Leyva-López
Faculty of Engineering
Autonomous University of Queretaro
Queretaro, Mexico
0000-0003-3634-8351

2nd Gerardo Hernández-Nava
Faculty of Engineering
Autonomous University of Queretaro
Queretaro, Mexico
0000-0002-8685-1003

3rd Enrique Mena-Camilo
Faculty of Engineering
Autonomous University of Queretaro
Queretaro, Mexico
0000-0001-8353-9017

4th Sebastián Salazar-Colores
Centro de Investigaciones en Óptica
León, México
0000-0002-6353-0864

2023 IEEE Conference on Artificial Intelligence (CAI) | 979-8-3503-3984-0/23/$31.00 ©2023 IEEE | DOI: 10.1109/CAI54212.2023.00078

***Abstract*—This work introduces a semantic segmentation model, UNet, as a preprocessing module to an algorithm predicting lung damage caused by Idiopathic Pulmonary Fibrosis. By modifying the model input through the incorporation of a guide image (a segmentation result) into the original image, we observed an improved performance of eight out of twelve tested backbones in the prediction model, with an improvement of up to 0.57 in the LLL_m metric. This study underscores the significance of data preprocessing in deep learning models' performance. The inclusion of additional data, such as segmented images, can significantly enhance a model's ability to perform specific tasks, emphasizing the need for careful data preprocessing to obtain precise and reliable results when implementing deep learning models for lung damage prediction.**

***Index Terms*—Segmentation, Deep Learning, Idiopathic Pulmonary Fibrosis, Computed Tomography.**

I. Introduction

Idiopathic Pulmonary Fibrosis (IPF) is a chronic, progressive, and debilitating disease that is primarily characterized by persistent coughing that evolves over several months, often preceded by an acute respiratory illness [1]. To date, IPF remains an incurable disease. However, an accurate and timely diagnosis can significantly contribute to reducing the mortality associated with this pathology and improving the life expectancy and quality of life of affected patients. It is worth noting that, after the onset of the first symptoms, the average survival interval ranges between 2 and 5 years [1] [2].

The main objective of this work is to improve the prediction of lung damage caused by IPF by implementing a UNet semantic segmentation module. This is a preprocessing block for the network proposed in [3], which is an end-to-end multimodal learning-based approach that predicts the decline of forced vital capacity (FVC) in patients with IPF, using Computed Tomography (CT) images and demographic information in convolutional neural network (CNN) frameworks with a stacked attention layer.

To evaluate the models, two main metrics are considered, Laplace Log-Likelihood (LLL_m) (Equation 1), which measures a model's ability to correctly predict the classes or values of a test dataset, using logarithmic probability as a measure of error. One advantage of this metric is that it avoids problems of overfitting and underestimating the model's accuracy. Moreover, it is especially useful when working with small or imbalanced datasets [4]. The root mean squared error ($RMSE$) (Equation 2) is also used, which evaluates the accuracy of a regression model. $RMSE$ measures the average squared difference between the actual values and the predicted values by the model [5].

$$LLL_m = log(P + \frac{1}{N+K}) \tag{1}$$

where:

LLL_m es la Laplace Log-Likelihood. P stands for the number of correct predictions made by the model on the test dataset. N is the size of the test dataset. K is the number of classes or possible values.

$$RMSE = \sqrt{\frac{\Sigma(y - y_{pred})^2}{n}} \tag{2}$$

where:

n represents the number of observations in the dataset. y is the true value of the response variable. y_{pred} is the predicted value of the response variable.

II. Database

The database provided in [6] comprises 176 training subjects and 5 test subjects. However, for this study, we decided to only utilize the training set as the remaining subjects are unlabeled. The segmentation of this training set is described in the methodology section.The main objective is to predict the last three forced vital capacity (FVC) measurements, as well

979-8-3503-3984-0/23/$31.00 ©2023 IEEE
DOI 10.1109/CAI54212.2023.00078

as a confidence value in the prediction. The data contained in the database consists of a chest CT scan taken at week zero of follow-up and clinical parameters, which were collected in follow-up sessions over one or two years as appropriate.

The clinical parameters that make up the database include: a unique patient identifier, the relative number of weeks before or after the reference CT, the lung capacity recorded in ml, a calculated field that approximates the patient's FVC as a percentage of the typical FVC for a person with similar characteristics, age, sex, and smoking status.

However, the test dataset provides a reference CT and only one initial measurement of FVC.

III. Methodology

As a first step, a central slice of the CT scan is randomly selected, omitting 15% of the initial and central slices. As the images are in DICOM format, they are converted to PNG format with dimensions of $256 \times 256 \times 1$ in order to optimize the use of graphic computing memory. Subsequently, this array is processed by a model from the library *Segmentation Models PyTorch*, which consists of a UNet that uses the DenseNet121 network architecture as the encoder. A value of one is assigned to both the number of classes and the number of input channels, suggesting that only one object is being segmented in the grayscale input image. Once the model is compiled, transfer learning is performed with the weights available in [7]. As a result of the evaluation of the input in the previous model, a mask is obtained, where the lungs of the CT slice are segmented.

Now that the mask is available, it is combined with the original input of the architecture and then processed by the network described in [3].

Finally, different encoder functions are tested within the Fig. 1 module, including EfficientNet, ResNet, and ResNeXt, with their respective variations. Finally, the LLL_m and $RMSE$ metrics, which are indicators of the reliability of the predictions, are calculated to perform a quantitative comparison between the experiments. It should be noted that the evaluations and calculated metrics were obtained using the 5-Fold technique, this implies that an 80% training and 20% testing data split was performed on the dataset used, with this process being repeated 5 times while varying the elements in the sets.

IV. Results and discussions

Based on the results reported by [3], the most significant improvement was found when using the ResNet34 encoder function with a value of 0.57 in the LLL_m metric, as shown in Table I, it should be noted that all metrics reported in the table were calculated using the test sets of the 5-fold cross-validation. However, this was not the only result where an improvement was found, as there was an improvement in eight out of the twelve encoder functions used, with a minimum improvement of 0.01.

Furthermore, when considering the results obtained in the replication of the [3] model, the improvements ranged from (0.06,0.36) in seven out of the twelve models.

Regarding the RMSE metric, there was a slight increase, except for four out of the twelve encoder functions tested.

In Table I, when referring to the Replicated Model, it is meant the Fibro-CoSANet model replicated locally, with the difference that the dicom images were converted to png format beforehand, and not simultaneously, that is, not during the training of the network. It should be noted that the reported value of LLL_m corresponds to its absolute value. The larger the difference from the guided image experiment, the better.

V. Conclusions

Based on the results obtained, which indicate a favorable difference when using a guided image at the input of the prediction model, it can be concluded that the inclusion of additional data, such as masks obtained through segmentation models, improves the input data to models and thus obtains precise and reliable results in predicting idiopathic pulmonary fibrosis damage. These can be used as preprocessing modules for future work.

Acknowledgment

The authors wish to thank Consejo Nacional de Ciencia y Tecnología (CONACYT) for their support.

References

[1] A. Xaubet, J. Ancochea, and M. Molina-Molina, "Idiopathic pulmonary fibrosis," *Medicina Clínica (English Edition)*, vol. 148, no. 4, pp. 170–175, 2017. [Online]. Available: https://doi.org/10.1016/j.medcle.2016.11.035

[2] G. DS, G. D, R. HA, A. P, S. P, D. J, and R. AB, "Idiopathic pulmonary fibrosis: Current and future treatment," *The Clinical Respiratory Journal*, vol. 16, no. 2, pp. 84–96, 2021. [Online]. Available: https://doi.org/10.1111/crj.13466

[3] Z. A. Nazi, F. R. Mashrur, M. A. Islam, and S. Saha, "Fibro-cosanet: pulmonary fibrosis prognosis prediction using a convolutional self attention network," *Physics in Medicine & Biology*, vol. 66, no. 22, 2021. [Online]. Available: https://dx.doi.org/10.1088/1361-6560/ac36a2

[4] Y. Tian, "Lung function decline predicting using improved efficientnet," *2022 7th International Conference on Intelligent Computing and Signal Processing (ICSP)*, pp. 924–927, 2022. [Online]. Available: https://doi.org/10.1109/ICSP54964.2022.9778391

[5] T. Chai and R. R. Draxler, "Root mean square error (rmse) or mean absolute error (mae)? – arguments against avoiding rmse in the literature," *Geoscientific Model Development*, vol. 7, no. 3, pp. 1247–1250, 2014. [Online]. Available: https://gmd.copernicus.org/articles/7/1247/2014

[6] O. S. I. C. (OSIC), "Osic pulmonary fibrosis progression," Open Source Imaging Consortium (OSIC), 2020, Visited: October 2022. [Online]. Available: https://www.kaggle.com/competitions/osic-pulmonary-fibrosis-progression/overview

[7] HFUT_YBX, "Unet-densenet121 lung of segmentation," Kaggle, 2020, Visited: October 2022. [Online]. Available: https://www.kaggle.com/code/hfutybx/unet-densenet121-lung-of-segmentation/notebook

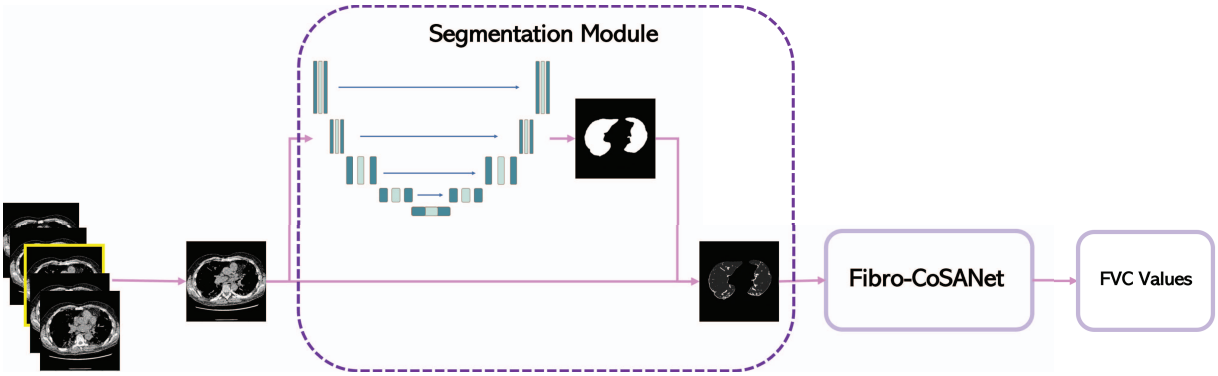

Fig. 1. Implementation of the segmentation module to the Fibro-CoSANet model

TABLE I
COMPARISON OF THE METRICS REPORTED BY THE AUTHORS OF FIBRO-COSANET, THOSE OBTAINED AFTER REPRODUCING THE SAME EXPERIMENT, AND AFTER MODIFYING THE INPUT TO THE NETWORK WITH A GUIDED IMAGE.

f-encoder	**Experiment**	$RMSE$			LLL_m		
		Mean	***Std***	***Difference from guided image***	***Mean***	***Std***	***Difference from guided image***
Efb0	Fibro-CoSANet	183.70	23.55	-1.38	6.70	0.29	-0.29
	Replicated Model	183.47	20.28	-1.51	6.71	0.28	-0.20
	Guided Image	186.24	21.27	-	6.72	0.28	-
Efb1	Fibro-CoSANet	183.96	22.89	-1.34	6.68	0.28	-0.71
	Replicated Model	184.99	0.26	-0.78	6.73	0.26	-0.03
	Guided Image	186.43	19.89	-	6.73	0.25	-
Efb2	Fibro-CoSANet	181.50	25.88	-1.34	6.68	0.31	-0.29
	Replicated Model	183.94	19.95	**0.01**	6.72	0.26	**0.36**
	Guided Image	183.93	20.34	-	6.70	0.26	-
Efb3	Fibro-CoSANet	183.28	22.87	-0.16	6.72	0.34	**0.01**
	Replicated Model	183.02	19.70	-0.30	6.70	0.27	-0.23
	Guided Image	183.58	20.34	-	6.72	0.29	-
Efb4	Fibro-CoSANet	184.86	22.60	**0.22**	6.73	0.30	**0.03**
	Replicated Model	183.21	19.43	-0.68	6.70	0.26	-0.38
	Guided Image	184.46	20.03	-	6.73	0.26	-
ResNet18	Fibro-CoSANet	184.71	23.79	-0.02	6.73	0.32	**0.49**
	Replicated Model	182.50	19.38	-1.23	6.71	0.27	**0.17**
	Guided Image	184.75	20.60	-	6.70	0.26	-
ResNet34	Fibro-CoSANet	184.79	21.45	-0.28	6.73	0.31	**0.57**
	Replicated Model	184.62	20.76	-0.37	6.71	0.25	**0.32**
	Guided Image	185.31	21.42	-	6.69	0.25	-
ResNet50	Fibro-CoSANet	184.07	21.74	-0.18	6.72	0.31	**0.35**
	Replicated Model	183.45	20.82	-0.52	6.69	0.24	-0.09
	Guided Image	184.40	21.24	-	6.70	0.26	-
ResNet101	Fibro-CoSANet	183.13	24.87	-0.23	6.71	0.30	**0.10**
	Replicated Model	184.44	21.20	**0.48**	6.71	0.26	**0.09**
	Guided Image	183.55	0.27	-	6.70	0.27	-
ResNet152	Fibro-CoSANet	183.89	22.25	-0.21	6.70	0.29	**0.04**
	Replicated Model	184.23	20.45	-0.02	6.71	0.27	**0.23**
	Guided Image	184.27	20.79	-	6.70	0.26	-
ResNext50	Fibro-CoSANet	184.75	22.75	-0.04	6.70	0.27	-0.10
	Replicated Model	184.03	20.55	-0.43	6.71	0.26	**0.06**
	Guided Image	184.82	20.85	-	6.71	0.25	-
ResNext101	Fibro-CoSANet	184.01	23.38	**0.34**	6.72	0.28	**0.28**
	Replicated Model	183.33	20.05	-0.03	6.71	0.26	**0.06**
	Guided Image	183.38	20.02	-	6.70	0.26	-

*The hyphen symbol is used when comparing the same experiment.
**The mean and standard deviation values were obtained from the results of the 5 experiments conducted using 5-fold cross-validation, with each of the f-encoder functions.