International Journal of Combinatorial Optimization Problems and Informatics, 15(2), May-Aug 2024, 139-146. ISSN: 2007-1558.
https://doi.org/10.61467/2007.1558.2024.v15i2.461

www.editada.org

Classification of Long COVID Pulmonary Fibrosis Based on Computed Tomography in Mexican Patients

Aron Hernandez-Trinidad[1], Rafael Guzman-Cabrera[2], José Ruiz Pinales[2], Teodoro Cordova-Fraga[1]
[1]División de Ciencias e Ingenierías
Universidad de Guanajuato campus
León, GTO, Mexico.
División de Ingenierías,
[2]Universidad de Guanajuato campus
Irapuato-Salamanca, GTO, Mexico.
E-mails: aron.hernandez@ugto.mx

Abstract. In this study, we focus on addressing the prolonged effects of COVID-19, specifically pulmonary fibrosis, through automated classification of computed tomography (CT) images from patients in León, GTO, Mexico. We employed a convolutional neural network (CNN) VGG16 along with image enhancement filters, such as Meijering and Roberts, to optimize image quality. Prolonged COVID has been linked to complications, including pulmonary fibrosis, underscoring the need for early detection. Our model, supported by a comprehensive dataset, achieved classification accuracy exceeding 97%, successfully distinguishing between patients with and without pulmonary fibrosis. The combination of enhancement filters and CNN VGG16 proved crucial in this success, highlighting the potential of our model for early detection and effective management of pulmonary fibrosis in long COVID patients. This promising approach may have a significant impact on clinical practice, enhancing outcomes and enabling timely interventions. In summary, we present an effective method contributing to timely diagnosis and treatment of pulmonary fibrosis in long COVID patients.

Keywords: Long COVID, pulmonary fibrosis, computed tomography, automated classification.

Article Info
Received May 20, 2024
Accepted Jun 1, 2024

1 Introduction

The COVID-19 pandemic, caused by the SARS-CoV-2 virus, has been a global phenomenon that has left a lasting impact on public health. By mid-2023, records show more than 769 million cases of coronavirus (SARS-CoV-2) since its emergence in 2020. Primarily characterized by acute respiratory manifestations, it has revealed a persistent facet in some individuals, known as "Long COVID". As studies have delved into the long-term sequelae of this disease, particular concern has arisen regarding pulmonary fibrosis, a condition characterized by the formation of scar tissue in the lungs. The evaluation of results was conducted by visually comparing them with interpretations by medical professionals specialized in identifying free fluid in abdominal CT images. Since the SARS-CoV-1 epidemic, there has been evidence of some medical sequelae causing these respiratory conditions, with pulmonary fibrosis being one of the most frequently occurring ailments.

Pulmonary fibrosis is the body's response to chronic lung injury, and its development has been observed in a significant number of convalescent COVID-19 patients one year after contracting the infection. This complication manifests through the accumulation of scar tissue in the lungs, compromising the elasticity and functional capacity of the vital respiratory organ. A third of patients admitted with this virus developed pulmonary fibrotic alterations over time.

Those who suffer from pulmonary fibrosis as sequelae of COVID-19 face a substantial challenge in their quality of life. Respiratory capacity is diminished, leading to difficulties in performing basic activities, reducing quality of life, affecting sleep, and generating constant fatigue. This situation impacts not only physical health but also the mental and emotional health of those affected, as they face a chronic condition that demands long-term management. Statistics reveal that a significant percentage of individuals who have experienced prolonged COVID develop pulmonary fibrosis one year after the initial infection. These figures underscore the pressing need to research and develop effective methods for early detection and appropriate management of this long-term complication.

In this context, our study focuses on addressing this issue through the implementation of advanced technologies, such as automated classification of computed tomography (CT) images, to improve detection and management of pulmonary fibrosis in Mexican patients affected by Long COVID. Through the application of a pre-trained convolutional neural network (CNN) VGG16 and image enhancement filters, we aim to contribute to knowledge and clinical practice, offering precise and effective tools for timely diagnosis and treatment of this post-COVID complication.

Our objective is to experiment with a series of convolutional axial tomographies with and without findings of Long COVID pulmonary fibrosis, applying different image enhancement filters that allow us to determine which one highlights the scars formed in the patients' lungs most clearly. Subsequently, we will use a pre-trained convolutional neural network to classify, among a set of images, those suffering from fibrosis.

2 Related work

Detecting pulmonary fibrosis accurately and timely is crucial for effective patient management and treatment planning. Deep learning techniques, coupled with advanced imaging modalities such as computed tomography (CT), have shown great promise in enhancing the detection and classification of pulmonary fibrosis. These techniques leverage the power of artificial intelligence to analyze intricate patterns and features in medical images, offering a more objective and consistent approach compared to traditional diagnostic methods. Given the complex nature of pulmonary fibrosis and its significant impact on patient outcomes, the development of accurate and efficient deep learning models is of paramount importance. In this context, we present a review of recent research articles that highlight the application of deep learning methodologies for the detection and evaluation of pulmonary fibrosis, emphasizing the advancements and insights gained in this field.

Table 1. Recent research on deep learning techniques for pulmonary fiborsis detection and evaluation.

Authors	Methodology	Main Finding
Christe et al. (2019)	Deep learning with CT images	Proposed a computer-aided diagnosis system for pulmonary fibrosis using deep learning techniques and CT images.
Walsh et al. (2018)	Deep learning with high-resolution CT images	Utilized deep learning for classifying fibrotic lung disease on high-resolution CT scans.
Yadav et al, (2022)	Deep learning with honeycombing features	Developed FVC-NET for automated diagnosis and predition of pulmonary fiborsis progression using honeycombing.
Walsh et al. (2020)	Deep learning for unsolverd problems in fibrotic lung disease	Applied deep learning techniques to unresolved issues in imaging research of fibrotic lung disease.
Levin (2018)	Deep learning evalutaion of pulmonary fibrosis	Discussed the application of deep learning in evalutaing pulmonary fibrosis, providing insights into its potential.

3 Theoretical framework

Pulmonary fibrosis is a condition whose treatment becomes more complex over time. Similarly, detecting it in early stages is not an easy task, as the scars that may form are not yet sufficiently visible, potentially being overlooked. Currently, its basic detection is using axial tomographies and breath tests. However, research for its early detection is still ongoing. Researchers at the University of Eberhard Karls Tübingen (Germany) conducted research proposing, in some cases, the use of a PET imaging agent, an imaging agent known as "64Cu-GPVI-Fc," acting as a non-invasive medium for patients' bodies, detecting pulmonary fibrosis in its early stages as it is directly associated with fibrosis activity and scar tissue, detecting patterns that betray its influence on the lungs.

Adhering to the method of computed axial tomographies, they are an indispensable tool in medical diagnosis, responsible for identifying types of traumas, strokes, various types of cancer, as well as respiratory pathologies and conditions. Their spatial resolution is around 0.3 mm, and their acquisition speed ranges from 0.5 s to 10 s (full volume). As a precedent and to improve the samples provided by the tomographies, the use of image enhancement filters has been put into practice. Giraldo, Fletcher, and McCollough (2010) point out that "the anisotropic bilateral filter allows for a better balance between noise, spatial resolution, and detectability of low contrasts" (pp.61-62). Reducing undesirable elements in CT scans clarifies and highlights points of interest, guaranteeing quality in the tests.

The application of instruments that alter the visual result of tomographies is supported by the manuscripts of various researchers in the field. This is evident in the study "System for comparison and advanced analysis of medical images obtained through computed axial tomography (CAT)," where Aguilar (2010) notes that "These concepts allow us to go beyond where the eye and human perception can reach, significantly assisting specialists in their decision-making" (p.1). The article rescues a series of programs, formulas, and algorithms whose advances can be used from the detection of tumors or damaged tissue types with the Snake algorithm, smoothing the finishes of CT scans using a Gaussian filter, or reducing noise presence with a Median filter, filters that, as needed, achieve satisfactory results.

4 Methodology

For the experiment, a total of 96 axial tomographies were collected from the IMSS T1 in León, GTO, Mexico. Among these, 39 correspond to lungs with findings of pulmonary fibrosis, while 57 images show no evidence of pulmonary fibrosis. It's noteworthy that all data were obtained from Mexican patients, encompassing both healthy and afflicted individuals.

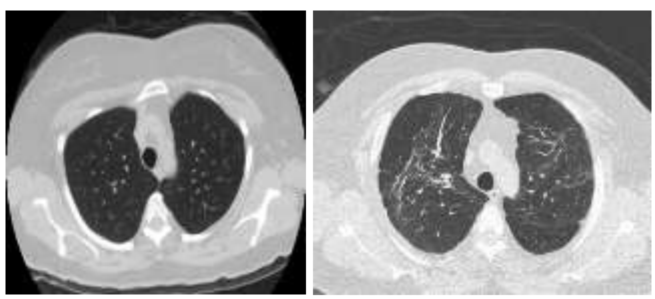

Figure 1. Example dataset of Mexican patients, labeled in two parts: healthy and fibrosis.

Image preprocessing of the dataset was conducted to enhance relevant features and improve class quality. Morphological filtering and image enhancement techniques were applied for this purpose. The decision to use these filters in our study is driven by the need to enhance image quality, reduce noise, and improve the accuracy of pulmonary fibrosis classification in CT scans. Their application contributes to the overall effectiveness of our methodology and the reliability of our study outcomes

- Morphological Dilation Filter: A morphological dilation filter was employed to enhance regions of interest and improve the detection of relevant features in the images. This image processing technique expands areas of interest pixels, facilitating the identification of distinctive patterns.
- Meijering Filter: The Meijering filter was applied to enhance edge features and tubular structures present in the images. This filter, based on Hessian matrix analysis, enhances edges and contours, which is useful for identifying specific anatomical patterns in axial tomography images.

- Roberts Filter: The Roberts filter was implemented to detect sudden intensity changes in the images and highlight the edges of objects present in them. This filter, which uses convolution masks to calculate intensity gradients in the image, provides an improvement in edge detection and structures of interest.

Image preprocessing was conducted systematically and uniformly across the entire dataset, ensuring consistency and quality of images for subsequent analysis and automated classification of pulmonary fibrosis. For image processing, a 3x3 convolutional mask was applied to CT images with dimensions of 800x700 pixels. This mask size was chosen to balance feature extraction effectiveness with computational efficiency. Morphological dilation filtering, a widely recognized technique in image enhancement, was employed to highlight relevant features in the images, enhancing the accuracy of subsequent classification tasks.

In terms of validation, a stratified 20-fold cross-validation approach was utilized to ensure robustness and reliability of the classification model. This strategy involves dividing the dataset into 20 equally sized subsets while preserving the distribution of classes, allowing for thorough evaluation of the model's performance across diverse data partitions. The decision to use 20 folds was based on extensive experimentation, where it consistently yielded competitive performance metrics, with classification accuracies surpassing 90%.

Furthermore, the parameter k in the k-nearest neighbors (k-NN) algorithm was set to 0.01 after rigorous experimentation. This value was found to produce optimal classification results, as deviations from it led to a noticeable decrease in model performance. By meticulously selecting parameter values and validation strategies, our experiments aimed to ensure transparency, reproducibility, and high-performance classification of pulmonary fibrosis.

In our study, we employed the VGG16 convolutional neural network (CNN) architecture as a feature extractor for pulmonary fibrosis classification. The VGG16 model, pre-trained on a large dataset of natural images, has demonstrated remarkable performance in various computer vision tasks. To adapt the VGG16 architecture to our specific task, we resized the CT images to 224x224 pixels, the input size expected by the VGG16 model. This resizing process ensures compatibility with the architecture's requirements while preserving essential information from the original images.

Once the images were resized, they were passed through the VGG16 network to extract high-level features. The VGG16 architecture consists of multiple convolutional and pooling layers, followed by fully connected layers, enabling it to capture intricate patterns and textures from the input images. By utilizing the pre-trained weights of the VGG16 model, we leveraged the learned representations of natural images to enhance the discriminative power of our classification model for pulmonary fibrosis. This approach allows us to exploit the hierarchical features learned by the VGG16 network, facilitating effective classification of CT images based on pulmonary fibrosis presence.

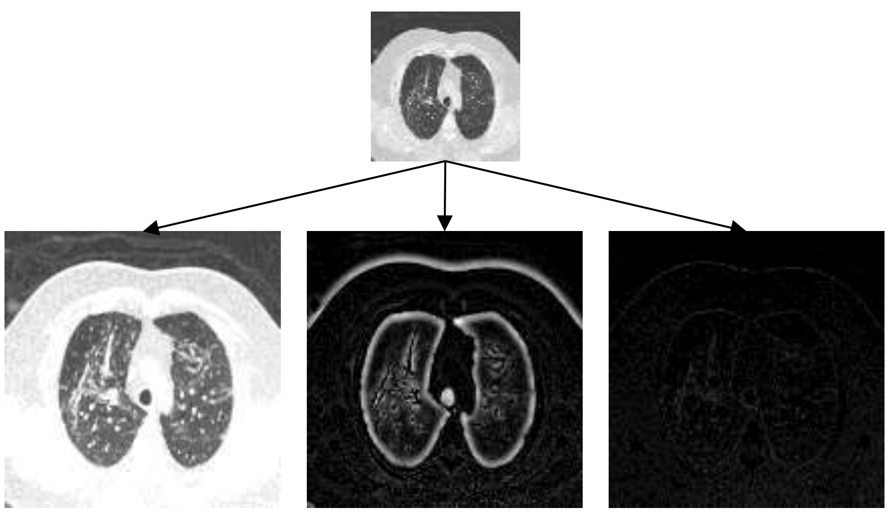

Figure 2. Preprocessing the images of the dataset to highlight characteristics and quality of thce classes. (Left) Morphological dilation filter, (Center) Meijering filter and (Right) Roberts filter.

Circular mask application

Circular masks were applied to the axial CT images to isolate and emphasize the alveoli structures. The objective behind this approach was to achieve enhanced classification accuracy between both classes. The circular masks were strategically positioned to focus exclusively on the alveolar regions within the lung images. By isolating and highlighting these specific structures, the classification model could better discern subtle variations and abnormalities associated with pulmonary fibrosis.

This technique aimed to improve the distinction between healthy lung tissue and regions affected by fibrosis, thereby enhancing the effectiveness of the classification process. The application of circular masks provided a targeted approach to feature extraction, contributing to the overall robustness of the classification methodology.

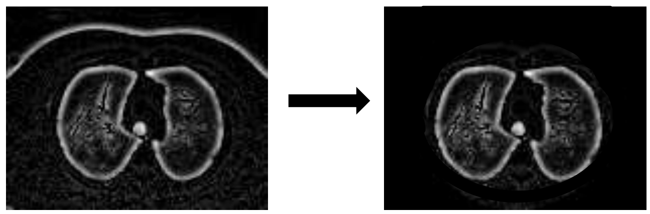

Figure 3. Application of masks to highlight the region of interest. (Left) Experiment one, with the original images and filters, (Right) Experiment two, with the results of the best filter and circular/ellipse mask.

5 Model

The proposed classification model consists of a sequential pipeline comprising three essential components. Firstly, the dataset comprises axial CT images sourced from pulmonary fibrosis patients. Secondly, feature extraction is performed using the VGG16 convolutional neural network (CNN). The VGG16 model serves as a powerful feature extractor, capable of capturing intricate patterns and textures within the images. By leveraging pre-trained weights and architectures, VGG16 generates high-dimensional feature vectors that encapsulate rich information about the underlying image content. Finally, these feature vectors are passed into the classification system, where machine learning algorithms or classifiers utilize them to discern distinct patterns associated with pulmonary fibrosis. The adoption of VGG16 as the feature extractor is paramount due to its proven efficacy in image recognition tasks and its ability to provide informative feature representations that facilitate accurate classification outcomes.

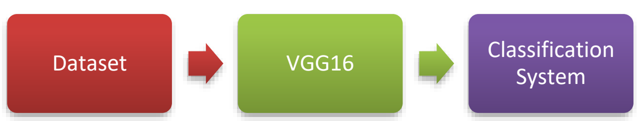

Figure 4. Proposed model to identify fibrosis in Mexican patients.

Future research

Circular masks were applied to the axial For future research endeavors, exploring the integration of a feed-forward classifier alongside the k-nearest neighbors (k-NN) algorithm for pulmonary fibrosis classification could be promising. Incorporating a feed-forward neural network as an alternative classifier offers the potential to capture more complex relationships in the feature space compared to k-NN.

A prospective study could involve constructing a feed-forward neural network architecture tailored to the specific characteristics of the dataset and the task of pulmonary fibrosis classification. This architecture could leverage the high-dimensional feature vectors extracted by the VGG16 network and incorporate various configurations of dense layers with different activation functions and regularization techniques.

Furthermore, conducting a comparative analysis between the feed-forward classifier and k-NN would be valuable to assess their respective strengths and weaknesses in pulmonary fibrosis detection. Metrics such as classification accuracy, precision, recall, and F1-score could be computed and compared to evaluate the relative performance of each classifier.

Exploring the potential of integrating a feed-forward classifier alongside k-NN represents an important avenue for future research in the field of pulmonary fibrosis detection and classification. This investigation could contribute to advancing the state-of-the-art, ultimately improving the accuracy and reliability of diagnostic tools for this critical medical condition.

6 Results

Cross-validation technique was employed to evaluate the results of the statistical analysis, ensuring independence from the partition between training and testing data. This approach provides a robust and reliable assessment of the classification model's performance. The obtained results are presented in Tables 2 and 3, showcasing two key evaluation metrics: accuracy and precision. Accuracy reflects the proportion of correct predictions made by the classifier overall, while precision indicates the proportion of true positive predictions, specifically concerning the classification of pulmonary fibrosis. The classification algorithm utilized in this study was k-NN (k-nearest neighbors), which was fine-tuned to obtain optimal values for the classification model. This choice was based on its ability to adapt to data patterns and its simplicity of implementation.

Analysis of the results reveals that by using masks (second experiment) the automatic classification of the model is improved, that is, by identifying the area of interest, the prediction of identifying fibrosis improves the model obtaining optimal values. These findings are pivotal for understanding the model's performance and its capability to discern between cases of pulmonary fibrosis and unaffected ones.

Table 2. Results of the first experiment.

Image	Accuracy	Precision
Original	0.87 (+/- 0.14)	0.87 (+/- 0.16)
Roberts	0.81 (+/- 0.12)	0.87 (+/- 0.12)
Meijering	0.92 (+/- 0.12)	0.92 (+/- 0.12)
Dilatación	0.87 (+/- 0.14)	0.87 (+/- 0.14)

Table 3. Results of the second experiment.

Meijering Filter	Accuracy	Precision
No mask	0.92 (+/- 0.12)	0.92 (+/- 0.12)
Mask	0.95 (+/- 0.09)	0.97 (+/- 0.05)

7 Discussion

Cross-validation technique The advancement of medical technologies has facilitated the detection of anomalies within the human body. This experiment underscored the potential of CT scans to enhance performance and effectiveness in diagnosing pulmonary fibrosis. The utilization of the Meijering filter on such medical images reveals indications of tissue injuries in the lungs, enabling neural networks to classify them accurately without significant difficulty. However, this tool does not aim to replace or supplant the clinical diagnosis proposed by medical professionals; instead, it serves as a visual aid, allowing these lesions to be visualized more clearly in each computed axial tomography scan.

8 Conclusion

By applying an advanced convolutional neural network, particularly the CNN VGG16 architecture, in conjunction with image enhancement techniques to our automatic classification model, we achieved a notable classification accuracy exceeding 97%. These results unveil promising potential for the early detection of fibrosis in individuals affected by long COVID, leading to a significant improvement in the management of this persistent health condition among the Mexican population. In conclusion, our research constitutes a valuable contribution to the realm of medical imaging and the exploration of the long-term effects of COVID.

By offering a reliable tool to identify fibrosis in Mexican patients, we are facilitating more effective management of this persistent health condition. This progress has a direct impact on healthcare provision by providing a solution that promotes early detection and, consequently, allows for more timely interventions to enhance the quality of life of those affected.

References

Achenbach S., Marwan M., Ropers D., Schepis T., Pflederer T., Anders K., Kuettner A., Daniel W.G., Uder M., Lell M.M. Coronary computed tomography angiography with a consistent dose below 1 mSv using prospectively electrocardiogram triggered high-pitch spiral acquisition, *European Heart Journal*, 31, 340-346, 2010.

Aguilar Peña, J. A. (2013). SLD194 Sistema para la comparación y análisis avanzado de imágenes médicas obtenidas por medio de tomografía axial computarizada (TAC) SLD194 Comparison System for Advanced Analysis of Medical Images. En Informática en Salud (pp. 1-10). Universidad de Oriente.

Amin, B. J. H., Kakamad, F. H., Ahmed, G. S., Ahmed, S. F., Abdulla, B. A., Mohammed, S. H., Mikael, T. M., Salih, R. Q., Ali, R. K., Salh, A. M., & Hussein, D. A. (2022). Post COVID-19 Pulmonary Fibrosis; A Meta-analysis study. *Annals of Medicine and Surgery*, 77.

Chérrez-Ojeda, I., Gochicoa-Rangel, L., Salles-Rojas, A., & Mautong, H. (2021). Seguimiento de los pacientes después de neumonía por COVID-19. Secuelas pulmonares. *Revista Alergia México*, 67(4).

Christe, A., Peters, A. A., Drakopoulos, D., Heverhagen, J. T., Geiser, T., Stathopoulou, T., ... & Ebner, L. (2019). Computer-aided diagnosis of pulmonary fibrosis using deep learning and CT images. *Investigative Radiology*, 54(10), 627.

Colectivo de autores. Grupo de Investigación EDMANS: Técnicas y algoritmos básicos de visión artificial. Universidad de la Rioja. Extracción de bordes; Operadores Sobel, Prewitt y Roberts. (n.d.).

Flores, N. (2022, November 7). Cross validation: qué es y su relación con machine learning. Tecnológico de Monterrey.

Levin, D. L. (2018). Deep learning and the evaluation of pulmonary fibrosis. *The Lancet Respiratory Medicine*, 6(11), 803-805.

Los sets de entrenamiento, validación y prueba | Codificando bits. (n.d.). Codificando Bits.

Marcos, A. (2020, May 21). Fibrosis pulmonar, una secuela en pacientes recuperados de COVID-19. Comisión Universitaria para la Atención de la Emergencia Coronavirus.

Na, & Na. (2020, July 15). Algoritmo K-Nearest Neighbor | Aprende Machine Learning. Aprende Machine Learning.
NayakShilp, BushomgVanna: Implementation of Snake Method for Lip Tracking. April 1, 2012. UCLA.

Orús, A. (August 8, 2023). Número acumulado de casos de coronavirus en el mundo desde el 22 de enero de 2020 hasta el 2 de agosto de 2023. Statista.

Pajares Gonzalo, M. de la cruz Jesús: Visión por computador. Universidad Complutense de Madrid.

Ramírez Giraldo, J. C., Fletcher, J. J., & McCollough, C. H. (2010). Reducción del ruido en imágenes de tomografía computarizada usando un filtro bilateral anisotrópico. *Revista Ingeniería Biomédica*, 4(7), 55-62.

MathWorks España. (n.d.). Red Neuronal Convolucional VGG-16 - MATLAB VGG16. Recovery of https://es.mathworks.com/help/deeplearning/ref/vgg16.html

SKImage.Filters — SKImage 0.22.0 documentation. (n.d.). Recovery of https://scikit-image.org/docs/stable/api/skimage.filters.html

Walsh, S. L., Calandriello, L., Silva, M., & Sverzellati, N. (2018). Deep learning for classifying fibrotic lung disease on high-resolution computed tomography: a case-cohort study. *The Lancet Respiratory Medicine*, 6(11), 837-845.

Walsh, S. L., Humphries, S. M., Wells, A. U., & Brown, K. K. (2020). Imaging research in fibrotic lung disease; applying deep learning to unsolved problems. *The Lancet Respiratory Medicine*, 8(11), 1144-1153.

Yadav, A., Saxena, R., Kumar, A., Walia, T. S., Zaguia, A., & Kamal, S. M. (2022). FVC-NET: an automated diagnosis of pulmonary fibrosis progression prediction using honeycombing and deep learning. Computational Intelligence and Neuroscience, 2022.