ISSN 1870-4069

Automatic COPD Detection through Vocal Emissions Using Intelligent Audio Analysis

Carlos Leyva-Bravo[1], Ivan Gallardo-Bernal[2], Humberto Pérez-Espinosa[1]

[1] Universidad Autónoma de Guerrero, Maestría en Tecnologías de la Información, Mexico

[2] Instituto Nacional, de Astrofísica, Óptica y Electrónica, Puebla, Mexico

carosleyvabravo@gmail.com, humbertop@inaoep.mx, gallardo@uagro.mx

Abstract. This study presents an application of artificial intelligence for the early detection of chronic obstructive pulmonary disease (COPD) through analyzing vocal emissions. Using audio processing techniques and machine learning, we analyzed 30 vocal recordings from individuals in vulnerable communities in Guerrero, Mexico. These recordings included clinically confirmed COPD patients, healthy controls, and individuals exposed to smoke from burning organic materials. Our approach employs a Transformer-based classifier to analyze vocal patterns and identify distinctive COPD characteristics, providing a non-invasive and accessible screening tool. Implementing AI in this context addresses significant healthcare access disparities by offering a scalable and cost-effective diagnostic tool for communities with limited access to advanced technologies. Preliminary results indicate that our transformer-based model effectively distinguishes between COPD patients and other groups, demonstrating its potential to enhance early detection and improve patient outcomes. This study underscores the transformative impact of artificial intelligence in promoting health equity and advancing public health in marginalized populations.

Keywords: COPD, respiratory diseases screening, intelligent audio analysis.

1 Introduction

Chronic Obstructive Pulmonary Disease (COPD) is a severe medical condition that critically impairs the respiratory system, significantly limiting the ability to breathe. According to the World Health Organization (WHO), COPD was the third leading cause of death worldwide in 2023, responsible for more than 3 million deaths [5]. Characterized by persistent obstruction of airflow in the lungs, COPD leads to breathing difficulties and substantially decreases quality of life.

The development of COPD is strongly associated with prolonged exposure to pulmonary irritants, smoking being the main risk factor. However, exposure to environmental pollutants, dust and chemicals also contributes to its onset [7]. COPD typically encompasses two main conditions: chronic bronchitis and emphysema, which often co-occur, exacerbating the disease's severity and the extent of lung damage in affected individuals. The situation is exacerbated in areas with high social marginalization and rural regions in Mexico, such as the southern and southeastern states such as Oaxaca, Guerrero, Chiapas, and Michoacán, where biomass (organic material derived from plants and animals) is a common energy source for cooking and heating homes [10]. The use of biomass contributes to poor air quality and increased respiratory health issues.

According to the National Institute of Statistics and Geography (INEGI) [11], COPD ranks tenth among the leading causes of death in these regions. In the state of Guerrero, chronic obstructive pulmonary diseases were the 11th leading cause of mortality between 2020 and 2022, as reported by the Epidemiological and Statistical Death System (SEED). The population of Guerrero is diverse, comprising various indigenous groups. However, they face significant vulnerabilities, including poverty, marginalization, discrimination, and limited access to basic health services due to inadequate infrastructure, a shortage of medical personnel, and difficulties in accessing medical care. These factors heighten the risk of poor health outcomes, underscoring the need for targeted interventions to improve healthcare access and equity.

2 Related Work

Studies on COPD have revealed notable differences between patients exposed to biomass and those affected by tobacco. Symptoms such as dyspnea, chronic bronchitis, rales, and wheezing were more common in the biomass group [3]. Spirometric measurements showed higher levels of severity of Forced Expiratory Volume in 1 Second (FEV1) and Forced Vital Capacity (FVC) in patients exposed to biomass compared to tobacco exposure. Radiographically, emphysema was more prevalent in the tobacco group, while bronchiectasis and atelectasis were more common in patients exposed to biomass. Another study [2] analyzed the effects of indoor biomass smoke pollution, concluding that biofuel smoke adversely affects the respiratory system. It affects the lung parenchyma and contributes to the burden of respiratory diseases, including COPD and tuberculosis, in exposed populations.

Artificial intelligence is transforming various medical specialties, with notable advances in neurology, oncology, radiology, and clinical pathology [8]. For instance, AI-based algorithms in oncology accurately diagnose cancer in computational histopathology. In radiology, deep learning software shows significant strides in image-based diagnosis compared to radiologists. Integrating AI in the diagnosis and management of COPD marks a substantial advance in contemporary medicine [9]. AI can reduce the global burden of COPD, reduce healthcare costs, and improve early diagnosis [19]. Improves clinical data interpretation, facilitating early intervention and more effective COPD management. In a recent study [21], the authors reviewed the role of AI in diagnosing and treating respiratory diseases such as COPD and asthma.

The findings highlight the utility of AI in clinical, functional, and imaging analysis and in developing clinical prediction models and remote patient monitoring. AI for respiratory disease analysis, such as the eRx tool [6] with its CNNs algorithms, utilizes innovative techniques for remote X-ray analysis in suspected tuberculosis cases. [15] investigated the early detection of respiratory diseases using advanced signal processing and machine learning [1]. The study achieved notable accuracy in detecting respiratory events such as crackles and wheezing through autoregression combined with SVM and CNN. CNNs for image analysis and classification algorithms for structured data were used in [20]. The final model, which integrated information from X-rays and clinical data, achieved high accuracy in the detection of COPD, which is a valuable tool to improve medical diagnosis and influence future research in respiratory health.

AI can identify early signs of diseases such as COPD, asthma, pneumonia, and COVID-19 [18]. by analyzing vocal emissions, such as cough contents and lung sounds [13]. This early detection capability is crucial for timely diagnosis and effective treatment, improving patient outcomes and quality of life. For example, a machine listening system was developed using deep learning to detect coughs and diagnose early respiratory diseases from noisy audio signals [17]. The system employed logarithmic spectrograms and a convolutional neural network, achieving high sensitivity.

We found limitations in the reviewed related works that our proposal aims to address. One of them is the lack of diversity and breadth in datasets, affecting the generalization and accuracy of diagnoses. In addition, challenges in external validation raise doubts about the reliability of previous findings. Our proposal addresses these issues by integrating diverse data sets and employing advanced AI and machine learning techniques. We prioritize rigorous validation through clinical studies and real-world testing, ensuring our solution's reliability across various medical contexts. By overcoming these limitations, our proposal significantly improves the early detection and management of COPD, advancing medical research.

3 Methodology

A dataset was constructed to train a COPD detection model based on audio recordings obtained from lung sounds, including cough recordings from individuals from the most relevant indigenous communities inhabiting the state of Guerrero with COPD and healthy subjects as the sample population. The study included three groups of patients, each consisting of 10 individuals from vulnerable communities in the state of Guerrero. The first group comprised healthy individuals without respiratory diseases or chronic conditions. The second group was formed by 10 individuals with suspected COPD. The third group included 10 individuals clinically diagnosed with COPD.

Sample collection was carried out in clinics of the Mexican Social Security Institute located in the state of Guerrero, where patients attended check-up consultations and complications in respiratory diseases. Data collection was supervised at all times by the specialist physicians and nurses on duty. In addition, a meticulous and ethical process of collecting personal, hereditary, and general health data was carried out, with special attention to their self-determination as Indigenous people, delivering to each patient an informed consent detailing the use and responsible management of this

information for academic and research purposes. In addition to the above, each patient was provided with a questionnaire based on the GOLD (Global Initiative for Chronic Obstructive Lung Disease) dyspnea scale to assess respiratory difficulty. In addition, relevant demographic data of the studied population were collected.

4 Capture of the Sample

The study included participants from the Amuzgos, Mixtecos, Tlapanecos, and Nahuas communities. The first group comprised 7 Amuzgos patients, the second group included 8 Mixtecos patients, the third group had 8 Tlapanecos patients, and the fourth group consisted of 9 Nahuas patients. For the study, a total of 420 cough data from patients in indigenous communities were used to classify between positive and negative cases of COPD. The data were divided into training sets (70%) and test sets (30%), stored in CSV files. An effort was made to maintain a representative proportion of positive and negative patients in both data sets to ensure that the classification model correctly learned the differences between the classes.

In the training set, having a larger amount of data allows the model to have more examples to learn from, thus improving the model's ability to generalize to unseen data. The testing set should be large and representative enough to evaluate the model's performance accurately. Using 126 samples for testing ensures that the model is not overfitting to the training data and that its performance can be evaluated on unseen data. Class balance in the training set is essential to avoid bias towards one class. With 7 positive patients and 14 negative ones, a certain balance is ensured, which helps the model effectively learn to distinguish between both classes.

In the testing set, maintaining a proportion of positive patients helps to evaluate whether the model has a high false negative rate or if it can accurately detect positive cases. Undoubtedly, this model will need to be tested with larger data sets; however, collecting data from patients with COPD is not an easy task, as it is not common for patients to attend regular medical check-ups but rather situations ranging from clinical complications to emergencies. Each CSV file contains relevant information about the audio files, such as the file path and name and the associated class (Positive or Negative). The data were processed to load and resample the audio files to 16 kHz and then used to train and evaluate a classification model. The main formula of the attention mechanism is Scaled Dot-Product Attention. It is defined as:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V, \tag{1}$$

where: QKT: First, the product of matrices between Q (Query) and the transpose of K (Key) is calculated, resulting in a matrix of similarity scores. Each score indicates how relevant a key is to a query value in the similarity matrix is divided by the square root of the dimension of the keys (Keys). This operation is performed to prevent scores from becoming too large, which could lead to very small gradients during backpropagation (vanishing gradients) [14]. Softmax: The softmax function is applied to the resulting matrix to normalize the scores from 0 to 1. This converts the scores into probabilities, indicating the relative importance of each key (Key) concerning the query (Query).

Multiplication by V: The probability matrix is multiplied by V (value) to obtain the weighted output. This operation combines the values according to their relative importance determined by the attention.

5 Wav2Vec 2.0 Model

Wav2Vec 2.0 is a speech recognition model developed by Facebook AI (now Meta AI) that has proven effective in automatic speech recognition (ASR) tasks. Wav2Vec 2.0 focuses on self-supervised pretraining, [4] which means it can learn robust speech representations using unlabeled data. It can then be fine-tuned with a much smaller amount of labeled data for specific speech recognition tasks, and this model can be adapted using audio signals for the detection of COPD or other respiratory diseases.

The structure of Wav2Vec 2.0 consists of two main parts: The Feature Encoder and the contextual network [12] The Feature Encoder transforms the raw audio signal into more manageable representations using a convolutional network that extracts local features from the audio wave. The signal is divided into small windows to be processed by different convolutional layers, and the output is a sequence of vectors representing the audio signal in a compact and latent form.

The contextual network of Wav2Vec 2.0 uses a Transformer Encoder to analyze speech and capture relationships between different elements in the audio sequence. This component is crucial to understanding the global context of speech [16]. In the Transformer Encoder of Wav2Vec 2.0, self-attention layers allow each latent vector in the sequence to attend to all other vectors. This allows the model to capture long-term dependencies in the audio sequence.

Additionally, each self-attention layer is followed by a feed-forward neural network applied to each latent vector independently. These feedforward networks help transform the representations nonlinearly, enriching the model's ability to capture complex patterns in the data. To stabilize and improve training, each sublayer in the Transformer Encoder of Wav2Vec 2.0 has residual connections and layer normalization. Residual connections help avoid gradient degradation issues and enable more effective training.

6 Data Collection Protocol

For this study, patients were asked to cough for 10 seconds, take a break, and cough again for 10 seconds until they completed a minute of recording the cough with a high-quality microphone. A capture application was developed, where the patients' coughs were recorded to later save them in WAV format for later processing with artificial intelligence methods and techniques.

7 Data Preprocessing

Subsequently, the sound data were preprocessed. The recordings were standardized to 1.4 seconds by repeating or cropping the sound signals. They were labeled as "Positive" (sick) or "Negative" (healthy) and separated into training and testing sets. The recordings were loaded and resampled to a frequency of 16 kHz.

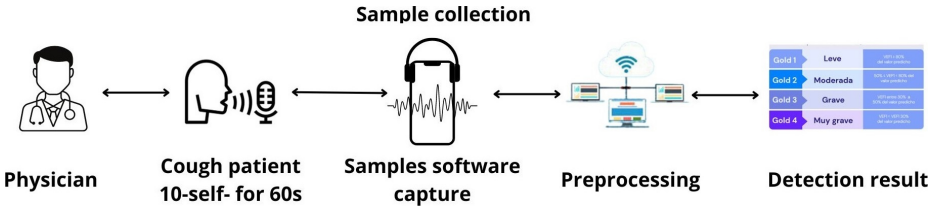

Fig. 1. Data collection protocol. Own creation.

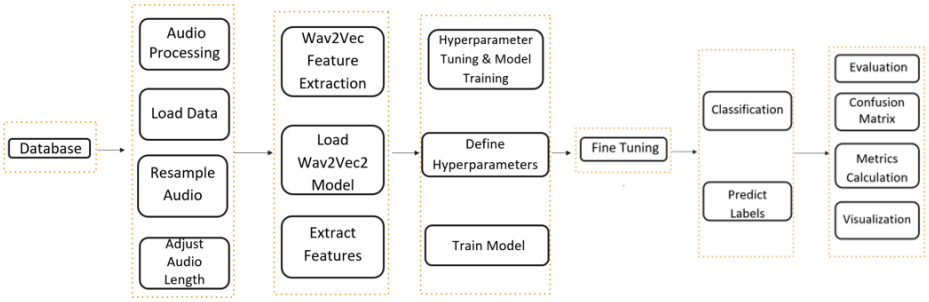

Fig. 2. Wav2Vec 2.0 process stages. Own creation.

8 Feature Extraction

Next, the pre-trained Wav2Vec 2.0 model, designed for voice recognition, was utilized to extract features from the audio recordings. This model transforms the audio signals into latent representations that capture important information for classification. To obtain the final classification, a global pooling layer and additional dense layers were applied to these representations, with regularization through dropout.

9 Optimization and Training

The training was optimized using the Hyperopt library to find the optimal hyperparameters. Finally, the performance of the model was evaluated in the test set, obtaining precision, recall, specificity, accuracy, and F1 score metrics. The results were displayed in precision and loss graphs, and a confusion matrix was generated to analyze the model's performance in COPD detection.

10 Results

Comparing the model's results with the pre-trained database and with the new specific data from the vulnerable population allowed the evaluation of the model's ability to generalize to different datasets. This is crucial to ensure that the model does not overfit a specific dataset and can be applied to diverse populations. The study demonstrates the potential of pre-trained AI models for COPD classification, showing promising results even when applied to data from vulnerable patients in the state of Guerrero.

Table 1. Pre-trained model results.

Results	Precision	Recall	F1-Score	Support
Healthy	0.72	0.77	0.74	30
Sicks	0.89	0.86	0.88	66
Macro avg	0.80	0.82	0.81	96
Weighted avg	0.84	0.83	0.83	96

Table 2. Study results in vulnerable populations.

Results	Precision	Recall	F1-Score	Support
Healthy	0.71	0.72	0.77	10
Sicks	0.88	0.87	0.88	10
Macro avg	0.80	0.81	0.81	20
Weighted avg	0.81	0.82	0.80	20

Despite some variations in performance, the model's ability to adapt to new data underscores its flexibility and utility. Although the model's accuracy varied between the original and new data, these results indicate the model's robustness. The observed decrease suggests specific areas for improvement, providing a solid foundation for future optimizations and adjustments. Differences in the sensitivity and specificity of the model when analyzing patient data highlight the importance of incorporating specific genetic and environmental factors.

These variations provide valuable insights that can be used to customize and improve the model. Training Set Results. The first section of the table shows the results obtained when the model was applied to the training data. These results are broken down by class, that is, "healthy" individuals and "sick" individuals. For each class, the following metrics are reported:

– Precision: The proportion of true positives among the total predicted positives. For the healthy individuals class, the precision was 0.72, while for the sick individuals, it was 0.89.

– Recall: The proportion of true positives among the total actual positives. The recall for healthy individuals was 0.77, and for sick individuals it was 0.86.

– F1-Score: The harmonic mean between precision and recall provides a balanced measure of the model's performance. The F1-Score for healthy individuals was 0.74, and for sick individuals it was 0.88.

– Support: The number of samples in each class, with 30 samples for healthy individuals and 66 for sick individuals.

In addition, aggregated metrics are presented to evaluate the overall performance of the model:

– Accuracy: The proportion of correct predictions out of the total predictions. The total accuracy of the model in the training set was 0.83, based on 96 samples.

– Macro Avg: The average of precision, recall, and F1-Score calculated in an unweighted manner across all classes, resulting in values of 0.80, 0.82, and 0.81 respectively.
– Weighted Avg: The average of precision, recall, and F1-Score weighted by the number of samples in each class, resulting in values of 0.84, 0.83, and 0.83 respectively.

Test Set Results The second section of the table presents the results of the model applied to a test dataset, specifically designed to evaluate its generalization capability. The following metrics are reported:

– Precision: The precision for the healthy individuals class in the test set was 0.71, and for sick individuals, it was 0.88.
– Recall: The recall for healthy individuals was 0.72, while for sick individuals, it was 0.87.
– F1-Score: The F1-Score for healthy individuals was 0.77, and for sick individuals it was 0.88.
– Support: The number of samples in each class was 10 for both healthy and sick individuals.

The aggregated metrics for the test set were as follows:

– Accuracy: The total accuracy of the model on the test set was 0.80, based on 20 samples.
– Macro Avg: The average of precision, recall, and F1-Score calculated in an unweighted manner across all classes, resulting in values of 0.80, 0.81, and 0.81 respectively.
– Weighted Avg: The average of precision, recall, and F1-Score weighted by the number of samples in each class, resulting in values of 0.81, 0.82, and 0.80 respectively.

11 Discussion of Results Pre-trained Model Results

The pre-trained model showed good performance in the classification of COPD. The results indicate that the precision for healthy individuals was 0.72, while for sick individuals it was 0.89. The model achieved an F1-score of 0.74 for healthy individuals and 0.88 for sick individuals, with support of 30 and 66 samples, respectively. The overall accuracy of the model was 0.83, with a macro average of 0.81 and a weighted average of 0.83. These results demonstrate the effectiveness of the pre-trained model in COPD classification, but they also suggest that the model could benefit from additional adjustments to improve precision and recall, especially in diverse populations. One of the primary limitations identified is the relatively small size of the dataset used.

Although the model has shown promising performance under controlled conditions, a larger and more diverse dataset is crucial for validating and enhancing its generalizability. a greater volume of data will allow the model to learn and adapt to a broader range of variations in vocal characteristics associated with COPD, thereby improving its accuracy and robustness. Additionally, validating the model in a realistic setting remains an outstanding issue. data obtained in a laboratory environment may not reflect the variable and potentially noisy conditions encountered in a real hospital setting. Environmental noise, medical equipment, and data capture devices could affect the quality of vocalization data, which may influence the accuracy of the model.

For example, background noise in a hospital or differences in the quality of hardware used for capturing vocalizations could introduce variations not observed under controlled laboratory conditions. The transformer model has been evaluated primarily in controlled environments, and more work is needed to adapt and validate it in real clinical situations. this includes conducting extensive tests in hospital settings and with data capture equipment that reflects practical use conditions. evaluating the model under these conditions will allow identification and addressing of possible biases or deficiencies, as well as making necessary adjustments to enhance its effectiveness in detecting COPD in diverse populations and clinical contexts

12 Study Results in Vulnerable Populations

When applying the model to specific data from vulnerable populations, the results showed a slight decrease in precision and F1-score. For healthy individuals in these populations, the precision was 0.71 and the F1-score was 0.77, with a support of 10 samples. For sick individuals, the precision was 0.88 and the F1-score was 0.88, also with a support of 10 samples. The overall accuracy in these populations was 0.80, with a macro average of 0.81 and a weighted average of 0.80. The observed decrease in precision and F1-score when applying the model to vulnerable populations indicates that the model, although robust, needs to be adjusted to account for specific factors in these populations, such as genetic and environmental characteristics. These results provide a solid foundation for future improvements to the model, to increase its precision and adaptability to different contexts and populations.

13 Conclusion

The study has demonstrated that the pre-trained model has a remarkable capacity to generalize to different datasets, which is crucial to avoid overfitting and ensure its applicability to diverse populations. This ability allows the model to maintain its effectiveness even when facing new and specific data from vulnerable populations, such as patients in the state of Guerrero. Despite some performance variations, the results indicate that the model possesses great flexibility and utility. Its ability to adapt to new data underscores the potential of pre-trained AI models for the classification of diseases such as COPD. The robustness of the model is evident in its consistency when comparing the original data with the new data, although a slight decrease in precision and F1-score was observed.

This decrease suggests specific areas for improvement, providing a solid foundation for future optimizations and adjustments. The differences observed in the model's sensitivity and specificity when analyzing patient data highlight the importance of incorporating specific genetic and environmental factors. These factors are essential for personalizing and improving the model, thus increasing its precision and applicability in different contexts and populations. The study's results offer valuable insights that can be used to fine-tune the model more precisely, enhancing its performance in diverse populations. Moreover, the study demonstrates that the pre-trained model can be a valuable tool in the classification of COPD, even in vulnerable populations.

This is especially relevant in contexts with resource limitations and specific population conditions, where a flexible and adaptable model can have a significant impact. It is important to note that these data are preliminary and that collaborations are currently underway with public and private hospitals to collect more data and with a population entirely different from the current one to diversify the model. this expansion of the database will allow for further evaluation and enhancement of the model's capacity to generalize to diverse populations, increasing its robustness and applicability in different clinical contexts.

This strategy will not only improve the model's precision and reliability but also contribute to its ability to adapt to real clinical scenarios, providing a valuable resource for the classification and management of COPD in various populations. Similarly, efforts will be made to validate the models with databases collected by other institutions. this validation will allow for the comparison and evaluation of the model's performance on varied and previously unseen datasets, ensuring that its application is truly universal and effective. External validation is essential for identifying potential biases and limitations of the model, providing opportunities for fine-tuning to improve its accuracy and clinical utility.

By integrating data from various sources, a more comprehensive and detailed understanding of COPD can be achieved, which in turn strengthens the model's ability to adapt to different epidemiological and socioeconomic realities, ensuring its effectiveness across a wide range of contexts and populations. In summary, the study's findings underscore the potential of pre-trained AI models for medical applications, demonstrating their effectiveness, flexibility, and ability to continuously improve through specific adjustments based on real and diverse data. These findings provide a solid foundation for future research and development, to optimize and adapt these models to the specific needs of different populations, thus improving precision and effectiveness in disease classification.

References

1. Alvarez-Casado, C., Lage-Cañellas, M., Pedone, M., Wu, X., Nguyen, L., Bordallo-López, M.: Respiratory disease classification and biometric analysis using biosignals from digital stethoscopes. In: Proceedings of the 32nd European Signal Processing Conference, pp. 1556–1560 (2024) doi: 10.23919/EUSIPCO63174.2024.10714964
2. Carbajal, D. S.: La contaminación intramuros del humo de biomasa. Universidad Y Sociedad, vol. 14, no. S1, pp. 396–402 (2022)

3. Castañeda-Torreblanca, M. J.: Diferencias clínicas, espirométricas y radiográficas en enfermedad pulmonar obstructiva crónica por biomasa y tabaco. Master's thesis, Universidad Privada Antenor Orrego (2022)
4. Chen, L. W., Rudnicky, A.: Exploring wav2vec 2.0 fine tuning for improved speech emotion recognition. In: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 1–5 (2023) doi: 10.1109/ICASSP49357.2023.10095036
5. Cheng, W., Zhou, A., Zeng, Y., Lin, L., Song, Q., Liu, C., Zhou, Z., Peng, Y., Yang, M., Yang, L., Chen, Y., Cai, S., Chen, P.: Prediction of hospitalization and mortality in patients with chronic obstructive pulmonary disease with the new global initiative for chronic obstructive lung disease 2023 group classification: A prospective cohort and a retrospective analysis. International Journal of Chronic Obstructive Pulmonary Disease, vol. 18, pp. 2341–2352 (2023) doi: 10.2147/copd.s429104
6. Curioso, W. H., Brunette, M. J.: Inteligencia artificial e innovación para optimizar el proceso de diagnóstico de la tuberculosis. Revista Peruana de Medicina Experimental y Salud Pública, vol. 37, no. 3, pp. 554–8 (2020) doi: 10.17843/rpmesp.2020.373.5585
7. Elonheimo, H. M., Mattila, T., Andersen, H. R., Bocca, B., Ruggieri, F., Haverinen, E., Tolonen, H.: Environmental substances associated with chronic obstructive pulmonary disease—a scoping review. International Journal of Environmental Research and Public Health, vol. 19, no. 7, pp. 3945 (2022) doi: 10.3390/ijerph19073945
8. Expósito-Gallardo, M. C., Ávila-Ávila, R.: Aplicaciones de la inteligencia artificial en la medicina: Perspectivas y problemas. Acimed, vol. 17, no. 5 (2008)
9. Feng, Y., Wang, Y., Zeng, C., Mao, H.: Artificial intelligence and machine learning in chronic airway diseases: Focus on asthma and chronic obstructive pulmonary disease. International journal of medical sciences, vol. 18, no. 13, pp. 2871 (2021)
10. García-Frapolli, E., Schilmann, A., Berrueta, V. M., Riojas-Rodríguez, H., Edwards, R. D., Johnson, M., Guevara-Sanginés, A., Armendariz, C., Masera, O.: Beyond fuelwood savings: Valuing the economic benefits of introducing improved biomass cookstoves in the Purépecha region of Mexico. Ecological Economics, vol. 69, no. 12, pp. 2598–2605 (2010) doi: 10.1016/j.ecolecon.2010.08.004
11. González-Block, M. Á., Reyes-Morales, H., Cahuana-Hurtado, L., Balandrán, A., Méndez, E.: Mexico: Health system review. Health Systems in Transition, vol. 22, no. 2 (2020)
12. Hsu, W. N., Sriram, A., Baevski, A., Likhomanenko, T., Xu, Q., Pratap, V., Kahn, J., Lee, A., Collobert, R., Synnaeve, G., Auli, M.: Robust wav2vec 2.0: Analyzing domain shift in self-supervised pre-training. Interspeech, pp. 721–725 (2021) doi: 10.21437/interspeech.2021-236
13. Koul, A., Bawa, R. K., Kumar, Y.: Artificial intelligence techniques to predict the airway disorders illness: A systematic review. Archives of Computational Methods in Engineering, vol. 30, no. 2, pp. 831–864 (2022) doi: 10.1007/s11831-022-09818-4
14. Lovisotto, G., Finnie, N., Munoz, M., Murnmadi, C. K., Metzen, J. H.: Give me your attention: Dot-product attention considered harmful for adversarial patch robustness. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 15213–15222 (2022) doi: 10.1109/CVPR52688.2022.01480
15. Mang, L. D.: Investigación y desarrollo de técnicas de procesado de señal e inteligencia artificial aplicadas a la recuperación de información biomédica a partir del análisis de señales sonoras respiratorias. Master's thesis, Universidad de Jaén (2024)
16. Novoselov, S., Lavrentyeva, G., Avdeeva, A., Volokhov, V., Gusev, A.: Robust speaker recognition with transformers using wav2vec 2.0 (2022) doi: 10.48550/ARXIV.2203.15095
17. Pérez-Alonso, D. A.: Análisis de señales de tos para detección temprana de enfermedades respiratorias. Master's thesis, Universidad de Valladolid, Escuela Técnica Superior de Ingenieros de Telecomunicación (2019)

18. Rani, S., Chaurasia, A., Dutta, M. K., Myska, V., Burget, R.: Machine learning approach for automatic lungs sound diagnosis from pulmonary signals. In: Proceedings of the 44th International Conference on Telecommunications and Signal Processing, pp. 366–371 (2021) doi: 10.1109/TSP52935.2021.9522663
19. Robertson, N. M., Centner, C. S., Siddharthan, T.: Integrating artificial intelligence in the diagnosis of COPD globally: A way forward. Chronic Obstructive Pulmonary Diseases: Journal of the COPD Foundation, vol. 11, no. 1, pp. 114–120 (2024) doi: 10.15326/jcopdf.2023.0449
20. Sierra-Villegas, S.: Caracterización de la enfermedad pulmonar obstructiva crónica (EPOC) a partir de imágenes de radiografías de tórax y datos clínicos. Master's thesis, Universidad Escuela de Ingeniería de Antioquia (2023)
21. Vargas-Ramírez, L., Ayazo, R. B.: Inteligencia artificial en neumología. Medicina, vol. 43, no. 4, pp. 570–581 (2021)