Pulmonary Disease Classification on Electrocardiograms Using Machine Learning

Aboubacar Abdoulaye Soumana[1], Prajwol Lamichhane[1], Mehlam Shabbir[1], Xudong Liu[1], Mona Nasseri[1], Scott Helgeson[2]
[1]University of North Florida, Jacksonville, FL 32224
[2]Mayo Clinic, Jacksonville, FL 32224
{n01483682, n01530010, n01474570, xudong.liu, mona.nasseri}@unf.edu, helgeson.scott@mayo.edu

Abstract

Pulmonary diseases, such as chronic obstructive pulmonary disease (COPD) and asthma are among the leading causes of death in the US. These lung diseases often are diagnosed by pulmonologists using physical exam (e.g., lung auscultation) and objective measurement of lung function with pulmonary function testing (PFT). These extensive tests, however, can be inaccessible to many patients due to limited resources and availability. In this paper, we explore the use of the easily accessible electrocardiograms (ECGs) to train machine learning models to classify pulmonary diseases. To this end, we developed and experimented with two approaches: deep neural network models trained (e.g., Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs)) on ECG signals directly, and non-neural models (e.g., support vector machines (SVMs) and logistic regression) trained on derived features from ECGs. In the task of classifying whether a patient has obstructive lung disease, our results show that deep neural network models outperformed the non-neural models, though the difference is within 3% on accuracy and F1-score metrics.

Introduction

This paper aims to explore the feasibility and effectiveness of using ECG monitors as a diagnostic tool for pulmonary diseases, leveraging the accessibility and convenience of wearable ECG technology and the advancement of machine learning techniques. By investigating the potential correlation between ECG data and pulmonary health, we seek to contribute to the growing body of knowledge in pulmonary critical care and offer insights that may improve the early detection and management of respiratory disorders. Through this exploration, we hope to enhance the accessibility and efficiency of pulmonary disease diagnosis and, ultimately, improve patient outcomes.

Machine learning has been extensively applied to ECG analytics of heart diseases, as noted in various survey papers (Jambukia, Dabhi, and Prajapati 2015; Berkaya et al. 2018; Adasuriya and Haldar 2023). Recent advancement in machine learning, especially deep learning, has resulted in research projects training deep neural network models using ECGs to detect respiratory diseases as well. COVID-19 detection from ECG data using deep learning has recorded high accuracy (Prashant et al. 2022; Sakr et al. 2023). On the topic of detecting COPD, Moran et al. (Moran et al. 2023) applied several deep neural networks to detect COPD or healthy using ECG images produced by signal-to-image transformation techniques, while Sarkar et al. (Sarkar et al. 2022) uses traditional machine learning models such as supported vector machines and k-nearest neighbors to detect obstructive, restrictive or healthy using a small number of features (e.g., P/R ratio and energy ratio) derived from ECG. Albeit these research projects obtained outstanding results, both are limited in the size of the subject population (usually with less than 100 patients) and in the time of ECG recordings (from minutes to hours). In our project, we analyze 10 seconds ECG segments recorded from 21,644 patients, in hope to train models for a large group of patients using short ECGs that are very accessible via wearable monitors.

Copyright © 2024 by the authors.
This open access article is published under the Creative Commons Attribution-NonCommercial 4.0 International License.

Research Methodology

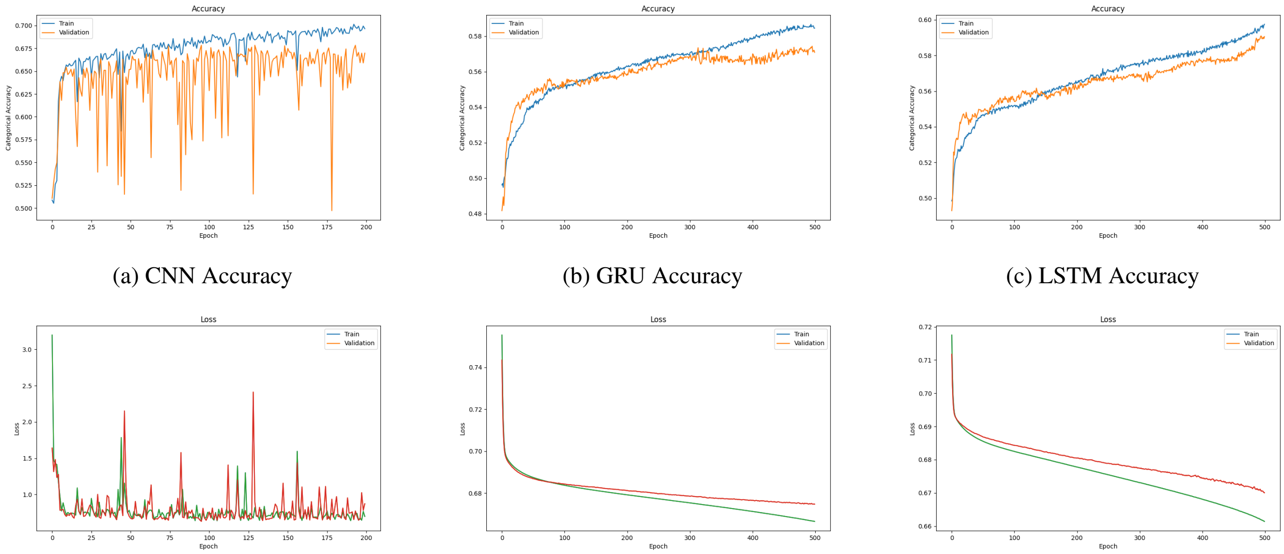
In pursuit of our research objectives, we designed and executed two distinct machine learning approaches to harness the potential of ECG data as a diagnostic tool for pulmonary diseases. These approaches encompassed deep neural models, specifically Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU), as well as non-neural models, including Logistic Regression, XGBoost, and Random Forest.

Feature Extraction

Our dataset consists of 21,644 patients' 12-lead ECGs, each a time-series sequence of amplitudes over 10 seconds. For the neural and non-neural models, we extracted relevant features to facilitate the diagnosis of pulmonary diseases.

For Recurrent Neural Models To prepare the data for recurrent neural models, each patient's 12-lead ECGs were transformed into a 3D array of shape [21644, 5000, 12], where the dimensions correspond to batch size, time stamps, and features, respectively. This transformation enabled our recurrent models to process sequential data effectively.

For Non-Neural Models For non-neural models, we derived 11 distinctive features (such as PR Ratio and P Energy)



(a) CNN Accuracy (b) GRU Accuracy (c) LSTM Accuracy

(d) CNN Loss (e) GRU Loss (f) LSTM Loss

Figure 1: Learning Curves for Deep Neural Network Models

from every patient's 12-lead ECGs based on PQRST segmentation. All these features are results of consultation with our medical expert and in accordance with ECG Library [1].

Model Training

For the training of our models, we adopted specific configurations (i.e., hyperparameters) tailored to their architecture and purpose. Due to space constraint, we omit the details of these configurations in this poster paper.

Table 1: Results of the selected model on the test set

Model	Accuracy	Precision	Recall	F1-Score
LSTM	**0.70**	0.71	**0.68**	**0.69**
CNN	0.69	**0.72**	0.62	0.66
SVM	0.67	0.69	0.62	0.66
Logistic Regression	0.65	0.67	0.64	0.65
XGBoost	0.64	0.66	0.61	0.64
Random Forest	0.64	0.66	0.61	0.64
Decision Tree	0.57	0.58	0.58	0.58
GRU	0.49	0.49	0.52	0.50

Results and Analysis

This section details our models' performance metrics on the test dataset, divided into 80% for training and 20% for testing. Table 1 lists accuracy, precision, recall, and F1-score for each model in diagnosing obstructive disease.

Firstly, our LSTM model, achieving an accuracy of 0.70 and an F1-score of 0.69, stands out as the top-performing model. It outperforms all other models across all metrics but precision for the binary classification task, demonstrating its effectiveness in diagnosing pulmonary obstruction. Our CNN model achieves the best precision among all models. Together our best results not surprisingly belong to the most complex deep learning models. What's surprising is GRU's performance, which seems to exhibit a nearly random guess.

Secondly, among the non-neural models, during the training process, various classification models were evaluated based on their performance metrics. Support Vector Machine (SVM) outperforms the other models, reaching 0.67 accuracy and 0.66 F1-score. While not reaching the same performance level as the deep neural network model, the SVM's utilization of only the 11 derived features reduces training time and dataset size, falling with only 3% behind the best LSTM model, which showcases the effectiveness of this non-neural model.

Thirdly, further comparing our deep neural network models in terms of their learning curves, as shown in Figure 1, we observe that the LSTM RNN model demonstrates appropriate-fitting compare to GRU which seems to plateau after 300 epochs, and that the CNN model presents fluctuating instability from epoch to epoch which seems to suggest high level of variance in our sheer number of patients.

Conclusion

Despite the best performance goes to the more complex deep neural network models, our findings indicate that the non-neural models consistently demonstrated comparable, and in some cases, even superior performance when compared to recurrent neural models.

[1] `https://ecglibrary.com/norm.php`

References

[Adasuriya and Haldar 2023] Adasuriya, G., and Haldar, S. 2023. Next generation ecg: the impact of artificial intelligence and machine learning. *Current Cardiovascular Risk Reports* 17(8):143–154.

[Berkaya et al. 2018] Berkaya, S. K.; Uysal, A. K.; Gunal, E. S.; Ergin, S.; Gunal, S.; and Gulmezoglu, M. B. 2018. A survey on ecg analysis. *Biomedical Signal Processing and Control* 43:216–235.

[Jambukia, Dabhi, and Prajapati 2015] Jambukia, S. H.; Dabhi, V. K.; and Prajapati, H. B. 2015. Classification of ecg signals using machine learning techniques: A survey. In *2015 International Conference on Advances in Computer Engineering and Applications*, 714–721. IEEE.

[Moran et al. 2023] Moran, I.; Altilar, D. T.; Ucar, M. K.; Bilgin, C.; and Bozkurt, M. R. 2023. Deep transfer learning for chronic obstructive pulmonary disease detection utilizing electrocardiogram signals. *IEEE Access*.

[Prashant et al. 2022] Prashant, K.; Choudhary, P.; Agrawal, T.; and Kaushik, E. 2022. Owae-net: Covid-19 detection from ecg images using deep learning and optimized weighted average ensemble technique. *Intelligent Systems with Applications* 16:200154.

[Sakr et al. 2023] Sakr, A. S.; Pławiak, P.; Tadeusiewicz, R.; Pławiak, J.; Sakr, M.; and Hammad, M. 2023. Ecg-covid: An end-to-end deep model based on electrocardiogram for covid-19 detection. *Information Sciences* 619:324–339.

[Sarkar et al. 2022] Sarkar, S.; Bhattacharyya, P.; Mitra, M.; and Pal, S. 2022. Automatic detection of obstructive and restrictive lung disease from features extracted from ecg and ecg derived respiration signals. *Biomedical Signal Processing and Control* 71:102791.