IST-Africa 2022 Conference Proceedings
Miriam Cunningham and Paul Cunningham (Eds)
IST-Africa Institute and IIMC, 2022
ISBN: 978-1-905824-69-4

Deep Learning Based Algorithms for Detecting Chronic Obstructive Pulmonary Disease

Endashaw Amsalu MELESE[1], Evarist NABAASA[2] , Matiwos Tekalign WONDEMAGEGN[3], Safari YONASI[4] ,Gemechis Melkamu NEGASA[5]
[1,2,4] *Mbarara University of Science and Technology, Mbarara, P.O BOX 1410, Uganda*
Tel: +256-4854-21373, Fax: +25648520782, Email: end21am@gmail.com, enabaasa@must.ac.ug, safij86@gmail.com
[3]*University of Gonder, Gonder, P.O BOX 196, Ethiopia*
Tel: +251581141232, Email: matiwos.tekalign@uog.edu.et
[5]*Haramaya University, Diredawa, P.O BOX 138, Ethiopia*
Tel: +2515530320, Email: gemechis.melkamu@haramaya.edu.et

Abstract: Chronic obstructive pulmonary disease (COPD) is a heterogeneous disease with various clinical presentations. The basic abnormality in all patients with COPD is airflow limitation. The main method for diagnosis of COPD is using spirometer and imaging equipment, which are expensive and not suitable for use. This study aims at developing algorithms for analysing cough sounds for detecting COPD. CNN and CRNN based deep learning techniques are used for developing the algorithm. We have used both augmented and non-augmented datasets with three different feature extraction methods: Mel-frequency cepstral coefficient, zero crossing rate, and harmonic change detection function. The developed CNN and CRNN scored an accuracy of 96.6% and 96.73% respectively. Conclusion: The proposed algorithms have improved classification performance that had been reported in the literature. Significance: The results of this study suggest that automatic diagnostic tools can be developed with less intervention from healthcare professionals.

Keywords: Chronic obstructive pulmonary disease, Deep learning methods, Cough Sound, Feature Extraction techniques

1. Introduction

Chronic obstructive pulmonary disease (COPD) is a heterogeneous disease with numerous clinical signs. The basic abnormality in all patients with COPD is airflow limitation [1]. Globally, it is estimated that about 3 million deaths were caused by the disease in 2016 (that is, five percent of all deaths globally in that year). More than 90% of COPD deaths occur in low and middle-income countries [2]. The primary cause of COPD is regular cigarette smoking, other factors such as prolonged exposure to air pollution, inhaling dust, or the fumes of fuel burned for cooking or heating purposes, chemicals, or fumes found in the workplace, and genetics problems.

Diagnosis with spirometry and imaging equipment (such as CT and X-ray) are the main techniques in COPD diagnosis. All of these methods are expensive in the context of developing countries, which costs 9.43-11.67 USD for spirometry [3], and 50-200 USD for CT scans and X-rays. In addition to this, they usually require specialist skills and laboratory facilities. Due to this reason, there is a need for an affordable and uncomplicated method of screening patients for further diagnostic testing [4]. Cough sound analysis using machine learning and deep learning-based techniques have shown promising results in diagnosis of COPD [5, 6, 7]. Automatic cough sound analysis using deep learning resolves this problem by taking only the cough sound of the patient. This will help to lower the diagnostic cost and make it suitable for use. The necessity of detecting cough is because it is one of the main reasons to visit a hospital in most countries across the world. Around 30 million patients that visit hospitals in the US are due to cough [8]. Cough can be aroused by accidental events for defending the respiratory system from foreign body intake. It is also one of the main symptoms of respiratory diseases [9, 10]. Cough sound based pulmonary disease detection is one of the oldest ways of diagnostic system for pulmonary diseases. In 1989, Piirila et al. [11] explored variations in acoustic and dynamic characteristics of involuntary cough in pulmonary diseases; following this, many pieces of research had been undertaken by different scholars.

Fernandez-Granero et al. [12] studied COPD exacerbations for timely identification of the severity of patients having COPD. The authors remotely monitored sixteen subjects having COPD for over six months. The cough sounds were taped using an automated system designed specifically for this purpose. The authors have trained and validated the recorded data using a decision tree forest classifier and achieved an accuracy of 75.8%. As the authors indicated, this had allowed an automatic forecast of symptom-based exacerbations [12]. Kanade and Bairagi [13] proposed a support vector machine classifier to classify the patients with COPD from normal ones using non-standard characters from electromyography along inhalation and exhalation progresses. Using this method, an accuracy rate of 87.8% was attained [13]. Badnjevic et al. [7] proposed new methods of automatic identification of asthma and COPD using an expert diagnostic system using clinical data. For coming up with more accurate algorithms, Badnjevic et al. [7] used data from 5307 patients gathered over two years. The algorithm was trained on 69% and tested on 31% of the dataset. The sensitivity and specificity of the developed algorithm were 96.45% and 98.71% respectively. The authors further implemented a system based on the developed algorithm. Amaral et al. [6] proposed an automatic identification of chronic obstructive pulmonary disease based on forced oscillation measurements (FOT) and artificial neural networks. The artificial neural network (ANN) inputs are the parameters given by the FOT, and the outputs are the implication of the parameters that indicate COPD or not. The dataset consists of 7 possible forced oscillation measurement parameters with an input of 90 features collected from 30 volunteers. The developed model is based on clinical data with a classification accuracy rate of 90% [6]. Using heart rate as a parameter, Newandee et al. [14] have classified COPD severity using the technique of PCA-CA (principal component and cluster analysis). The authors demonstrated that the PCA-CA technique is capable of differentiating normal subjects from subjects with COPD with an accuracy rate of 88%. They used clinical information (laboratory ECG, blood pressure, and respiration signal) of 55 subjects [14]. Daniel et al. [15] developed an algorithm for automated screening of COPD and asthma disease using a logistic regression model. The value from lung volume was measured with peak flow meter and clinical information. By combining the data from lung volume and regression models, they developed a mobile application. The developed algorithm scored an AUC of 0.95. The developed system is semi-automated, and it needs other equipment for measuring lung volume [15]. Altan et al. [5] developed an automatic diagnostic technique for identifying the stage of COPD severity, using the Hilbert-Hung transform. The authors used the cough sound of 50 subjects, of which 25 are COPD patients with different disease severity stages, and 25 are healthy personnel. The result from the developed system shows 93.67% accuracy [5]. Srivastava et al. [31] developed deep learning-based respiratory sound analysis for the detection of chronic obstructive pulmonary disease using CNN algorithm. The authors have used different types of feature extraction methods found in the librosa library. Of the different feature extraction methods, MFCC based works well with a sensitivity and specificity of 92%. Ahmed et al. [32] proposed a variant of 3D VoxResNet for COPD and emphysema classification. The model uses volume-wise annotations without any further feature enhancement or addition of meta-data. The authors tested the developed network with and without transfer learning. Without the use of transfer learning the developed network achieved 68.5% validation and 58.8% test accuracy. In comparison, with transfer learning, they were able to achieve 78.3% validation and 70.0% test accuracy.

Previous studies undertaken have four problems as a limitation that needs attention. These include noise-free data, small datasets, semi-automatic, and low accuracy. Therefore, there is a need to develop algorithms that are tested by a diverse dataset that shows the standard hospital scenario, better accuracy, and low computational time. In this paper, we implemented two deep learning techniques: CNN and CRNN for detecting COPD disease using cough sound.

Table 1: Summary of different COPD detection methods

Publication	**Objective**	**Result**	**Method**	**Gap**
Fernandez-Granero et al. [12]	COPD exacerbations for timely identification of severity of patient	Accuracy of 75.8%	DT	Accuracy is low and it's developed mainly for COPD exacerbation, not detection
Kanwade and Bairagi [13]	Separating patient with COPD from normal subject	Accuracy of 87.8%	SVM	Used data from EMG, unaffordable for developing countries
Badnjevic et al.[7]	Identification of asthma and COPD	Sensitivity of 96.4% Specificity of 98.7%	Expert diagnostic system using clinical data	Used clinical information and lung function test, unaffordable for developing countries
Amaral et al. [6]	Identification of COPD	Accuracy of 90%	FOT and ANN	Used clinical information and lung function test, unaffordable for developing countries
Newandee et al.[14]	Differentiating patient having COPD	Accuracy of 88%	PCA-CA	Used clinical information for identifying the stages of COPD; not the detection
Daniel et.al. [15]	Algorithm for automated screening of asthma disease	AUC of 0.95	LR models, value from LV and clinical information	The developed system is semi-automated and it needs another equipment for measuring lung volume
Shahmoradi adi et.al [33]	COPD cases into four phenotypes	Accuracy of 96%	Bayesian network	It is a semi-automatic and and it's developed mainly for identifying COPD phenotypes
Altan et.al. [5]	Detect COPD from cough sound	Accuracy of 93.6%	HHT	Dataset consisted of forced coughs, it's also small and is recorded in noise free room
Srivastaval	Detect COPD	Sensitivity and	CNN	Used only CNN algorithm,

et al.[31]	from cough sound	specificity of 92%		and the author suggested using data augmentation will improve the accuracy of the result
Ahmed et.al [32]	Classi*fi*cation of COPD using CT images and a 3D *CNN*	Using transfer learning (Validation:78.3%, Test accuracy:70%) Without transfer learning (Validation:68.5%, Test accuracy:58.8%)	CNN using CT scan images	The developed model was tested only on one algorithm, comparison with another network is required.

2. Materials and Methods

2.1- Research Design

The developed diagnostic model in this research follows different steps. Firstly, the dataset was organized, and time-domain based data augmentation was conducted. Following that, the features in the time domain of the audio signals are extracted. The extracted spectrogram signals were then trained and tested in the deep neural network architecture. Table 1 shows the flow diagram for each step.

Table 2: Steps for developing algorithms for COPD cough sound analysis

Audio dataset

↓

Pre-processing

↓

Feature Extraction

↓

Design of neural network

↓

Train and test the algorithms

↓

Selecting the best result

Each step is discussed in the following sub sections.

2.2 Dataset

In this research, the dataset of Challenge ICBHI 2017 published by Rocha et.al [16] was used. The audio samples had been collected for many years from two different countries (Greece and Portugal). It contains five and half hour's cough sound data with 6,898 annotated respiratory cycles from 126 subjects. The dataset contains healthy subjects and 7 different respiratory diseases: Bronchiectasis, Bronchiolitis, Pneumonia, COPD, lower respiratory tract infection, upper respiratory tract infection, and Asthma. The demographic information shows that 46 subjects are female, 80 are male, and 1 is unidentified gender. Children, adults, and elders are included in the study, with a distribution of ages - 47 subjects of 15 years of age, 78 subjects aged> 15 years, and 1 unidentified age group. The cough sounds range between ten and ninety seconds. Some respiration cycle comprises of a high noise level showing the capability of the dataset to simulate real-life situations.

2.3 Data Augmentation

We used both original data and augmented datasets for training and testing our algorithms. Data augmentation is a method of creating artificial data from the original dataset by transforming the original annotated training data, which helps in preventing overfitting. The most frequently used data augmentation technique in deep learning is image augmentation. Audio augmentation is used for overcoming data scarcity in audio fields, which is getting popular in recent times. Different audio augmentation tasks were done by different researchers [17] [18]. The first proposed audio augmentation is mapping the input frequency to another frequency [19]. Audio augmentation is prepared on either time or spectral domain. In this article, we have used five different time-domain based audio augmentations, based on a library built by Iver Jordel [20]. These are:

1) Adding Gaussian noise: Gaussian noise adds noise to the input audio signal with a specified signal-to-noise ratio, it is a statistical noise with the same probability-density function as normal-distribution [21].

2) Frequency mask: Frequency masking (FM) works by masking a given interval frequency. If a given frequency channel of f_m, $f_{m_o} + f_i$ are masked. f_i is selected from 0 to the FM parameter, and f_{m_o} is chosen from $(0, n_f + f)$ where n_f is the number of frequency channels [18].

3) Adding Pitch shift: This transformation changes the pitch of the input audio without changing the duration. The user can give the ratio by which the pitch needs to be scaled. The pitch of the input audio is changed without changing the duration [22].

4) Time mask: Time masking (TM) works by masking a given interval of time, it is implemented in the way that, for t_{m_i} time steps $(t_{m_0}, t_{m_o} + t_{m_i})$ are masked, where t_{m_i} is first selected from an even distribution starting from 0 to TM parameter [18].

5) Time stretch: Time stretch changes the duration of the input audio signal by either stretching or compressing the audio without changing the pitch [22].

2.4-Feature Extraction

1) Mel-Frequency Cepstral Coefficient (MFCC): MFCC is a perceptually inspired, short-time spectral feature, which is one of the commonly used feature extraction methods in machine learning and different audio/speech analysis. Steps followed in MFCC feature extraction are illustrated in Figure 1.

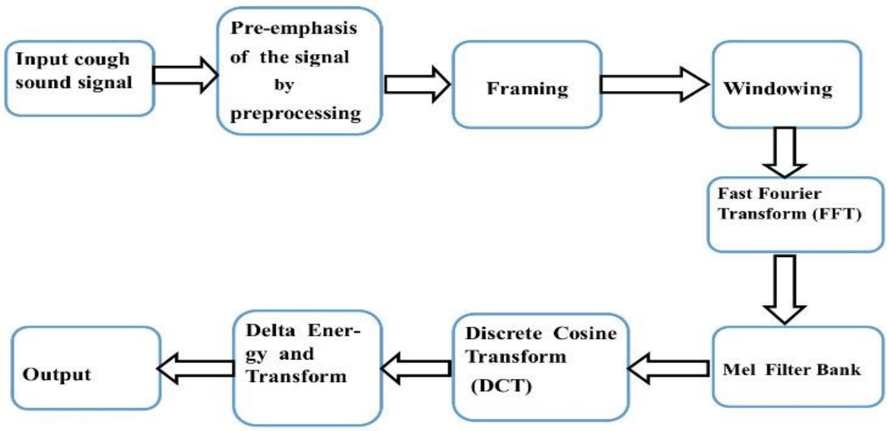

Figure 1: Steps for calculating MFCC of an audio signal

2) Zero-Crossing Rate (ZCR): ZCR shows the spectrogram image by measuring the rate at which the amplitude of the speech signals passes through a value of zero [23].

3) Harmonic Change Detection Function (Tonnetz): It extracts meaningful features by using event-driven feature analysis by chord recognition from the audio signal using a digital signal processing method [24].

3. Deep Learning

The extracted features of the audio signals are trained and tested with two different deep neural network architectures, which are Convolutional Neural Network and Convolutional Recurrent Neural Network.

3.1- Convolutional Neural Network

Convolutional Neural Network (CNN) is a commonly used deep learning mechanism that uses convolution rather than matrix multiplication, in at least one of its layers. A convolution is an operation that processes groups of neighboring pixels. A convolved feature map is designed by multiplying the kernel weights by the features. A convolutional layer in the neural network is a collection of features each having an aggregation of distinct features adjusted by the training process. CNN's ability to learn translation-invariant patterns and the ability to detect the spatial order of patterns, gives it a power in classifying digital signals [25]. The developed model for the CNN layer used in this research is depicted in Figure 3 below.

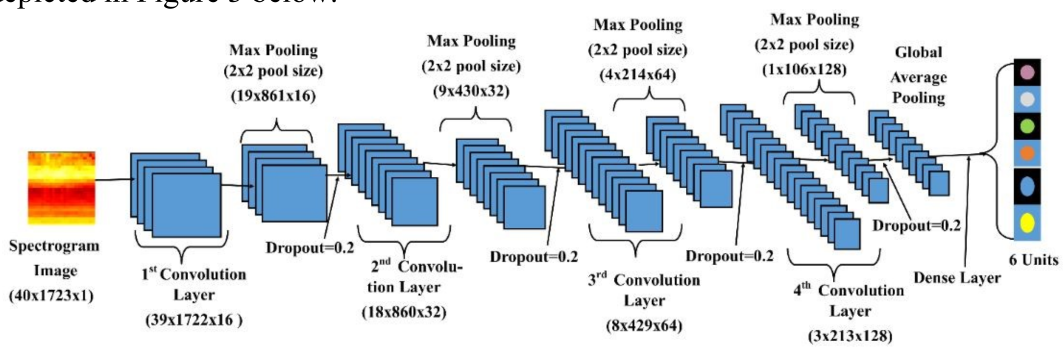

Figure 2: Step of CNN algorithm

3.2- Convolutional Recurrent Neural Network

The Convolutional Recurrent Neural Network (CRNN) is one of the latest neural network architectures that works by combination of the two well-known neural networks, CNN and RNN. CRNN works by taking CNN layers for feature extraction on the given input data joined with RNN to support sequence prediction. If we can combine both the CNN and RNN we can get a better result that can address the limitation of both architectures [26]. CRNN has advantage of both the spatial domain of CNN and temporal domain of RNN. The following diagram illustrates the CRNN architecture developed in this research.

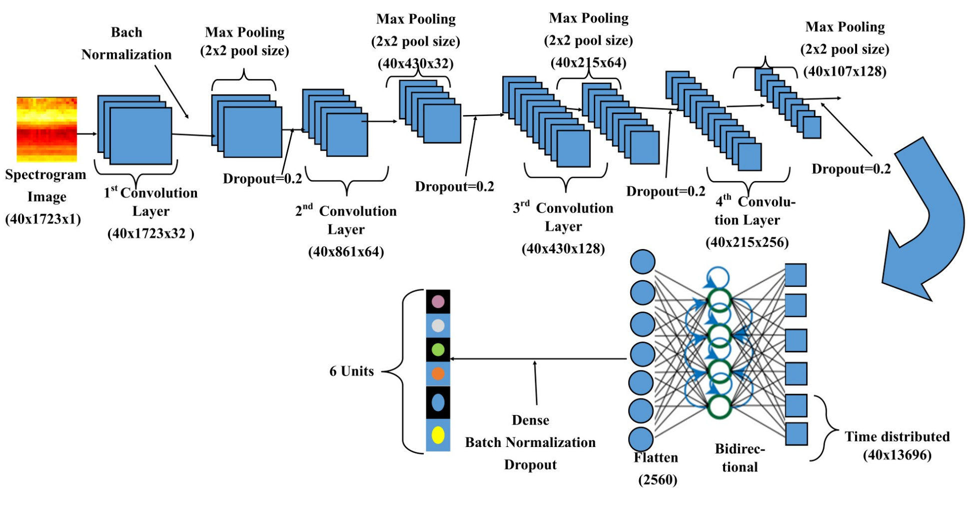

Figure 3: Steps of CRNN algorithm

4. Experimental Setup

For training and testing, we primarily imported all-important python libraries and application programming interfaces (API) such as Tensor flow library, Keras, Sklearn library, Librosa (audio analysis library), NumPy, Pandas, Matplotlib, and Seaborn libraries. The dataset is first augmented with five different time-domain data augmentation systems including; Gaussian noise, time mask, frequency mask, pitch shift and time stretch. After augmentation is done, three different feature extraction techniques were used to extract important features. The used feature extraction methods are the Mel-frequency cepstral coefficient, zero-crossing rate, and harmonic detection function. The extracted features are reshaped into a NumPy array format. The description of the CNN model of figure 3 is that: The input spectrogram image is fed into a convolution layer of 2x2 kernel size and 2x2 pool size with a dropout of 0.2 using Relu activation function. The feature map we get from this layer is convolved three different times. After getting the last feature map, we make a global pooling and obtain 6 feature maps by using the soft-max activation function. This gives us the model's accuracy in predicting the diagnostic ability of each input cough sounds.

The description of CRNN model of figure 3 is that the input spectrogram image is fed into a convolution layer of 3x3 kernel size with bath normalization and 2x2 pool size with a dropout of 0.2 using Relu activation function. The feature map we get from this layer is convolved four different times. After getting the last feature map, it feeds into a time distributed layer. This was followed by application of bidirectional GRU with 32 hidden units. Finally, a flatten layer with batch normalization is implemented to give a 6 units of output dense layer.

5. Evaluation Metrics

Apart from accuracy, examining precision and recall is necessary for determining the efficiency of machine learning algorithms [27]. In this research, we have analyzed the following different evaluation metrics:

1) Accuracy: This is used for comparing the effectiveness of the developed model in contrast with the actual value.
2) Precision: This is used for showing the accuracy of an algorithm with positive identification.
3) Recall: This is used for showing the accuracy of an algorithm with negative identification.
4) F1-Score: This is used for showing the weighted average of precision and recall.
5) ROC-Curve: This shows the relationship between sensitivity (precision) and 1-specificity. The higher the area under the ROC curve the more accurate the algorithm is.

6. Result and Discussion

In the developed algorithms, we divided the dataset into training and testing, 80% for training, and 20% for testing. We have used three different feature extraction methods which makes our algorithm more robust. To prevent overfitting with the training data, we have used early stopping callback. In the following tables, we have illustrated the result of the developed algorithms basing on each feature extraction method. For a better understanding of the impact of data augmentation, we have included the result of both augmented and non-augmented datasets. In most cases, the best performing algorithms are those of augmented data, which shows the impact of having big data in deep-learning tasks. In table 4, we are able to see the CRNN algorithm using TNT feature extraction method has performed well in non-augmented data during testing phase, which requires further analysis. The computational time for the non-augmented datasets is better than that of augmented data. This indicates that the computational time of deep learning algorithms are dependent on the size of the data. Since the dataset used in this study is big, this study used a GPU for better computational analysis. The GPU used is NVIDIA TESLA P100, with a GPU memory of 16GB and performance of 9.3TFLOPS and 1.32GHz GPu memory clock, which is provided by Kaggle and available for free.

Table 3: Result of CNN algorithm using augmented and non-augmented dataset (GPU was used)

Feature extraction method	Dataset Used	Computational Time(h:mm:ss)	Training Accuracy	Testing Accuracy
MFCC	Augmented	0:15:55	99.45%	96.64%
	Non-Augmented	0:08:08	95.36%	90.21%
TNT	Augmented	0:02:53	88.77%	87.92%
	Non-Augmented	0:00:36	87.44%	87.5%
ZCR	Augmented	0:01:48	86.07%	86.37%
	Non-Augmented	0:00:24	86.49%	86.41%

Table 4: Result of CRNN algorithm using augmented and non-augmented dataset (GPU was used)

Feature extraction method	Dataset Used	Computational Time(h:mm:ss)	Training Accuracy	Testing Accuracy
MFCC	Augmented	0:45:38	99.86%	96.73%
	Non-Augmented	0:27:32	99.86%	89.13%
TNT	Augmented	0:10:21	97.45%	91.0%
	Non-Augmented	0:01:58	99.59%	85.32%
ZCR	Augmented	0:05:13	90.6%	84.9%
	Non-Augmented	0:02:38	88.67%	83.15%

As seen in Table 3, the best performing algorithm is the CRNN algorithm with MFCC feature extraction method, with an accuracy of 96.73%. In all algorithms the MFCC based feature extraction technique has the best result. MFCC property of being decorrelated by discrete cosine transform has enhanced its detection ability, as research suggests decorrelated signals remove noise from them [28]. We hypothesized that CRNN's ability to take advantage of the spatial domain of CNN and the temporal domain of RNN had helped the CRNN algorithm to score the best result.

We have examined another evaluation metrics for the two best algorithms. Table 4 shows the precision, recall, and F1-scores of the two best algorithms (CRNN and CNN) using MFCC feature extraction method. The precision, recall, and F1-support of measure on the table 3 and 4 below indicates as the performance of the developed algorithms are good.

Table 5: Precision, recall, and F1-score of the CNN algorithm using MFCC feature extraction

Label Names	Precision	Recall	F1-Score
Bronchiectasis	89%	84%	86%
Bronchiolitis	87%	81%	84%
COPD	98%	100%	99%
Healthy	87%	79%	82%
Pneumonia	82%	73 %	77%
URTI	85%	79%	81%
Weighted avg	96%	97%	97%

Table 6: Precision, recall, and F1-score of the CRNN algorithm using MFCC feature extraction

Label Names	Precision	Recall	F1-Score
Bronchiectasis	88%	79%	83%
Bronchiolitis	92%	75%	83%
COPD	98%	100%	99%
Healthy	76%	90%	83%
Pneumonia	100%	75 %	86%
URTI	94%	61%	74%
Weighted avg	97%	97%	97%

The other parameter that was used to measure the performance of an algorithm is receiver operating characteristics (ROC). The ROC contains the true positive rate and false-positive rate. The higher the area under ROC, the more accurate the algorithm is [29]. From figures 4 and 5 we realize that CNN and CRNN obtained accuracies of 100% and 99% respectively, which were the highest scores that show the best techniques at performance. We have selected only MFCC based feature extraction method from the two algorithms, for figurative illustration

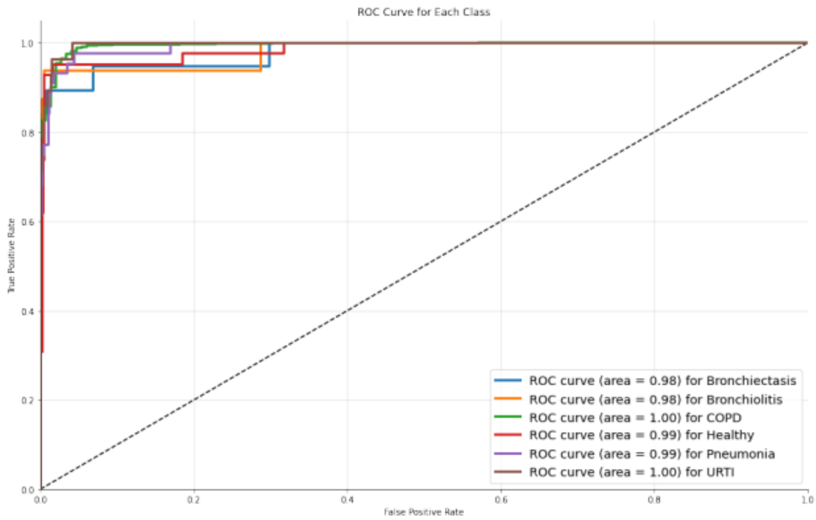

Figure 4: ROC of CNN algorithm using MFCC feature extraction.

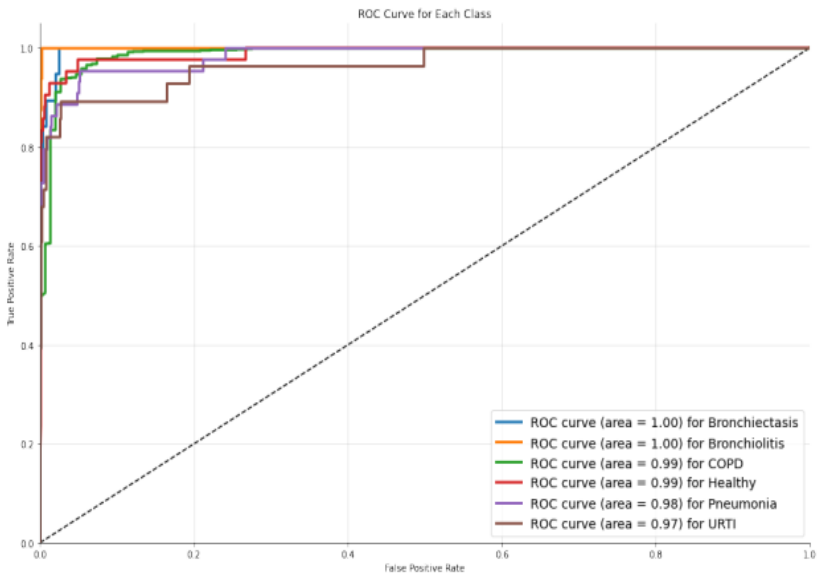

Figure. 5: ROC of CRNN algorithm using MFCC feature extraction

7. Conclusion

We designed an algorithm using convolutional neural network and recurrent neural network for diagnosing COPD disease. Original cough sound and five different time-domain based data augmentation techniques were used. On the developed algorithms, we have used three different feature extraction methods. Our experiment revealed that MFCC based feature extraction had worked well on all algorithms, CRNN being the best with an accuracy of 96.73%. We observed that the MFCC property of being decorrelated by discrete cosine transform had increased its statistical ability since decorrelated signals remove noise and let the audio signal to be accurately predicted [28]. We hypothesized that CRNN's ability to take advantage of the spatial domain of CNN and the temporal domain of RNN had helped the CRNN algorithm to score the best result.

Despite the good performance of the developed algorithms, we identify a limitation to this method. The dataset used in this research was collected from Greece and Portugal (which are developed countries), but training and testing the algorithm with cough sound datasets from low resource settings is necessary. In low resource settings, the medical wards that are used for diagnosis are narrow. Narrow rooms create an echo; this echo should be added to the cough sound and analyzed. In this scenario, echo cancellation methods should be developed to improve the efficiency of the algorithm. For this reason, it is recommended to collect data from low resource settings, which increases the generalizability of the algorithm, if a system is to be implemented and used in real world for diagnosis purposes. Furthermore, data needs to be collected from more subjects to further evaluate the developed algorithm.

Objective detection of cough sound by lessening intervention from health care practitioners during diagnosis can be easily achieved by digital respiratory sound analysis. This non-invasive way of detecting COPD disease using deep learning methods will provide a significant contribution to eHealth and telemedicine. This would also solve the problem of limited healthcare workers in developing nations. In addition to this, automatic detection of cough sound enhances long period therapy at an affordable price and a suitable method.

Generally, this research has enhanced existing approaches of COPD diagnoses using the cough sound of the patient subject.

References

[1] Peter J Barnes. Immunology of asthma and chronic obstructive pulmonary disease. *Nature Reviews Immunology*, 8(3):183–192, 2008.

[2] WHO.chronic-respiratory-diseases.https://www.afro.who.int/health-topics/chronic-respiratory-diseases, addendum= accessed:15.10.2019,.

[3] J Aguirre Ruiz, E Garrofa Vilert, P Saura Solanas, C Jordán Morera, C Beltrán Mallorquí, and M Marqués Mas. Costs of spirometry as a screening test for chronic obstructive pulmonary disease in primary care. *Atencion primaria*, 36(7):373–377, 2005.

[4] Sandra V Kik, Claudia M Denkinger, Martina Casenghi, Caroline Vadnais, and Madhukar Pai. Tuberculosis diagnostics: which target product profiles should be prioritised? *European Respiratory Journal*, 44(2):537–540, 2014.

[5] Gokhan Altan, Yakup Kutlu, and Novruz Allahverdi. Deep learning on computerized analysis of chronic obstructive pulmonary disease. *IEEE journal of biomedical and health informatics*, 2019.

[6] Jorge LM Amaral, Alvaro CD Faria, Agnaldo J Lopes, Jose M Jansen, and Pedro L Melo. Automatic identification of chronic obstructive pulmonary disease based on forced oscillation measurements and artificial neural networks. In *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, pages 1394–1397. IEEE, 2010.

[7] Almir Badnjevic, Lejla Gurbeta, and Eddie Custovic. An expert diagnostic system to automatically identify asthma and chronic obstructive pulmonary disease in clinical settings. *Scientific reports*, 8(1):11645, 2018.

[8] Samantha J Barry, Adrie D Dane, Alyn H Morice, and Anthony D Walmsley. The automatic recognition and counting of cough. *Cough*, 2(1):8, 2006.

[9] Francesco De Blasio, Johann C Virchow, Mario Polverino, Alessandro Zanasi, Panagiotis K Behrakis, Gunsely Kilinç, Rossella Balsamo, Gianluca De Danieli, and Luigi Lanata. Cough management: a practical approach. *Cough*, 7(1):7, 2011.

[10] Peter G Gibson, Anne B Chang, Nicholas J Glasgow, Peter W Holmes, Andrew S Kemp, Peter Katelaris, Louis I Landau, Stuart Mazzone, Peter Newcombe, Peter Van Asperen, et al. Cicada: Cough in children and adults: Diagnosis and assessment. australian cough guidelines summary statement. *Medical Journal of Australia*, 192(5):265–271, 2010.

[11] Päivi Piirilä and Ansi RA Sovijärvi. Differences in acoustic and dynamic characteristics of spontaneous cough in pulmonary diseases. *Chest*, 96(1):46–53, 1989.

[12] Miguel Angel Fernandez-Granero, Daniel Sanchez-Morillo, and Antonio Leon-Jimenez. An artificial intelligence approach to early predict symptom-based exacerbations of copd. *Biotechnology & Biotechnological Equipment*, 32(3):778–784, 2018.

[13] Archana Kanwade and VK Bairagi. Classification of copd and normal lung airways using feature extraction of electromyographic signals. *Journal of King Saud University-Computer and Information Sciences*, 2017.

[14] DA Newandee, SS Reisman, AN Bartels, and RE De Meersman. Copd severity classification using principal component and cluster analysis on hrv parameters. In *2003 IEEE 29th Annual Proceedings of Bioengineering Conference*, pages 134–135. IEEE, 2003.

[15] Daniel B Chamberlain, Rahul Kodgule, and Richard Ribón Fletcher. A mobile platform for automated screening of asthma and chronic obstructive pulmonary disease. In *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 5192–5195. IEEE, 2016.

[16] BM Rocha, D Filos, L Mendes, I Vogiatzis, E Perantoni, E Kaimakamis, P Natsiavas, A Oliveira, C Jácome, A Marques, et al. A respiratory sound database for the development of automated classification. In *Precision Medicine Powered by pHealth and Connected Health*, pages 33–37. Springer, 2018.

[17] Anton Ragni, Katherine Knill, Shakti P Rath, and Mark Gales. Data augmentation for low resource languages. 2014.

[18] Daniel S Park, William Chan, Yu Zhang, Chung-Cheng Chiu, Barret Zoph, Ekin D Cubuk, and Quoc V Le. Specaugment: A simple data augmentation method for automatic speech recognition. *arXiv preprint arXiv:1904.08779*, 2019.

[19] Navdeep Jaitly and Geoffrey E Hinton. Vocal tract length perturbation (vtlp) improves speech recognition. In *Proc. ICML Workshop on Deep Learning for Audio, Speech and Language*, volume 117, 2013.

[20] Iver Jordal. A python library for audio data augmentation. https://github.com/iver56/audiomentations. Accessed: 2020-04-21.

[21] Tudor Barbu. Variational image denoising approach with diffusion porous media flow. In *Abstract and Applied Analysis*, volume 2013. Hindawi, 2013.

[22] Siddharth Bhardwaj. Audio data augmentation with respect to musical instrument recognition, 2017.

[23] Theodoros Giannakopoulos and Aggelos Pikrakis. *Introduction to Audio Analysis: a MATLAB® approach*. Academic Press, 2014.

[24] Christopher Harte, Mark Sandler, and Martin Gasser. Detecting harmonic change in musical audio. In *Proceedings of the 1st ACM workshop on Audio and music computing multimedia*, pages 21–26, 2006.

[25] Francois Chollet. *Deep Learning with Python*. MITP-Verlags GmbH & Co. KG, 2018.

[26] Siwei Lai, Liheng Xu, Kang Liu, and Jun Zhao. Recurrent convolutional neural networks for text classification. In *Twenty-ninth AAAI conference on artificial intelligence*, 2015.

[27] Jesse Davis and Mark Goadrich. The relationship between precision-recall and roc curves. In *Proceedings of the 23rd international conference on Machine learning*, pages 233–240, 2006.

[28] Gary S Kendall. The decorrelation of audio signals and its impact on spatial imagery. *Computer Music Journal*, 19(4):71–87, 1995.

[29] Anthony K Akobeng. Understanding diagnostic tests 3: receiver operating characteristic curves. Acta paediatrica, 96(5):644–647, 2007.

[30] Leila Shahmoradi, Amos Otieno Olwendo, Hussein Arab-Alibeik, Khosrow Agin, et al. A probabilistic model for copd diagnosis and phenotyping using bayesian networks. Journal of Community Health Research, 6(1):34–43, 2017.

[31] Srivastava, Arpan, Sonakshi Jain, Ryan Miranda, Shruti Patil, Sharnil Pandya, and Ketan Kotecha. "Deep learning based respiratory sound analysis for detection of chronic obstructive pulmonary disease." PeerJ Computer Science 7 (2021).

[32] Ahmed J, Vesal S, Durlak F, Kaergel R, Ravikumar N, Rémy-Jardin M, Maier A, Tolxdorff T, Deserno T, Handels H, Maier A. "COPD classification in CT images using a 3D convolutional neural network." In: Maier-Hein K, Palm C, eds. Bildverarbeitung für Die Medizin 2020—Informatik Aktuell. Wiesbaden: Springer Vieweg, 39–45. 2020

[33] Leila Shahmoradi, Amos Otieno Olwendo, Hussein Arab-Alibeik, Khosrow Agin, et al. A probabilistic model for copd diagnosis and phenotyping using bayesian networks. Journal of Community Health Research, 6(1):34–43, 2017.