J. Integr. Sci. Technol. 2024, 12(4), 780 Article

Journal of Integrated
SCIENCE & TECHNOLOGY

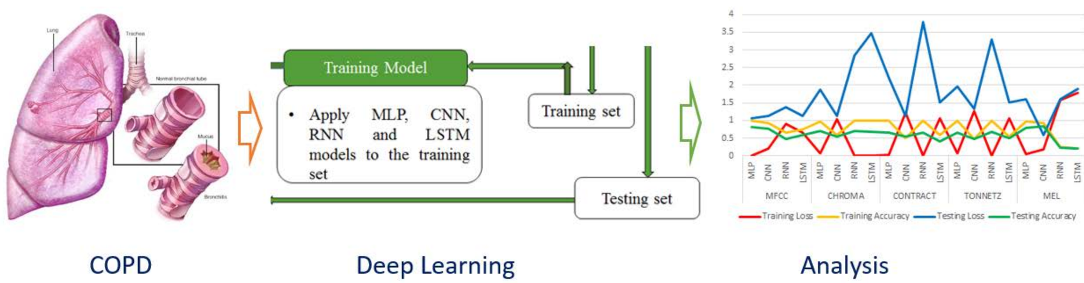
Multi feature fusion for COPD classification using Deep Learning algorithms

Pinal J. Patel[1*], Daksha Diwan[2], Kinjal A. Patel[3], Shashi Ranga[4], Niral J. Modi[5], Samay Dumasia[1]

[1]Computer Engineering Department, [2]General Department, Government Engineering College, Gandhinagar, Gujarat, India. [3]Faculty of Information Technology, Monash University, Clayton, Victoria, Australia. [4]Chemical Department, Government Engineering College, Valsad, Gujarat, India, [5]Government Engineering College, Gandhinagar, Gujarat Technological University, Gujarat. India.

Received on: 04-Sep-2023, Accepted and Published on:18-Dec-2023

ABSTRACT



Machine learning (ML) and deep learning (DL) are becoming pivotal for providing solutions to healthcare issues. Due to their accurate and quick forecasting models and discoveries, ML and DL algorithms are being used for disease classification by healthcare experts. Along with life-threatening illnesses like cancer, respiratory problems such as Chronic Obstructive Pulmonary Disease (COPD) have been growing more prevalent and endangering the survival of human society. According to the World Health Organization, COPD will be the third-leading cause of death and the seventh-leading cause of illness globally by 2030. Therefore, early detection and fast treatment are essential. The primary methods for diagnosing COPD need inadequate and pricy spirometer and imaging equipment. In this paper, an attempt is made to determine the severity of COPD disease using ML and DL algorithms using the cough sound of the patient. To extract audio features like Mfcc, Chroma, Contract, Mel, and Tonnetz, we have used the Librosa Python Library. To address the issues of imbalanced dataset, we have used the SMOTE algorithm. To find the most effective multi feature fusion for classifying COPD, numerous experiments have been carried out using various fusions of audio features. For the purpose of evaluating the multi-feature fusion's performance, we have run MLP, CNN, RNN, and LSTM models on fusion of two audio features and three audio features. Results of experiments suggest that the LSTM model with Adam as an optimization function gives 100% training accuracy and 87% testing accuracy for fusion of Mfcc and Mel features. As a result of the fusion of the three features of Tonnetz, Chroma, and Mel, CNN model performs better with training accuracy of 90% and testing accuracy of 82%.

Keywords: COPD, Deep Learning, Algorithms, Classification, Multi Feature Fusion

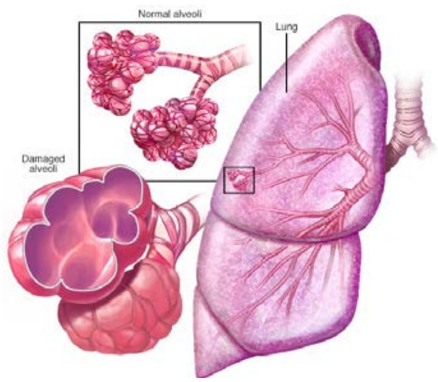
INTRODUCTION

According to the World Health Organization, chronic obstructive pulmonary disease (COPD) is the third most common cause of mortality in the world.[1] It affects 64 million people and results in an estimated 3.2 million deaths annually. Apart from it, the impact of COPD in terms of disability and reduced quality of life is substantial.[2] Because of widespread tobacco use,[3] environmental exposures such smoke from biomass fuels,[4] and an aging population,[5] prevalence is rising in both developing and developed nations. Comorbid illnesses and COPD are known to frequently coexist.[6,7] Although COPD is regarded to be a disease of later years, estimates suggest that 50% of persons with COPD are younger than 65 years old,[8] many of whom are likely to be in paid employment. Nearly 90% of deaths due to COPD are found in nations with low or middle incomes, where accessible or consistently applied effective prevention and control techniques are frequently lacking.. The chronic inflammatory lung disease COPD lowers the amount of airflow from the lungs.[9,10] Among the early warning signs and symptoms are wheezing, breathing issues, and sputum coughing. It develops as a result of repeated contact with irritating chemicals or particles, typically cigarette smoke. People with COPD have a higher risk of developing heart disease, lung

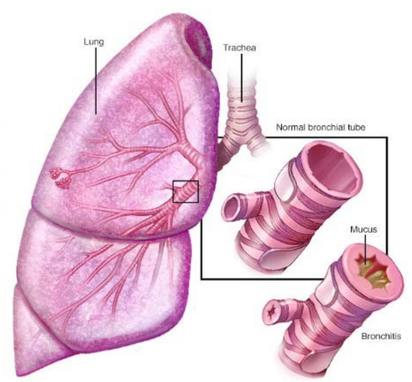
*Dr. Pinal J. Patel, Computer Enginerring Department, Government Engineering College, Gandhinagar, Gujarat, India
Email: pinalpatel@gecg28.ac.in

Cite as: *J. Integr. Sci. Technol., 2024, 12(4), 780. URN:NBN:sciencein.jist.2024.v12.780*

cancer, and a variety of other disorders. a variety of lung illnesses that progress. Emphysema and persistent bronchitis may lead to COPD.[11]



(a)



(B)

Figure1. (a) Emphysema (b) bronchitis under COPD.

Figure 1 depicts the COPD condition: (a) Emphysema (b) Chronic Bronchitis. Emphysema is a lung ailment carried on by damage to the lungs' alveolar walls. There's a chance that a clog will form, trapping air in the lungs. If there is too much air trapped in the lungs, the chest can seem bigger or barrel-chested. Less oxygen gets supplied to the bloodstream when there are fewer alveoli. Finally, patient unable to expel it. Chronic bronchitis is a condition marked by constant coughing that lasts longer than three months and appears more than twice in a two-year period. Smoking is a substantial cause for it and is frequently a contributing factor in chronic obstructive pulmonary disease.[12] Inflammation forms in the lining of the bronchial tubes, which transport oxygen to and from the lungs.

LITERATURE SURVEY

Acoustic signals related to the lungs have frequencies between 100 Hz and 2 kHz.[13] The human ear, however, is only sensitive to waves between 20 Hz and 20 kHz. Many disorders could be misdiagnosed or go unnoticed while using a manual stethoscope since you can't hear the corresponding respiratory sounds. The effectiveness of the tool, the expertise of the doctor, and the surrounding environment are frequently factors in the diagnosis of lung disorders.[14] As a result, electronic stethoscopes have been emerging progressively to take the role of conventional diagnostic equipment. Lung sounds can be stored as signals within a computer, enabling experts to examine these signals in time-frequency analysis with a more accurate interpretation.[14]

Many researchers have used machine and/or deep learning algorithms for automatic detection of respiratory disorders and the classification of lung sounds K-nearest neighbors (KNNs),[15] Support vector machines (SVMs),[16] artificial neural networks (ANNs), a naive Bayes classifier, are only a few of the models that have been used in machine learning. Convolutional neural networks (CNNs) were used by researchers to categorize respiratory sounds, and it has been demonstrated that the CNN model outperformed more common machine learning models (such as SVM and KNN) in terms of accuracy.[17] A combination of ANN and a back propagation-based Multi-Layer Perceptron algorithm has been used to forecast respiratory disease, mainly asthma and COPD.[18] Deep learning algorithms have been used to analyze respiratory sounds with the purpose of detecting chronic obstructive pulmonary disease.[19] Different stages of COPD patients were detected using ML Techniques. Researchers have also used CT scan images for detecting COPD. A 3D Convolutional Neural Network was used to classify COPD in CT images.[20] COPD has been recognized by utilizing deep CNN to generate 3D lung airway trees from CT images.[21] Analysis of the effect of cardiac color ultrasound on COPD under mask region was done using deep learning methods.[22] The present methods, such as manual diagnosis by a doctor, take a long time and numerous hospital visits to ascertain if a patient has COPD or not. By taking into account a variety of training factors, we have examined the effectiveness of deep learning approaches for determining the severity level of COPD disorders. For the investigation, we have used an audio dataset. The dataset is described in the section below.

DATABASE DESCRIPTION

Dataset comprises the patient's recorded lung sounds in various stages of COPD. Each patient's 12-channel lung sounds are included in the data collection. There are 504 .wav files and corresponding labels for it. Two pulmonologists used a Littmann3200 digital stethoscope to concurrently record the left (L) and right (R) channels in each lung region to gather the respiratory data. To collect the respiratory data, two pulmonologists used a Littmann3200 digital stethoscope to simultaneously record the left (L) and right (R) channels in each lung region. The dataset includes five different COPD severity levels: COPD0, COPD1, COPD2, COPD3, and COPD4. COPD0 denotes an absence of COPD and COPD4 denotes the most severe stage, each severity level corresponds to a particular stage of the progression of the illness. The lung sound recordings are short-term, lasting at least 17 seconds each. The recordings were captured using electronic stethoscopes and were collected from a diverse population of patients from around the world. The database includes lung sounds from 42 COPD patients, aged 38 to 68, with varied degrees of severity, including 34 men and 8 women. This dataset is intended to be used for the development and evaluation of deep learning models for COPD severity analysis, particularly for the classification of COPD severity levels using lung sound recordings. The dataset is publicly available and can be accessed for research purposes.[23] Turkey's Mustafa Kemal University's ethical committee has approved RespiratoryDatabase@TR. Table 1 depicts the no. of COPD recording per COPD severity levels.

Table 1: COPD audio dataset

Sr.no	COPD levels	Total number of recordings / level
1	COPD0	6
2	COPD1	5
3	COPD2	7
4	COPD3	7
5	COPD4	17

DESIGNED SYSTEM

Figure 2 shows the steps of the proposed system. Each step is discribed below.

(A) Data Acquisition

Wav audio files are used as an input during the data acquisition phase.

(B) Feature Extraction

Using the librosa library, we extracted various features from it, including mfcc, chroma, contract, mel, tonnetz. Details of the features is given below.

1) MFCC (Mel Frequency Cepstral Coefficients) is a feature extraction type commonly used in audio files.[24] MFCC is generally suggested to be used as an identifier for monosyllables in audio without identifying the speaker.[25] The MFCC feature extraction process in audio begins with the Pre-emphasis stage, namely amplifying the audio signal at high frequencies, followed by the framing and windowing stages, where framing stage aims to divide the length of the audio into several time intervals between 20 ms to 30 ms while the windowing technique is used to limit the occurrence of disturbances at the beginning and end of the audio. The next stage is the implementation of the Fast Fourier Transform, Mel Filter Bank, and Discrete Cosine Transform as a process of transforming the windowing results into MFCC. MFCC is a feature used in speech emotion recognition which has the advantage of representing the acoustic properties of the human voice. The MFCC uses the mel scale, which is similar to the human auditory perception of frequency. MFCC features are generated by taking the logarithm of the power spectrum and converting it to cestrum, thereby helping to reduce feature dimensionality and processing complexity. MFCC can represent temporal information in speech signals through short-duration frame splitting techniques to capture variations in speech signals associated with temporal emotional changes.

2) Chroma is a feature that focusing on music oriented audio tones.[26] This feature can provide a distribution of tonal variations in audio. The Chroma feature's result is a chromagram built based on 12 (twelve) tone levels.[27] The use of chroma is expected to recognize the high and low pitch of the actor's speech in audio, where the tone of the speech can indicate a certain type of emotion.

3) Mel (Mel-Spectrogram) is an audio feature that was built to overcome the problem of limited human hearing ability in distinguishing high-frequency values.[27] The use of the Mel-Spectrogram in this study is to extract information on differences in frequency values, particularly in identifying the types of emotions expressed by patients.

4) Tonnetz is a feature derived from Chroma that also focuses on audio harmony and tone classes.

5) Contrast is a feature in audio that is useful for estimating the average sound energy based on each sub-band's peak and valley spectral values.[28]

(C) Data Augmentation

For the Data Augmentation stage, we used the synthetic minority oversampling technique (SMOTE) to address the problem of imbalanced dataset.

(D) Preprocessing

The primary objective of this stage is to improve the data quality. The missing data has simply been eliminated. In the scaling and standardization steps, all values are normalized to fall between [1, 1]. For label encoding, we used one hot encoding method.

(E) Training and Testing Set

The dataset is finally divided into train and test sets. On the train dataset, we have run the four models - CNN, RNN, MLP, and LSTM. The accuracy of the trained models is evaluated using test data.The architecture of the model for the fusion of two and three features is shown in figures 3 and 4 respectively.

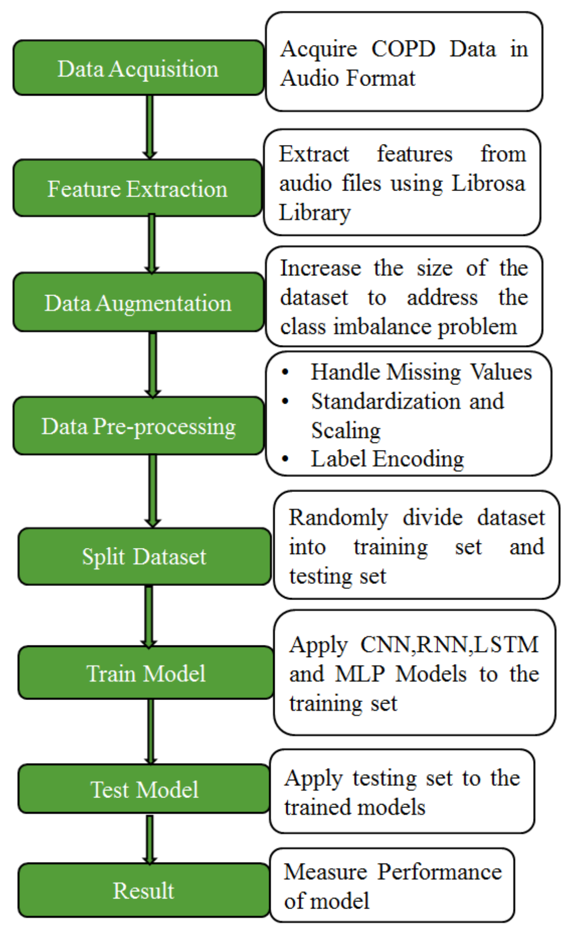


Figure 2. Designed system for audio dataset

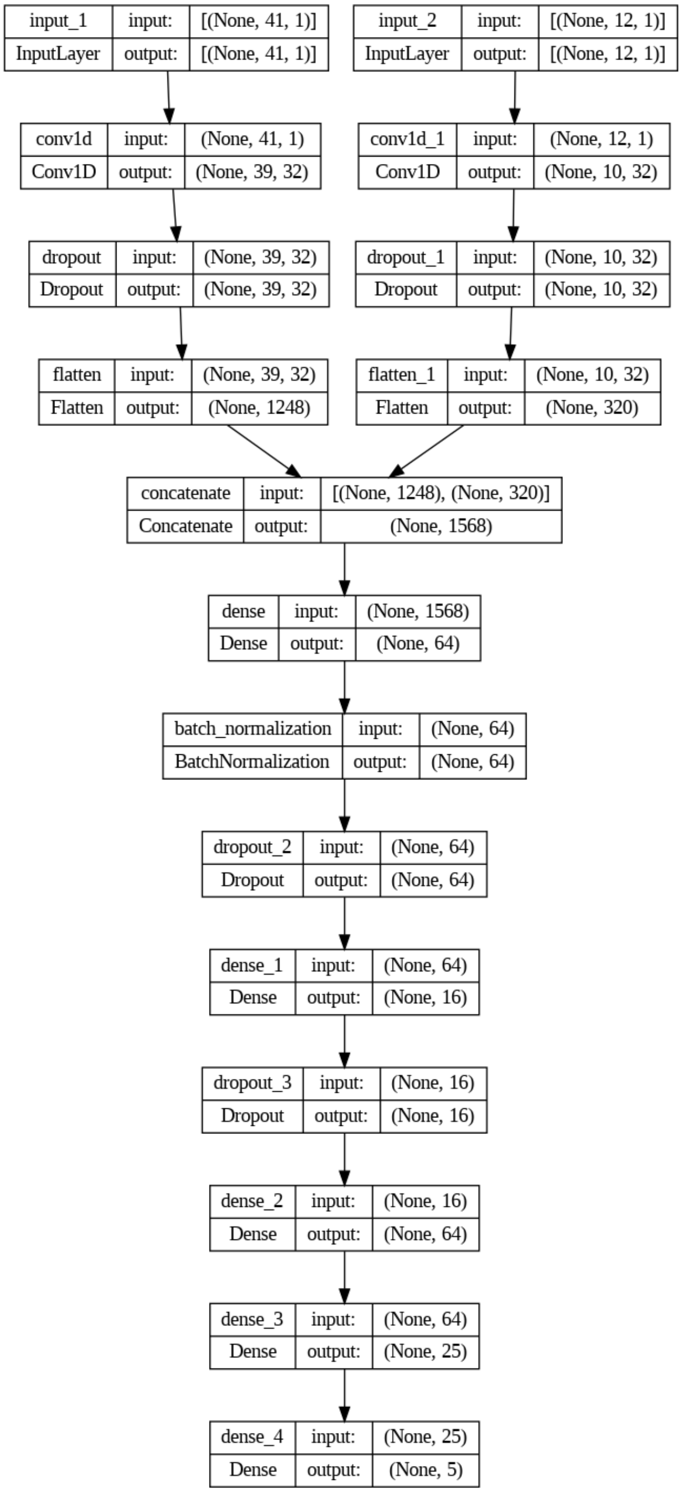


Figure 3. Architecture of the CNN model for fusion of Mfcc and Chroma features

RESULTS

An epoch is the total number of iterations of all the training data in a single cycle. Loss function is a function for figuring out the difference or error between expected and actual data. In order to minimize a loss function, optimizers are algorithms that change the model's parameters during training. We have used ADAM and SGD optimizers. The number of training instances in a batch is referred to as the batch size. Different experiments have been conducted by taking different values of the training parameters like epoch, loss function, optimizer, and batch size.

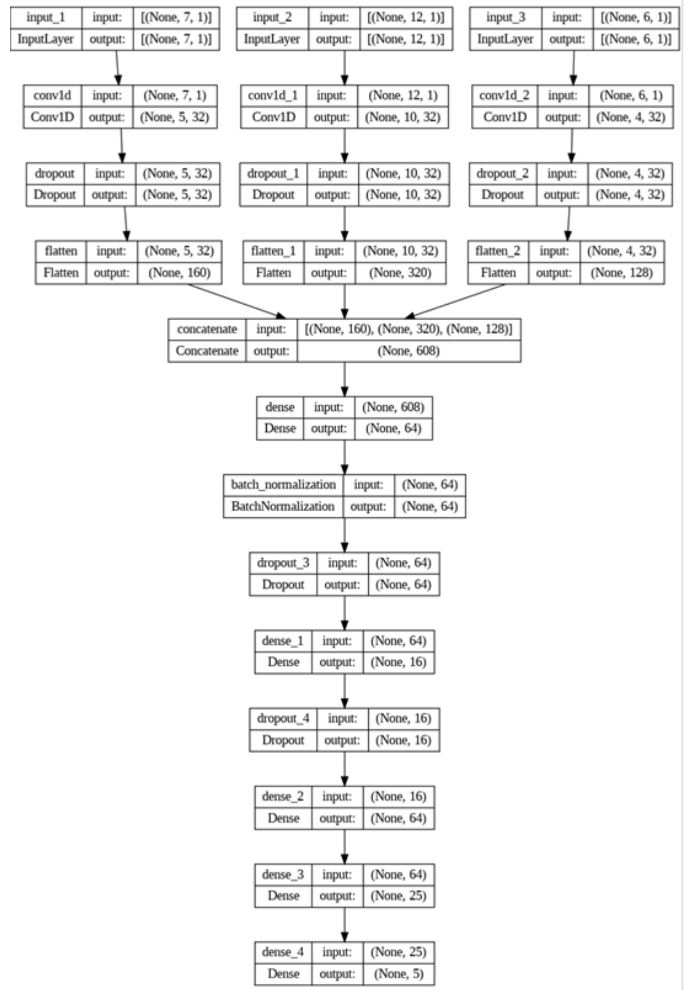


Figure 4. Architecture of the CNN model for fusion of Contrast, Chroma, Tonnetz features

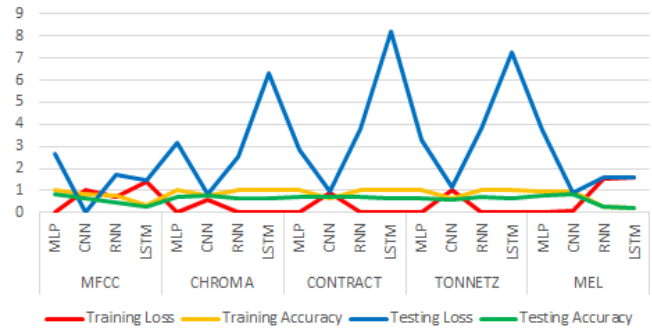
Initially, we have run DL models on the individual audio features using SGD and Adam as an optimizers with different batch sizes. Figure 5 depicts the result. Figure 6 shows the accuracy and loss curves of the MLP model with different batch sizes and optimizers for the individual audio features. Then we have combined two features of the audio file and run different DL algorithms on it to find the suited feature fusion for detecting COPD levels. We started merging mfcc with chroma, mel, contract, toneetz. Result is shown in table 2. We have also combined three features of the audio file, run DL algorithms on it, gather the result. Table 3 depicts the result of it. Figure 7 and figure 8 show accuracy and loss curves of the two feature fusion and three feature fusion using DL models respectively.

Table 2: Performance of the DL algorithms 1000 epoch, Optimizer Adam, Loss function categorical cross entropy, batch size 16

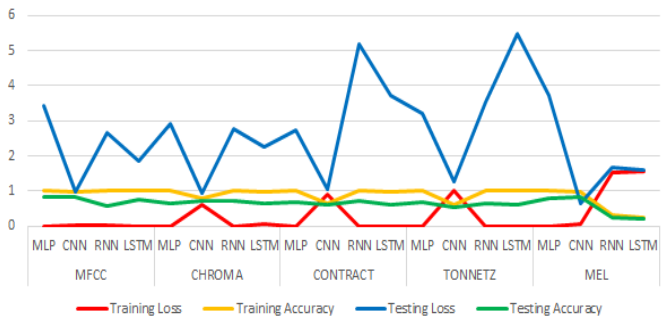
Audio feature name	Model Name	Training Loss	Training Accuracy	Testing Loss	Testing Accuracy
Mfcc, Chroma	MLP	0.0008	1.0000	3.2931	0.7353
	CNN	0.3760	0.8689	0.6925	0.7157
	RNN	0.0063	1.0000	5.0404	0.2598
	LSTM	0.0021	1.0000	7.7174	0.3775
Mfcc, Mel	MLP	0.0058	1.0000	4.5119	0.7402
	CNN	0.2679	0.9118	1.3176	0.5588
	RNN	0.2541	0.9228	2.6436	0.4118
	LSTM	0.0010	1.0000	2.0942	0.8725
Mfcc, Contrast	MLP	0.0022	1.0000	2.3289	0.7892
	CNN	0.3271	0.8860	0.8887	0.6618
	RNN	0.0056	1.0000	11.5652	0.1961
	LSTM	0.0001	1.0000	8.7462	0.4461
Mfcc, tonnetz	MLP	0.0002	1.0000	1.6445	0.7892
	CNN	0.3727	0.8799	0.5106	0.8039
	RNN	0.1508	0.9645	6.0813	0.2402
	LSTM	0.0001	1.0000	5.6773	0.5637
Chroma, mel	MLP	0.0001	1.0000	2.2512	0.6814
	CNN	0.3315	0.8848	0.6607	0.7304
	RNN	0.0006	1.0000	1.7304	0.7206
	LSTM	0.0001	1.0000	3.8097	0.7206
Chroma, Contrast	MLP	0.0013	1.0000	2.6907	0.7059
	CNN	0.7376	0.7304	1.6319	0.3922
	RNN	0.0001	1.0000	2.5774	0.6814
	LSTM	0.0026	1.0000	6.4665	0.3873
Chroma, Tonnetz	MLP	0.0004	1.0000	1.6113	0.7402
	CNN	0.7979	0.7108	1.3373	0.4706
	RNN	0.0006	1.0000	1.4076	0.7157
	LSTM	0.0006	1.0000	8.8786	0.5147
Mel, Contrast	MLP	0.0002	1.0000	1.2164	0.8088
	CNN	0.3181	0.8934	0.4285	0.8529
	RNN	0.0005	1.0000	1.5043	0.7843
	LSTM	0.0005	1.0000	5.5939	0.6275
Mel, Tonnetz	MLP	0.0005	1.0000	1.1257	0.7941
	CNN	0.3469	0.8848	0.3281	0.8676
	RNN	0.0002	1.0000	1.2916	0.8137
	LSTM	1.5558	0.2426	2.1263	0.0001
Contrast, Tonnetz	MLP	0.0005	1.0000	2.6259	0.7255
	CNN	0.8337	0.6814	1.3721	0.4118
	RNN	0.0030	1.0000	1.6326	0.7304
	LSTM	4.06120	1.0000	7.2401	0.5098

Table 3 : Performance of the DL algorithms 1000 epoch, Optimizer Adam, Loss function categorical cross entropy, batch size 16

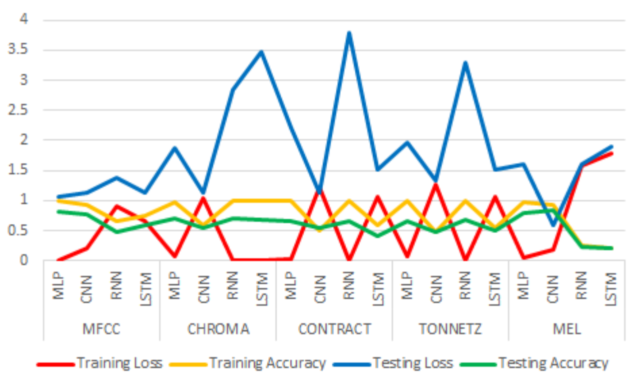
Audio feature name	Model Name	Training Loss	Training Accuracy	Testing Loss	Testing Accuracy
Mfcc, Contrast, Mel	MLP	0.0009	1.0000	5.0919	0.5931
	CNN	0.2299	0.9216	1.4824	0.5588
	RNN	0.0001	1.0000	12.9605	0.2647
	LSTM	1.5559	0.2402	2.1256	0.0001
Mfcc, Mel, Tonnetz	MLP	0.0001	1.0000	5.9857	0.6029
	CNN	0.2439	0.9228	1.6958	0.4902
	RNN	0.1408	0.9522	9.3978	0.2108
	LSTM	0.1592	0.9534	9.7502	0.1814
Mfcc, Contrast, Mel	MLP	0.0001	1.0000	5.6984	0.5490
	CNN	0.1773	0.9363	1.4839	0.5931
	RNN	0.0001	1.0000	12.9605	0.2647
	LSTM	1.5559	0.2439	2.1248	0.0001
Mfcc, Chroma, Contrast	MLP	0.0003	1.0000	3.6328	0.6961
	CNN	0.3527	0.8725	0.6949	0.7255
	RNN	0.0008	1.0000	15.9124	0.2549
	LSTM	0.0007	1.0000	13.8600	0.3529
Mfcc, Chroma, Tonnetz	MLP	0.0002	1.0000	3.1849	0.7255
	CNN	0.3352	0.8824	0.8002	0.6863
	RNN	0.0010	1.0000	8.2455	0.2500
	LSTM	1.5559	0.2463	2.1275	0.0001
Mfcc, Contrast, Tonnetz	MLP	0.0002	1.0000	1.8531	0.7745
	CNN	0.3383	0.8676	0.6640	0.7157
	RNN	0.0010	1.0000	8.2455	0.2500
	LSTM	0.0039	1.0000	5.6198	0.2990
Contrast, Chroma, Mel	MLP	0.0007	1.0000	2.0598	0.7255
	CNN	0.2978	0.9044	0.6620	0.7696
	RNN	0.0002	1.0000	2.0019	0.7745
	LSTM	0.0005	1.0000	8.6340	0.5294
Tonnetz, Chroma, Mel	MLP	0.0001	1.0000	2.4011	0.7451
	CNN	0.2716	0.9056	0.5743	0.8235
	RNN	0.0006	1.0000	0.8360	0.8137
	LSTM	1.5559	0.2390	2.1267	0.0001
Contrast, Chroma, Tonnetz	MLP	0.0002	1.0000	1.8047	0.7745
	CNN	0.6204	0.7917	1.4565	0.4608
	RNN	0.0002	1.0000	0.9061	0.7696
	LSTM	0.0001	1.0000	7.3003	0.5098
Contrast, Tonnetz, Mel	MLP	0.0001	1.0000	1.6484	0.7696
	CNN	0.3227	0.8934	0.4534	0.8284
	RNN	0.0008	1.0000	1.7168	0.7647
	LSTM	0.0004	1.0000	5.9438	0.5588



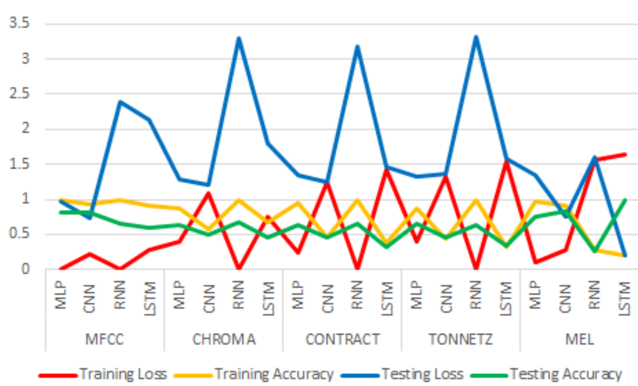
(a) Batch Size 16, ADAM with all DL models



(b) Batch Size 32, ADAM with all DL models

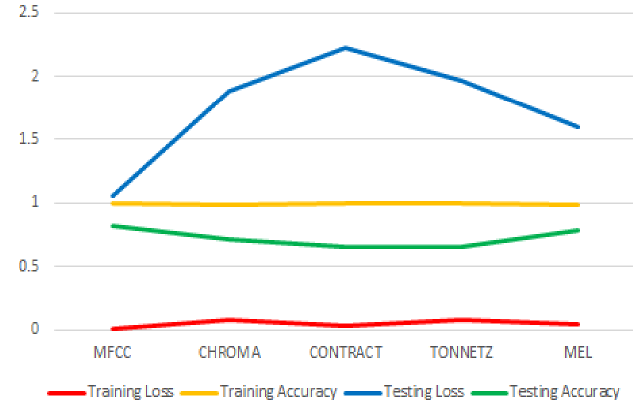


(c) Batch Size 16, SGD with all DL models

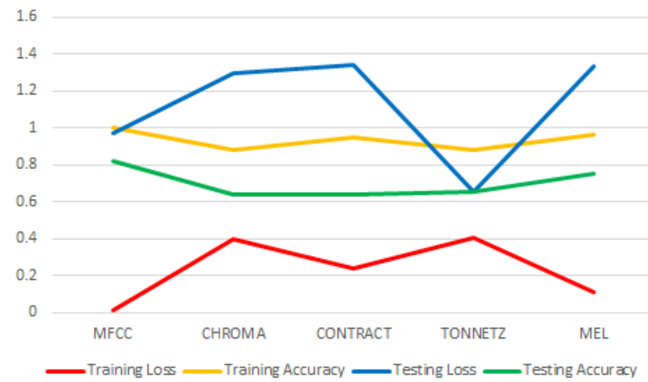


(d) Batch Size 32, SGD with all DL models

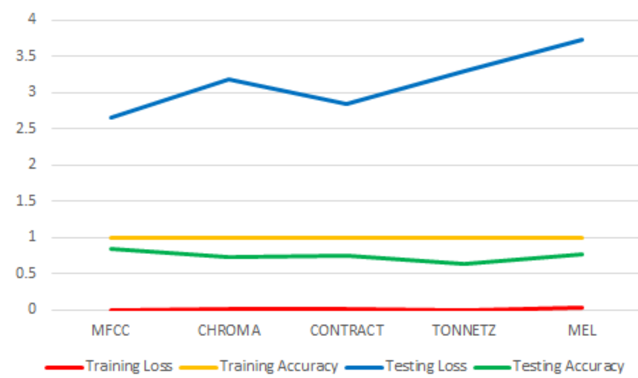
Figure 5: Batch wise performance curves of all individual five features with Adam and SGD as an optimization functions



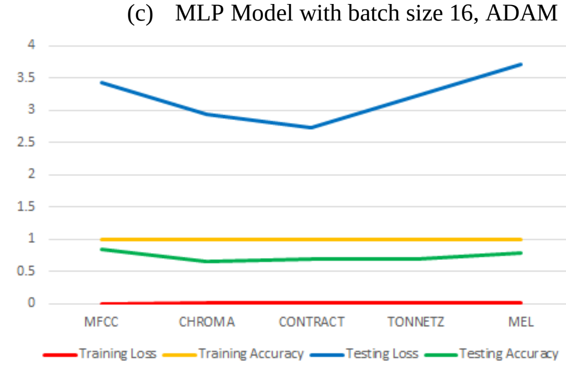
(a) MLP Model with batch size 16 , SGD



(b) MLP Model with batch size 32, SGD

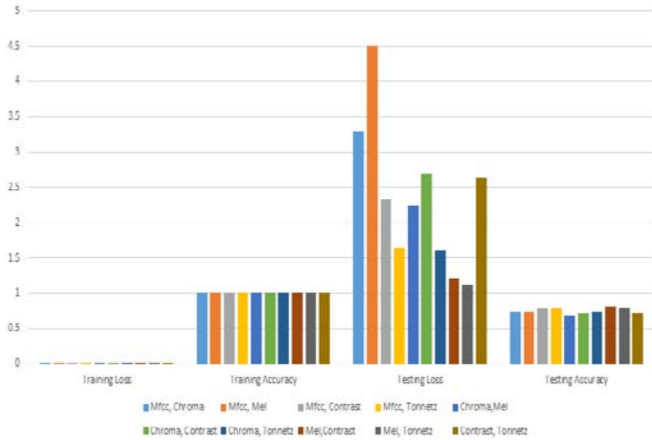


(c) MLP Model with batch size 16, ADAM

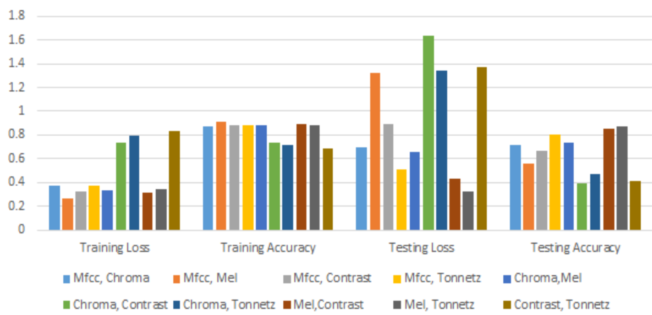


(d) MLP Model with batch size 32, ADAM

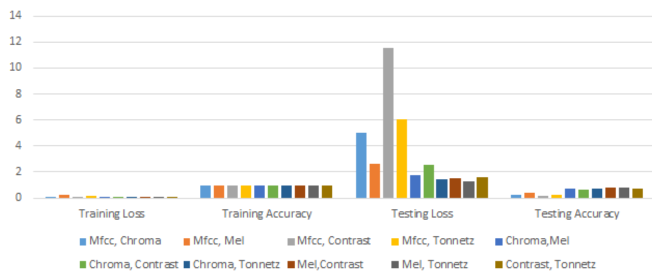
Figure 6: Batch wise performance curves of all five features using MLP model with Adam and SGD as an optimization functions



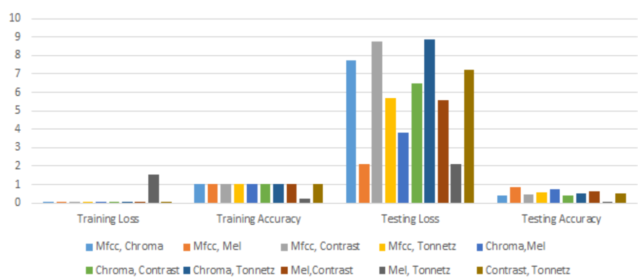
(a) MLP Model



(b) CNN Model



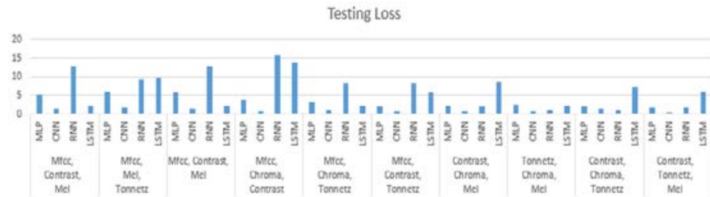
(c) RNN Model



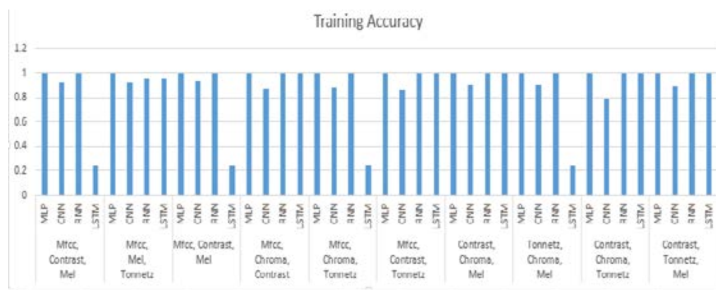
(d) LSTM Model

Figure 7: Accuracy and loss curves of DL models for two feature fusion for (a) MLP Model (b) CNN Model (c) RNN Model (d) LSTM Model

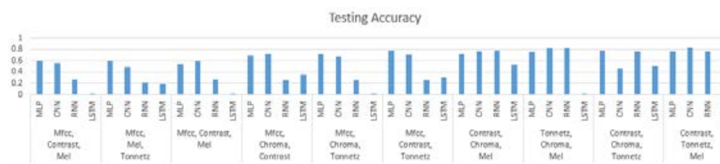
(a) Training Loss



(b) Testing Loss



(c) Training Accuracy



(d) Testing Accuracy

Figure 8: Accuracy and loss curves of DL models for three feature fusion for (1) MLP Model (2) CNN Model (3) RNN Model (4) LSTM Model

Discussion

We have analysed the cough sounds of patients using the librosa python library and determined the severity levels of COPD disease using Deep Learning algorithms.[29] In order to investigate the efficacy of different deep learning algorithms such as CNN, RNN, LSTM, and MLP for classifying COPD severity levels using audio dataset, a number of experiments have been carried out with varying values for the learning parameters, such as batch size, epochs, etc. during the model-learning process. We have extracted five different features including Mfcc, Chroma, Mel, Tonnetz, and Contrast using the librosa python library from audio of the patient. To handle the issue of the imbalanced samples of the severity levels of the COPD disease dataset, we have used the SMOTE algorithm.

We have run DL algorithms on the fusion of two audio features and three audio features. The findings show that the LSTM model using Adam as the optimization function provides 100% training accuracy and 87% testing accuracy for the fusion of Mfcc and Mel features. CNN model with ADAM as an optimization function provides 90% of training accuracy and 82% of the testing accuracy. In the absence of feature fusion, the MLP model with Adam as the optimization function provides 100% training accuracy and 84% testing accuracy. Better results are obtained when training models are learned with a batch size of 32.

CONCLUSION

We have conducted numerous experiments by taking different values for the learning parameters of the models like epoch, batch size, optimization functions etc. to analyze the performance of the various DL algorithms such as CNN, RNN, LSTM, and MLP for classifying COPD severity levels. We have run DL models using combinations of the audio features. Result of experiments suggests that LSTM model with Adam as an optimization function give 100% training accuracy and 87% testing accuracy for fusion of Mfcc and Mel features. Three features fusion of Tonnetz, Chroma and Mel with CNN model gives training accuracy 90% and testing accuracy 82%. In the absence of the feature fusion, MLP model with Adam as an optimization function provides training accuracy 100% and testing accuracy 84%. Training models provide better result when they are learnt with batch size 32. The findings of the experiments indicate that the amalgamation of various features enhances the accuracy of the model by taking suitable values for the model parameters.

CONFLICT OF INTEREST STATEMENT

We (authors) do not have any conflict of interest (financial or academic) for this work.

REFERENCES

1. C.D. Mathers, D. Loncar. Projections of global mortality and burden of disease from 2002 to 2030. *PLoS Med.* **2006**, 3 (11), 2011–2030.
2. M. Ferrer, J. Alonso, J.M. Anto. Relationship between chronic obstructive pulmonary disease stage and health-related quality of life. *Cardiol. Rev.* **1998**, 15 (10), 42–45.
3. D.M. Mannino, A.S. Buist. Global burden of COPD: risk factors, prevalence, and future trends. *Lancet* **2007**, 370 (9589), 765–773.
4. Y. Hu, B. Chen, Z. Yin, et al. Increased risk of chronic obstructive pulmonary diseases in coke oven workers: Interaction between occupational exposure and smoking. *Thorax* **2006**, 61 (4), 290–295.
5. P. Burney, S. Suissa, J.B. Soriano, et al. The pharmacoepidemiology of COPD: Recent advances and methodological discussion. *Eur. Respir. Journal, Suppl.* **2003**, 22 (43), 1 –44.
6. L.M. Fabbri, F. Luppi, B. Beghé, K.F. Rabe. Complex chronic comorbidities of COPD. *Eur. Respir. J.* **2008**, 31 (1), 204–212.
7. T. Goto, C.A. Camargo, M.K. Faridi, B.J. Yun, K. Hasegawa. Machine learning approaches for predicting disposition of asthma and COPD exacerbations in the ED. *Am. J. Emerg. Med.* **2018**, 36 (9), 1650–1654.
8. A.S. Buist, M.A. McBurnie, W.M. Vollmer, et al. International variation in the prevalence of COPD (The BOLD Study): a population-based prevalence study. *Lancet* **2007**, 370 (9589), 741–750.
9. V.K. Shanwal. Assessing impact of Air Pollution on behavior of school children in Greater Noida, India. *J. Integr. Sci. Technol.* **2021**, 9 (1), 1–8.
10. M. Shrivastava, P. Sharma, R. Singh. Identification of potential CYP51 inhibiting anti-Aspergillus phytochemicals using molecular docking and ADME/T studies. *Chem. Biol. Lett.* **2021**, 8 (1), 18–21.
11. C. Pal. Molecular mechanism facets of Oxidative stress mediated pathogenesis. *J. Mol. Chem.* **2023**, 3 (2 SE-Medicinal Chemistry), 587.
12. F. Mejza, L. Gnatiuc, A.S. Buist, et al. Prevalence and burden of chronic bronchitis symptoms: results from the BOLD study. *Eur. Respir. J.* **2017**, 50 (5), 1700621.
13. V. Gross, A. Dittmar, T. Penzel, F. Schüttler, P. Von Wichert. The relationship between normal lung sounds, age, and gender. *Am. J. Respir. Crit. Care Med.* **2000**, 162 (3 I), 905–909.
14. A.P. Benguerel. Signals and Systems for Speech and Hearing. *Language and Speech*. Brill, Leiden 1991, pp 381–382.
15. G. Serbes, C.O. Sakar, Y.P. Kahya, N. Aydin. Pulmonary crackle detection using time-frequency and time-scale analysis. *Digit. Signal Process. A Rev. J.* **2013**, 23 (3), 1012–1021.
16. F. Jin, F. Sattar, D.Y.T. Goh. New approaches for spectro-temporal feature extraction with applications to respiratory sound classification. *Neurocomputing* **2014**, 123, 362–371.
17. D. Bardou, K. Zhang, S.M. Ahmad. Lung sounds classification using convolutional neural networks. *Artif. Intell. Med.* **2018**, 88, 58–69.
18. K.L. Khatri, L.S. Tamil. Early Detection of Peak Demand Days of Chronic Respiratory Diseases Emergency Department Visits Using Artificial Neural Networks. *IEEE J. Biomed. Heal. Informatics* **2018**, 22 (1), 285–290.
19. A. Srivastava, S. Jain, R. Miranda, et al. Deep learning based respiratory sound analysis for detection of chronic obstructive pulmonary disease. *PeerJ Comput. Sci.* **2021**, 7, 1–22.
20. J. Ahmed, S. Vesal, F. Durlak, et al. COPD classification in CT images using a 3D convolutional neural network. *Informatik aktuell*. 2020, pp 39–45.
21. R. Du, S. Qi, J. Feng, et al. Identification of COPD from Multi-View Snapshots of 3D Lung Airway Tree via Deep CNN. *IEEE Access* **2020**, 8, 38907–38919.
22. Y. He, F. Chen, Q. Chen. Deep Learning-Based Analysis of the Effect of Cardiac Color Ultrasound on Chronic Obstructive Pulmonary Disease under Mask Region. *Sci. Program.* **2021**, 2021, 1–7.
23. G. Altan, Y. Kutlu, Y. Garbi, A.Ö. Pekmezci, S. Nural. Multimedia Respiratory Database (RespiratoryDatabase@TR): Auscultation Sounds and Chest X-rays. *Nat. Eng. Sci.* **2017**, 2 (3), 59–72.
24. R.M. Hanifa, K. Isa, M. Mohamad. Comparative Analysis on Different Cepstral Features for Speaker Identification Recognition. *2020 IEEE Student Conf. Res. Dev. SCOReD 2020* **2020**, No. September, 487–492.
25. S. Ajibola Alim, N. Khair Alang Rashid. Some Commonly Used Speech Feature Extraction Algorithms. In *From Natural to Artificial Intelligence - Algorithms and Applications*; **2018**.
26. J.V.T. Abraham, A.N. Khan, A. Shahina. A deep learning approach for robust speaker identification using chroma energy normalized statistics and mel frequency cepstral coefficients. *Int. J. Speech Technol.* **2023**, 26 (3), 579–587.
27. U. Garg, S. Agarwal, S. Gupta, R. Dutt, D. Singh. Prediction of Emotions from the Audio Speech Signals using MFCC, MEL and Chroma. *Proc. - 2020 12th Int. Conf. Comput. Intell. Commun. Networks, CICN 2020* **2020**, 87–91.
28. B. Kabra, C. Nagar. Attention-Emotion-Embedding BiLSTM-GRU network based sentiment analysis. *J. Integr. Sci. Technol.* **2023**, 11 (4), 563.
29. W. Xie, Y. Fang, G. Yang, K. Yu, W. Li. Transformer-Based Multi-Modal Data Fusion Method for COPD Classification and Physiological and Biochemical Indicators Identification. *Biomolecules* **2023**, 13 (9), 1391.