

Cascaded PFLANN Model for Intelligent Health Informatics in Detection of Respiratory Diseases from Speech Using Bio-inspired Computation

Jagannath Dayal Pradhan,¹ L. V. Narasimha Prasad,² Tusar Kanti Dash,¹
Manisha Guduri,³ and Ganapati Panda¹

¹Electronics and Communications Engineering, C V Raman Global University, Bhubaneswar, India

²Department of Computer Science and Engineering, Institute of Aeronautical Engineering, Hyderabad, India

³School of Computing and Informatics, University of Louisiana, Lafayette, LA, USA

(Received 29 July 2023; Revised 13 September 2023; Accepted 11 November 2023; Published online 21 February 2024)

Abstract: Due to the recent developments in communications technology, cognitive computations have been used in smart healthcare techniques that can combine massive medical data, artificial intelligence, federated learning, bio-inspired computation, and the Internet of Medical Things. It has helped in knowledge sharing and scaling ability between patients, doctors, and clinics for effective treatment of patients. Speech-based respiratory disease detection and monitoring are crucial in this direction and have shown several promising results. Since the subject's speech can be remotely recorded and submitted for further examination, it offers a quick, economical, dependable, and noninvasive prospective alternative detection approach. However, the two main requirements of this are higher accuracy and lower computational complexity and, in many cases, these two requirements do not correlate with each other. This problem has been taken up in this paper to develop a low computational complexity-based neural network with higher accuracy. A cascaded perceptual functional link artificial neural network (PFLANN) is used to capture the nonlinearity in the data for better classification performance with low computational complexity. The proposed model is being tested for multiple respiratory diseases, and the analysis of various performance matrices demonstrates the superior performance of the proposed model both in terms of accuracy and complexity.

Keywords: AI-ML; ASR; FLANN; health informatics; neural network; PFLANN

I. INTRODUCTION

With the development of cognitive-inspired computing and interaction systems, cognition is emerging as a promising paradigm that has the potential to transform our way of life. The process through which a system develops strong adaptive, anticipatory, and autonomous behavior that involves embodied perception and action is known as cognition [1]. These techniques are widely used in distributed denial of service defense mechanisms, sentiment analysis and classification, 5G communications, and healthcare applications [2–5]. The Internet of Things (IoT) is now widely used in various applications, and as a result, its significance in daily life is growing. For the purpose of effectively supplying patients with emergency services, IoT technology is also being developed in the healthcare monitoring system [6]. The IoT revolution has also changed how modern healthcare systems and management are organized. The greatest promise of IoT is to significantly advance the field of healthcare [7]. Several techniques including Radio-Frequency Identification, Edge Computing, Semantics, Cloud Computing, Big Data, Grid Computing, and Augmented Reality are used for IoT-based healthcare systems, and it has been observed that IoT-based healthcare will increase day by day in the upcoming years [8].

Remote health monitoring systems now have a much wider range of applications because of the quick advancement of technology. Due to the low doctor-to-patient ratio, the primary goal of this study is to make it easier for cardiac patients who live in distant areas to access the most up-to-date medical care [9]. The body area sensor network framework is intended to be adopted by a wearable IoT-cloud-based health monitoring system in support of real-time health monitoring. The heartbeat, body temperature, and blood pressure sensors are only a few of the embedded wearable sensors [10]. A real-time health monitoring system called HealthSOS is presented for stroke prognostics. Thirty-seven stroke victims who had been admitted to a hospital's emergency room and 36 senior volunteers in good health were used to test this technique [11]. Environmental causes and associated air pollution are thought to be major contributors to the rise in respiratory disorders, particularly asthma, which is still a major issue. These disorders are frequently accompanied by symptoms like airway obstruction, chest discomfort or pain, cough, wheezes, or other strange sounds made when breathing [12].

A method for characterizing and detecting respiratory diseases is present using respiratory sounds saved in audio format with a supervised model. Two classifier models are used where a feature vector is an input for a second classifier intended to characterize the lung disease if the patient is diagnosed with a (generic) lung disease [13]. In another work, an effective deep learning framework for auscultation analysis is presented and examined in this research.

Corresponding author: Jagannath Dayal Pradhan (e-mail: jdpradhan@cgu-odisha.ac.in).

This uses respiratory sound recordings to categorize abnormalities in respiratory cycles and identify illnesses. The spectrogram characteristics are then categorized into categories of respiratory abnormality cycles or diseases using a back-end deep learning network [14]. To compare the classification of interstitial lung disease patterns using shallow and deep learning, a convolution neural network (CNN) is used with six learnable layers, four of which were convolutional and two of which were completely linked.

An objective of study was to develop and validate a classification algorithm for the early identification of patients, with a background of chronic obstructive pulmonary disease (COPD), who appear to be at high risk of an imminent exacerbation event [15]. A secured custom reusable intellectual property core for the CNN using facial biometrics has been proposed in ref. [16]. A multilayer perceptron-based artificial neural network classifier model using a backpropagation algorithm is used to predict the peak events in hospital emergency departments for respiratory diseases in Dallas County of Texas due to respiratory diseases. The ANN model achieves an overall accuracy of 81% in predicting peak and nonpeak events [17].

In modern healthcare, audio classification can be used for cardiac assessment and respiratory check-ups using a stethoscope and sonography. This study is taken for heart and lung disease detection with 85% accuracy for helping doctors in making better diagnoses using the hidden Markov model [18]. Computational complexity plays a crucial role in the implementation of real-time healthcare systems. For remote cardiovascular monitoring, a simple method with low computational complexity for extracting the fiducial points from the electrocardiogram has been developed. Here, continuous sensing and processing are performed in computationally limited, low-power devices; therefore, power consumption and algorithm complexity should be kept to a minimum [19]. Big healthcare data analytics that use huge and diverse data have a trade-off between the value of conventional hypothesis-driven inference and statistical significance with computational efficiency, protocol complexity, and methodological validity [20]. In the past, decisions on patient care were solely dependent on the doctor's professional experience, subject-matter expertise, physical symptoms, and diagnostic test results. In contrast, new tools, items, and technologies have emerged that play a vital part in healthcare monitoring and aid medical professionals in making wiser decisions. In U-healthcare monitoring system topologies, the cloud paradigm serves as the foundation for on-demand network utilization of a shared pool of reconfigurable computing resources. However, it has several drawbacks that are resolved by fog [21].

Deep neural networks have a very small computational and power budget; therefore, their adoption on low-power-embedded devices is hampered in particular by the memory and computing needs. A second rapid convolution stage is added to the first stage of computational intensity reduction using a modified Toom-Cook technique. At this point, there are far fewer powerful operations overall without any approximations that might compromise precision. Strong proof that a large speedup may be obtained without sacrificing baseline accuracy by combining optimizing the network's topology and the underlying implementation of basic convolution operations is provided by the evaluation's findings running on a variety of real hardware [22]. In another interesting study, it is observed that neural networks take higher computational complexity for training. The importance of computational complexity has been studied with both positive and negative results to

provide new provably efficient and practical algorithms for training certain types of neural networks [23]. Using a neural network expert system, computationally effective classifiers for classifying the Frank lead ECG as normal or one of six illness states are studied [24]. The Daubechies wavelet of order 10 (dB 10) and the Symlet wavelet of order 8 (sym8) applied to a single beat of the X, Y, and Z Frank leads are two discrete wavelets included in this. These findings were obtained at a substantially lower computing cost than those obtained from neural networks trained with more than 229 scalar parameters. Cascaded neural networks are being used instead of a single neural network for higher accuracy, and it has been used for several tasks including land-use classification, improved protein contact map prediction, face detection, and speech emotion recognition [25–28].

A. RESEARCH GAP AND CONTRIBUTION

It has been observed from the literature review that in the domain of smart healthcare using IoT, cognitive-inspired computing plays a crucial role. Real-time healthcare monitoring is an important criterion specifically in the case of respiratory diseases. To achieve this, many speech recognition models are designed. However, very few works are reported in the context of reducing the computational complexity without compromising on the accuracy. To solve this problem, a cascaded neural network architecture is proposed in this paper using a perceptual functional link artificial neural network (PFLANN). Based on the research gap, the main contributions of the paper are the use of cascaded PFLANN for effectively capturing the perceptual nonlinearity of human speech perception technique, training and testing the model with eight datasets covering multiple respiratory diseases, reduction of the computational complexity of the model, and statistical analysis of the classifier models used in multiple respiratory diseases and multiple datasets.

The paper is divided into four sections, with Section I covering the introduction, literature overview, research motives, and research goals. Section II deals with the specifics of the tools and techniques used. The results analysis and contributions in terms of study findings are found in Section III. Section IV presents the research's findings, restrictions, and areas of potential future study.

II. MATERIAL AND METHODS

A. DATASETS

The dataset-1 contains the voiced samples for the diseases such as hyperfunctional dysphonia, functional dysphonia, laryngitis, vocal fold polyps, and leukoplakia. The samples are taken from Saarbruecken Voice Database which is created by Phonetics at the University of Saarland, Germany, which includes voice samples of more than 2000 people. The vowels /a/, /i/, and /u/ at normal pitch are recorded 259 times each in healthy and sick conditions with a maximum duration of 4 seconds, sampling at a rate of 50 kHz, and 16 bits per sample with equal number of male and female voices. For the study of COVID-19, five datasets have been used, including Coswara (dataset-2) [29], Crowd-sourced University of Cambridge (dataset-3) [30], Virufy (dataset-4) [31], recorded interviews from online platforms in telephone quality speech (dataset-5) [32], and Coughvid (dataset-6) [33]. These datasets contain speech samples of subjects from more than 52 countries. A total of 4398 speech samples have been used in the simulation study which includes 280, 796, 636, 121, 702, and 2310

samples from datasets 1, 2, 3, 4, 5, and 6, respectively, with approximately equal numbers of speech samples from healthy and diseased categories. Dataset-7 has been taken from the respiratory and drug actuation dataset which was recorded at the University of Patras under a controlled environment [34]. The audio signals have been recorded during breathing and drug actuation. A total of 370 audio files have been recorded with an 8-KHz sampling frequency and 8-bit depth in WAV format. For experimentation, in the current implementation, 193 drug actuation segments, 319 inhalation, and 620 exhalation segments are used. The dataset-8 is collected from the ALS-D-Net dataset which was recorded at Fortis Hospital, India. The pulmonary sound dataset is divided into two categories: abnormal (pulmonary patients) and normal (healthy). An e-stethoscope that was connected to a laptop with an amplifier to record the lung sound was used. An electronic stethoscope with a chest piece that contacts the subject and a microphone to record the signal was used. The audio samples are recorded in WAV format with a 44.1-kHz sampling rate, 16 bits per sample, and 10 s duration of each file.

B. PFLANN MODEL

FLANN is one of the computationally efficient and effective forms of the traditional neural network model [35,36]. Recently, the PFLANN model has been developed for intelligent water fountain sound pleasantness monitoring [37]. The whole analysis is divided into three frequency regions: low, medium, and high, and each region’s output is associated with its specific weight. This modification in the FLANN model is done considering the response of human speech perception to different frequency regions. The trigonometric functional expansion is used for its impressive results in ref [37]. The human auditory system’s frequency response has been found to be nonlinear in speech processing. The frequency unit Mel, which is based on perception, serves as a representation of this nonlinearity. The audio signal is transformed from the linear frequency to the perceptual domain by employing the Mel filter bank [37]. The relationship between the linear scale (f) and Mel scale (f_m) can be expressed as:

$$f_m = 2595 \cdot \log_{10} \left(1 + \frac{f}{700} \right) \tag{1}$$

The low-frequency region (0 Hz–300 Hz), mid-frequency region (300 Hz –5 kHz), and high-frequency region (5 kHz–20 kHz) have the importance, or the weighted parameters are found to be 0.33, 0.58, and 0.08, respectively. These values are calculated by using the number of filters in each region and the average peak amplitude of filters. The weights of the PFLANN model have been updated using the particle swarm intelligence technique as mentioned in ref. [37,38].

C. CASCADED PFLANN MODEL

Instead, using a single hidden layer in the case of PFLANN model, the overall accuracy in the detection of respiratory diseases can be greatly enhanced by cascading PFLANN [39,40]. This approach has been considered in this paper with PSO as the weight updation scheme, and the details are mentioned in Fig. 1.

The traditional PFLANN model has one hidden layer with weights w , input, and outputs x and y . The relevant equations of the FLANN are listed in Equation (2):

$$\begin{aligned} \varphi(k) &= [X(k)\sin(\pi X(k)) \dots \cos((2J - 1)\pi X(k))]^T \\ W(k) &= [W_{ij}(k)], 1 \leq i \leq I, 1 \leq j \leq J \\ Y(k) &= W(k)\varphi(k) \end{aligned} \tag{2}$$

The input features are divided into three categories such as: low, medium, and high-frequency ranges. The output y has been further expanded by using another triangular expansion. Finally, the predicted output is compared with the actual output and the error is calculated. For decreasing the error, the weights are updated using the bio-inspired computing. In this case, PSO has been used for the weight updation. For the activation function, tanh has been used in the current implementation after the weighted summation at output of both the layers. The mean square error (MSE), which is derived from the error terms, is computed; once the MSE reaches the minimum value, the training procedure is terminated.

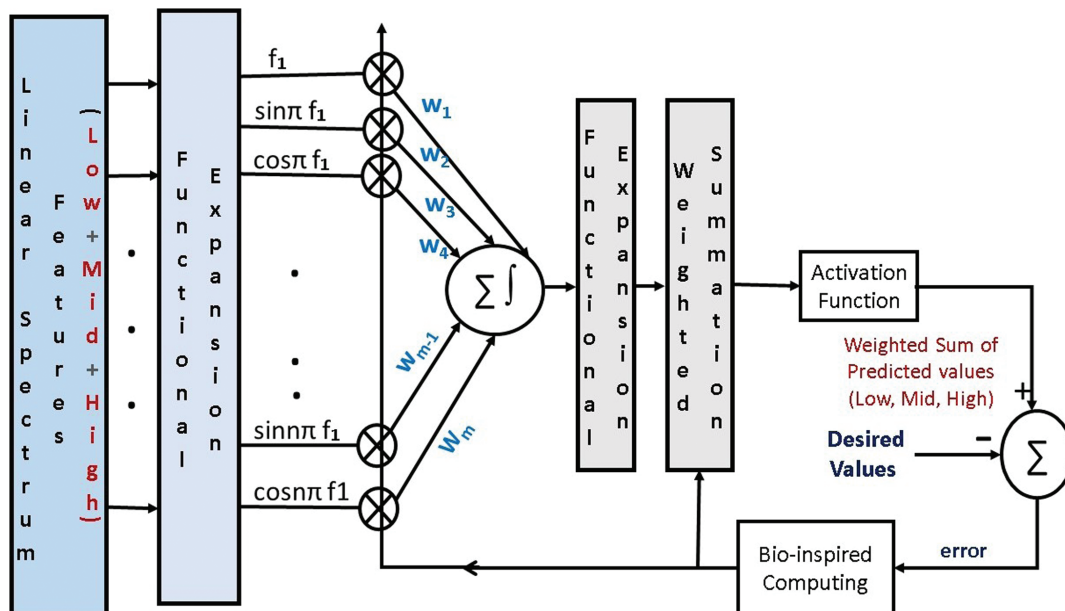


Fig. 1. Block diagram of implementation of cascaded PFLANN model.

III. SIMULATION RESULTS AND DISCUSSIONS

In this section, the details of the experimental design and performance evaluation measures are presented.

Step 1: Dataset preparation — In the first step, the eight respiratory datasets have been downloaded and divided mainly into two categories such as: healthy and diseased except dataset-7. For the dataset-7, three classes have been considered, including drug actuation, inhalation, and exhalation.

Step 2: Feature extraction frame-wise — In this step, framing and windowing is applied to the input speech samples. Out of different available windows, the Hamming window is used because of low spectral leakage, and the 25 ms window size is used with 50% overlapping. Then, at the frame level, one-sided linear spectrum features are extracted and divided into three categories considering frequency ranges such as low (0 Hz–300 Hz), mid

(300 Hz –5 kHz), and high (5 kHz–20 kHz). At the sample level, the two statistical features such as mean and variance are extracted.

Step 3: Classification using cascaded PFLANN and baseline classifiers — The speech sample level feature vectors extracted in step 2 are given as input to the proposed cascaded PFLANN. The classifier is then trained and evaluated using two cross-validation schemes such as leave-one-out and k-fold. The proposed model is shown in Fig. 2. The results of the proposed model are compared with the standard classifiers used in the original datasets.

Step 4: Performance evaluation and model validation — A numeric metric called confusion matrix is widely used to assess the performance of classifiers.

The confusion matrix is a classification metric that is calculated from four values: T+ (True Positive), F+ (False Positive), F– (False Negative), and T– (True Negative). In the present model, T+ is the number of diseased patients predicted accurately as diseased, T– is the number of healthy patients predicted accurately

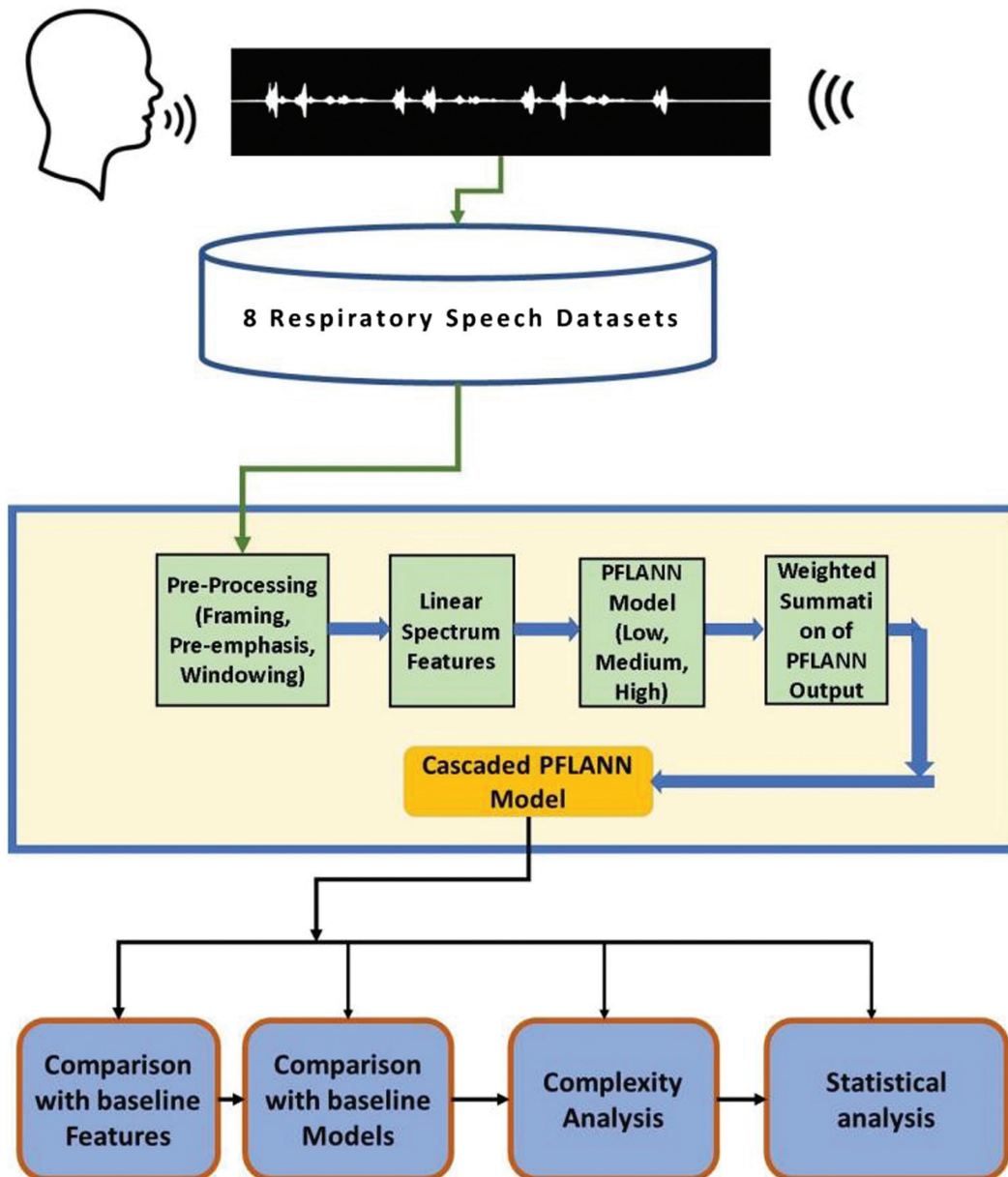


Fig. 2. Block diagram of the proposed implementation model.

as healthy, $F+$ is the number of healthy patients predicted as diseased, and $F-$ is the number of diseased patients predicted as healthy. The $F-$ is a crucial parameter in medical diagnosis since it identifies the patients who have respiratory disease but were incorrectly classified as healthy by the model. However, the $F+$ is less important because the patient will go back for a second round of tests to confirm. Diseased patients, on the other hand, should not be misinterpreted as negative. Standard performance measures as reported in [41,42] such as classification accuracy (ACC), the area under the curve (AUC), precision (P), recall (R), and F-2 score (F-2) are used:

$$Acc = \frac{(T+) + (T-)}{(T+) + (T-) + (F+) + (F-)} \quad P = \frac{(T+)}{(T+) + (F+)} \quad (3)$$

$$R = \frac{(T+)}{(T+) + (F-)} \quad F2 = \frac{5 \times P \times R}{(4 \times P) + R}$$

A. COMPARISON USING LEAVE-ONE-OUT CROSS-VALIDATION SCHEME

The performance of the proposed cascaded PFLANN model is compared using leave-one-out cross-validation (LOOCV) with several standard machine learning based models such as: classifier-1 (support vector machine) [29], classifier-2 (random forest) [30,32], and classifier-3 (K-nearest neighbor) [43] used for the speech classification task using same set of linear spectrum feature vectors. LOOCV provides almost unbiased estimator of statistical models' generalization properties and makes it a useful criterion for model selection and comparison. In this technique, from n samples of a dataset, $n-1$ samples are used to train the model, and 1 sample is used for testing. This process is repeated n times till all the samples are tested once [44]. The overall performance of the model is dependent on the size of the dataset. When the data size is small and a reliable estimation of model output is more critical, then LOOCV is performed. As in the present paper, the accuracy of testing is more important due to the research in medical field and the size of the datasets are comparatively small, the LOOCV model is used for cross-validation [45]. To select the best parameters of the classifiers, grid search is used. The simulation results are listed in Table I. It has been observed that the proposed cascaded PFLANN classifier shows improved performance with an average accuracy of 0.94, and the classifier-2 is at the second position followed by classifier-3 and 1. Similarly, for the evaluation of F-2 score and AUC, the rankings in the descending order are cascaded PFLANN, classifier-2, 3, and 1. In the dataset-2, all the classifiers are working better than other datasets. The simulation results demonstrate the superiority of cascaded PFLANN classifier. From the dataset analysis, it has been observed that the proposed classification model is working best at dataset-2 and dataset-8, while the accuracy is comparatively less for the dataset-7. For dataset-7, the classification is mainly between inhalation and exhalation rather than healthy and diseased.

B. COMPARISON WITH BASELINE FEATURES

After investigating the performance of the proposed cascaded PFLANN model with baseline models, the effectiveness of the linear spectrum feature extraction technique is analyzed in this section. Here, different baseline features are used along with the cascaded PFLANN model. The baseline features used are

Table I. Evaluation results using LOOCV for baseline models

Category	Evaluation measures	EnFLANN	Model-1	Model-2	Model-3
Dataset-1	ACC	0.94	0.85	0.91	0.88
	F-2	0.94	0.85	0.91	0.88
	P	0.93	0.85	0.91	0.88
	R	0.93	0.85	0.91	0.88
	AUC	0.93	0.84	0.92	0.88
Dataset-2	ACC	0.95	0.84	0.93	0.87
	F-2	0.95	0.84	0.93	0.87
	P	0.95	0.84	0.93	0.87
	R	0.94	0.84	0.93	0.87
	AUC	0.95	0.83	0.94	0.85
Dataset-3	ACC	0.93	0.85	0.91	0.89
	F-2	0.93	0.85	0.91	0.89
	P	0.93	0.85	0.91	0.89
	R	0.93	0.85	0.91	0.89
	AUC	0.92	0.84	0.91	0.89
Dataset-4	ACC	0.94	0.83	0.92	0.87
	F-2	0.94	0.83	0.92	0.87
	P	0.94	0.83	0.92	0.87
	R	0.94	0.83	0.92	0.87
	AUC	0.94	0.84	0.91	0.85
Dataset-5	ACC	0.93	0.82	0.91	0.85
	F-2	0.93	0.83	0.91	0.85
	P	0.93	0.82	0.91	0.85
	R	0.93	0.82	0.91	0.85
	AUC	0.94	0.82	0.91	0.85
Dataset-6	ACC	0.94	0.84	0.92	0.87
	F-2	0.94	0.84	0.92	0.87
	P	0.94	0.84	0.92	0.87
	R	0.94	0.84	0.92	0.87
	AUC	0.94	0.84	0.92	0.85
Dataset-7	ACC	0.91	0.83	0.90	0.84
	F-2	0.92	0.84	0.90	0.85
	P	0.91	0.83	0.91	0.84
	R	0.91	0.83	0.91	0.84
	AUC	0.91	0.83	0.90	0.84
Dataset-8	ACC	0.95	0.87	0.93	0.88
	F-2	0.94	0.87	0.93	0.88
	P	0.94	0.87	0.93	0.88
	R	0.94	0.87	0.93	0.88
	AUC	0.94	0.88	0.94	0.89

The best performing model and the evaluation measures are highlighted in bold.

Featureset-1 (MFCC), Featureset-2 (DFT), and Featureset-3 (GTCC). For the three feature sets, the mean values are taken at the sample level from the frame level values. But in the case of linear spectrum features, the statistical values of mean and variance are being considered. The simulation results are listed in Table II. It can be observed that the detection accuracy is comparatively higher

Table II. Comparison with baseline features

Dataset	Evaluation measures	Proposed features	Feature set-1	Feature set-2	Feature set-3
Dataset-1	ACC	0.94	0.92	0.93	0.90
	F-2	0.94	0.92	0.93	0.89
	P	0.93	0.92	0.93	0.89
	R	0.93	0.92	0.93	0.89
	AUC	0.93	0.92	0.93	0.89
Dataset-2	ACC	0.95	0.91	0.94	0.91
	F-2	0.95	0.91	0.94	0.91
	P	0.95	0.92	0.94	0.91
	R	0.94	0.90	0.94	0.91
	AUC	0.95	0.92	0.94	0.91
Dataset-3	ACC	0.93	0.93	0.92	0.92
	F-2	0.93	0.93	0.92	0.92
	P	0.93	0.93	0.92	0.92
	R	0.93	0.93	0.92	0.92
	AUC	0.92	0.93	0.93	0.92
Dataset-4	ACC	0.94	0.91	0.93	0.92
	F-2	0.94	0.91	0.93	0.92
	P	0.94	0.92	0.93	0.92
	R	0.94	0.92	0.93	0.92
	AUC	0.94	0.90	0.93	0.94
Dataset-5	ACC	0.93	0.90	0.92	0.92
	F-2	0.93	0.90	0.92	0.92
	P	0.93	0.89	0.92	0.92
	R	0.93	0.89	0.92	0.92
	AUC	0.94	0.89	0.93	0.94
Dataset-6	ACC	0.94	0.91	0.92	0.92
	F-2	0.94	0.91	0.92	0.92
	P	0.94	0.91	0.92	0.92
	R	0.94	0.91	0.92	0.92
	AUC	0.94	0.91	0.93	0.92
Dataset-7	ACC	0.91	0.89	0.90	0.89
	F-2	0.92	0.89	0.90	0.89
	P	0.91	0.89	0.90	0.89
	R	0.91	0.89	0.90	0.89
	AUC	0.91	0.89	0.91	0.89
Dataset-8	ACC	0.95	0.91	0.94	0.92
	F-2	0.94	0.91	0.94	0.92
	P	0.94	0.91	0.94	0.92
	R	0.94	0.91	0.94	0.92
	AUC	0.94	0.91	0.93	0.92

on a consistent basis for the linear spectrum features. However, the MFCC and GTCC features are also performing well in the second position in most of the cases. The GTCC features where only mean are taken are at the last rank. Out of the eight datasets, for dataset-3, MFCC and the proposed features are in the same rank. The improved performance of the linear spectrum statistical features

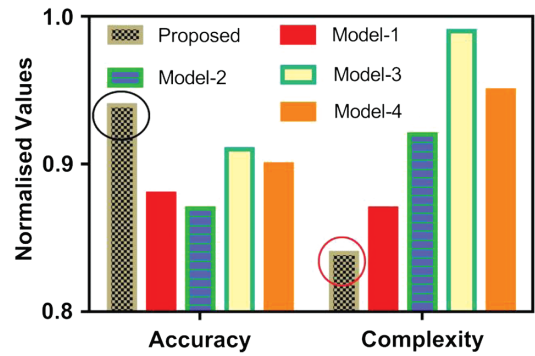


Fig. 3. Computational complexity analysis.

shows the superiority of the proposed model compared to the baseline models and features.

C. COMPARATIVE ANALYSIS ON COMPUTATIONAL COMPLEXITY

In IoT networks, computational complexity plays a crucial role. To compute the computational complexity, computational speed (CS) is used, and it can be calculated as mentioned below:

$$CS = \frac{n}{T} \tag{4}$$

where n is the time duration (length) of the speech signal and T is the execution time of the model for that speech signal [46]. The models for comparing the computational complexity are model-1 (MFCC+SVM) [47], model-2 (GTCC+KNN) [48], model-3 (MFCC+RF) [49], and model-4 (ERB+NB) [50], and the results are plotted in Fig. 3. The proposed model has been compared with standard models from the literature which are being used for speech recognition [47–50]. It can be observed that the proposed model provides the highest detection accuracy as well as lowest computational complexity. It is showing superior performance as compared to others. In terms of higher accuracy, the descending rank of models are model-3, 4, 1, and 2 and in terms of lower computational complexity, the ranking of models are model-1, 2, 4, and 3.

D. STATISTICAL ANALYSIS OF CLASSIFIER MODELS

For statistical analysis comparison, t-statistic analysis is used. It is computed between two classifiers as:

Table III. T-statistic analysis of proposed model with standard ML-based models

Dataset	CPFLANN vs SVM	CPFLANN vs RF	CPFLANN vs KNN
1	4.4	1.3	3.5
2	2.1	4.6	1.7
3	3.2	2.3	6.1
4	2.3	0.5	2.3
5	3.4	2.2	6.1
6	5.2	2.6	3.4
7	4.3	2.8	2.6
8	3.7	1.6	3.4

$$t = \frac{m_1 - m_2}{\sqrt{v_1^2 + v_2^2}} \quad (5)$$

where m_1, m_2 and v_1^2, v_2^2 are the mean and variance of the fivefold classification accuracy of classifier 1 and 2, respectively [50]. The proposed cascaded PFLANN is being denoted as CPFLANN. Most of the t-statistic values listed in Table III are positive. It justifies the improved performance of the proposed model over baseline models.

The results of the aforementioned classification tasks with baseline models and features and statistical analyses show the superiority of the proposed CPFLANN model. The primary cause of this is using the perceptual model for capturing the nonlinearity of frequency response and application of cascaded structure with bio-inspired computing for weight updation.

IV. CONCLUSION

In the current study, a low-complexity-based health monitoring system is developed using cascaded PFLANN model with bio-inspired computing. It is mainly used for the respiratory disease detection from recorded speech signal. The major contribution of the study is the development of cascaded PFLANN model which has not been used till now by any researcher to the best of knowledge. It provides a new dimension of designing a neural network and extension to the basic PFLANN model. Another contribution is the use of bio-inspired techniques for the weight updation. The analysis of accuracy and computational complexity justifies the use of the proposed model, and it has been verified in eight standard datasets. However, it is advised that the medical practitioner can confirm the detection method in hospitals by using the proposed intelligent model before coming to the final prescription. Other speech recognition tasks can also be evaluated and implemented further using the proposed model.

ACKNOWLEDGMENTS

The voice database of COVID-19 was shared by Professor Cecilia Mascolo of the Department of Computer Science and Technology and the University of Cambridge, who is also the Chancellor, Master, and Scholar [30].

CONFLICT OF INTEREST STATEMENT

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

REFERENCES

- [1] R. Zhu, L. Liu, M. Ma, and H. Li, "Cognitive-inspired computing: advances and novel applications," *Future Gener. Comput. Syst.*, vol. 109, pp. 706–709, 2020.
- [2] J. Cui, M. Wang, Y. Luo, and H. Zhong, "DDoS detection and defense mechanism based on cognitive-inspired computing in SDN," *Future Gener. Comput. Syst.*, vol. 97, pp. 275–283, 2019.
- [3] D. K. Jain, P. Boyapati, J. Venkatesh, and M. Prakash, "An intelligent cognitive-inspired computing with big data analytics framework for sentiment analysis and classification," *Inf. Process. Manage.*, vol. 59, no. 1, p. 102758, 2022.
- [4] U. R. Kamboh, U. Ullah, S. Khalid, U. Raza, C. Chakraborty, and F. Al-Turjman, "Path loss modelling at 60 GHz mmWave based on cognitive 3D ray tracing algorithm in 5G," *Peer-to-Peer Network Appl.*, vol. 14, pp. 3181–3197, 2021.
- [5] J. Amin, M. Sharif, N. Gul, S. Kadry, and C. Chakraborty, "Quantum machine learning architecture for COVID-19 classification based on synthetic data generation using conditional adversarial neural network," *Cognit. Comput.*, vol. 14, no. 5, pp. 1677–1688, 2022.
- [6] S. Selvaraj and S. Sundaravaradhan, "Challenges and opportunities in IoT healthcare systems: a systematic review," *SN Appl. Sci.*, vol. 2, no. 1, p. 139, 2020.
- [7] A. Ukil, S. Bandyopadhyay, C. Puri, and A. Pal, "IoT healthcare analytics: the importance of anomaly detection," in *2016 IEEE 30th Int. Conf. Adv. Inf. Netw. Appl. (AINA)*, 2016, pp. 994–997.
- [8] M. N. Bhuiyan, M. M. Rahman, M. M. Billah, and D. Saha, "Internet of things (IoT): a review of its enabling technologies in healthcare applications, standards protocols, security, and market opportunities," *IEEE Internet Things J.*, vol. 8, no. 13, pp. 10 474–10 498, 2021.
- [9] P. Kakria, N. K. Tripathi, and P. Kitipawang, "A real-time health monitoring system for remote cardiac patients using smartphone and wearable sensors," *Int. J. Telemed. Appl.*, vol. 2015, pp. 8–8, 2015.
- [10] J. Wan, M. A. Al-awlaqi, M. Li, M. O'Grady, X. Gu, J. Wang, and N. Cao, "Wearable IoT enabled real-time health monitoring system," *EURASIP J. Wireless Commun. Netw.*, vol. 2018, no. 1, pp. 1–10, 2018.
- [11] I. Hussain and S. J. Park, "HealthSOS: real-time health monitoring system for stroke prognostics," *IEEE Access*, vol. 8, pp. 213 574–213 586, 2020.
- [12] P. Mayorga, C. Druzgalski, R. L. Morelos, O. H. Gonzalez, and J. Vidales, "Acoustics based assessment of respiratory diseases using GMM classification," in *2010 Annu. In. Conf. IEEE Eng. Med. Biol.*, 2010, pp. 6312–6316.
- [13] L. Brunese, F. Mercaldo, A. Reginelli, and A. Santone, "A neural network-based method for respiratory sound analysis and lung disease detection," *Appl. Sci.*, vol. 12, no. 8, p. 3877, 2022.
- [14] L. Pham, H. Phan, R. Palaniappan, A. Mertins, and I. McLoughlin, "CNN-MoE based framework for classification of respiratory anomalies and lung disease detection," *IEEE J. Biomed. Health. Inf.*, vol. 25, no. 8, pp. 2938–2947, 2021.
- [15] G. B. Kim, K.-H. Jung, Y. Lee, H.-J. Kim, N. Kim, S. Jun, J. B. Seo, and D. A. Lynch, "Comparison of shallow and deep learning methods on classifying the regional pattern of diffuse lung disease," *J. Digital Imaging*, vol. 31, pp. 415–424, 2018.
- [16] A. Sengupta and R. Chaurasia, "Secured convolutional layer IP core in convolutional neural network using facial biometric," *IEEE Trans. Consum. Electron.*, vol. 68, no. 3, pp. 291–306, 2022.
- [17] K. L. Khatri and L. S. Tamil, "Early detection of peak demand days of chronic respiratory diseases emergency department visits using artificial neural networks," *IEEE J. Biomed. Health. Inf.*, vol. 22, no. 1, pp. 285–290, 2017.
- [18] G. Chambres, P. Hanna, and M. Desainte-Catherine, "Automatic detection of patient with respiratory diseases using lung sound analysis," *IEEE 2018 Int. Conf. Content-Based Multimed. Indexing (CBMI)*, 2018, pp. 1–6.
- [19] E. B. Mazomenos, D. Biswas, A. Acharyya, T. Chen, K. Maharatna, J. Rosengarten, J. Morgan, and N. Curzen, "A low-complexity ECG feature extraction algorithm for mobile healthcare applications," *IEEE J. Biomed. Health Inf.*, vol. 17, no. 2, pp. 459–469, 2013.
- [20] I. D. Dinov, "Methodological challenges and analytic opportunities for modeling and interpreting Big Healthcare Data," *GigaScience*, vol. 5, no. 1, pp. s13 742–016, 2016.

- [21] C. S. Nandyala and H.-K. Kim, "From cloud to fog and IoT- based real-time U-healthcare monitoring for smart homes and hospitals," *Int. J. Smart Home*, vol. 10, no. 2, pp. 187–196, 2016. P. Maji and R. Mullins, "On the reduction of computational complexity of deep convolutional neural networks," *Entropy*, vol. 20, no. 4, p. 305, 2018.
- [22] R. Livni, S. Shalev-Shwartz, and O. Shamir, "On the computational efficiency of training neural networks," *Adv. N.A. Inf. Process. Syst.*, vol. 27, pp. 1–9, 2014.
- [23] B. G. Celler and P. de Chazal, "Low computational cost classifiers for ECG diagnosis using neural networks," in *Proc. 20th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. Vol. 20 Biomed. Eng. Towards Year 2000 Beyond (Cat. No. 98CH36286)*, vol. 3, 1998, pp. 1337–1340.
- [24] J. Liang, J. Xu, H. Shen, and L. Fang, "Land-use classification via constrained extreme learning classifier based on cascaded deep convolutional neural networks," *Eur. J. Remote Sens.*, vol. 53, no. 1, pp. 219–232, 2020.
- [25] W. Ding, J. Xie, D. Dai, H. Zhang, H. Xie, and W. Zhang, "CNNcon: improved protein contact maps prediction using cascaded neural networks," *PLoS One*, vol. 8, no. 4, p. e61533, 2013.
- [26] X. Li, Z. Yang, and H. Wu, "Face detection based on receptive field enhanced multi-task cascaded convolutional neural networks," *IEEE Access*, vol. 8, pp. 174 922–174 930, 2020.
- [27] S. Li, X. Xing, W. Fan, B. Cai, P. Fordson, and X. Xu, "Spatiotemporal and frequential cascaded attention networks for speech emotion recognition," *Neurocomputing*, vol. 448, pp. 238–248, 2021.
- [28] N. Sharma, P. Krishnan, R. Kumar, S. Ramoji, S. R. Chetupalli, N. R. P. K. Ghosh, and S. Ganapathy, "Coswara — a database of breathing, cough, and voice sounds for COVID-19 diagnosis," in *Proc. Interspeech*, 2020, pp. 4811–4815.
- [29] C. Brown, J. Chauhan, A. Grammenos, J. Han, A. Hasthanasombat, A. Spathis, T. Xia, P. Cicuta, and C. Mascolo, "Exploring automatic diagnosis of COVID-19 from crowdsourced respiratory sound data," in *Proc. ACM SIGKDD Int. Conf. Knowl. Disc. Data Mining*, pp. 3474–3484, August 2020.
- [30] G. Chaudhari, X. Jiang, A. Fakhry, A. Han, J. Xiao, S. Shen, and A. Khanzada, "Virufy: global applicability of crowdsourced and clinical datasets for AI detection of COVID-19 from cough," *arXiv preprint arXiv:2011.13320*, 2020.
- [31] K. V. S. Ritwik, S. B. Kalluri, and D. Vijayasanen, "COVID-19 patient detection from telephone quality speech data," *arXiv preprint arXiv:2011.04299*, 2020.
- [32] L. Orlandic, T. Teijeiro, and D. Atienza, "The COUGHVID crowdsourcing dataset: a corpus for the study of large-scale cough analysis algorithms," *Sci. Data, Nat. Publishing Group*, vol. 8, no. 1, pp. 1–10, 2021. [Online]. DOI: [10.1038/s41597-021-00937-4](https://doi.org/10.1038/s41597-021-00937-4).
- [33] S. Nousias, G. Arvanitis, A. Anastasiou, and K. Moustakas, "Respiratory and drug actuation dataset," IEEE Dataport, 2022. [Online]. Available: <https://iee-dataport.org/documents/respiratory-anddrug-actuation-dataset>, doi: 10.21227/g84h-ma24.
- [34] T. K. Dash and S. S. Solanki, "Investigation on the effect of the input features in the noise level classification of noisy speech," *J. Sci. Ind. Res.*, vol. 78, no. 12, pp. 868–872, 2019. [Online]. Available at <http://nopr.niscares.in/handle/123456789/52213>.
- [35] T. K. Dash, S. S. Solanki, and G. Panda, "Multi-objective approach to speech enhancement using tunable Q-factor-based wavelet transform and ANN techniques," *Circuits Syst. Signal Process.*, vol. 40, no. 12, pp. 1–31, 2021.
- [36] B. Brahma, T. K. Dash, G. Panda, L. N. Prasad, and R. Kulkarni, "Design of P-FLANN model for intelligent water fountain sound pleasantness monitoring using bio-inspired computing and human speech perception," *J. Artif. Intell. Technol.*, vol. 40, no. 12, 2023.
- [37] J. Kennedy and R. Eberhart, "Particle swarm optimization," *Proc. ICNN'95-Int. Conf. Neural Netw.*, vol. 4, pp. 1942–1948, 1995.
- [38] M. Rout, A. K. Jena, J. K. Rout, and H. Das, "Teaching-learning optimization based cascaded low-complexity neural network model for exchange rates forecasting," in *Smart Intell. Comput. Appl.: Proc. Third Int. Conf. Smart Comput. Inf., Vol. 1*, 2020, pp. 635–645.
- [39] N. V. George and G. Panda, "Active control of nonlinear noise processes using cascaded adaptive nonlinear filter," *Appl. Acoust.*, vol. 74, no. 1, pp. 217–222, 2013.
- [40] J. Lever, M. Krzywinski, and N. Altman, "Points of significance: model selection and overfitting," *Nat. Methods*, vol. 13, no. 9, pp. 703–705, 2016.
- [41] T. K. Dash, C. Chakraborty, S. Mahapatra, and G. Panda, "Mitigating information interruptions by COVID-19 face masks: a three-stage speech enhancement scheme," *IEEE Trans. Comput. Soc. Syst.*, pp. 1–10, 2022.
- [42] T.-L. Pao, Y.-T. Chen, J.-H. Yeh, Y.-M. Cheng, and Y.-Y. Lin, "A comparative study of different weighting schemes on KNN-based emotion recognition in Mandarin speech," in *Int. Conf. Intell. Comput.*, 2007, pp. 997–1005. DOI: [10.1007/978-3-540-74171-8_101](https://doi.org/10.1007/978-3-540-74171-8_101)
- [43] G. C. Cawley and N. L. Talbot, "Fast exact leave-one-out cross-validation of sparse least-squares support vector machines," *Neural Netw.*, vol. 17, no. 10, pp. 1467–1475, 2004.
- [44] A. Tahmassebi, A. H. Gandomi, I. McCann, M. H. Schulte, A. E. Goudriaan, and A. Meyer-Baese, "Deep learning in medical imaging: FMRI big data analysis via convolutional neural networks," in *ACM Proceedings of the practice and experience on advanced research computing*, pp. 1–4, 2018.
- [45] Y.-H. Wang, C.-H. Yeh, H.-W. Vincent Young, K. Hu, and M.-T. Lo, "On the computational complexity of the empirical mode decomposition algorithm," *Physica A*, vol. 400, pp. 159–167, 2014.
- [46] A. Milton, S. S. Roy, and S. T. Selvi, "SVM scheme for speech emotion recognition using MFCC feature," *Int. J. Comput. Appl.*, vol. 69, no. 9, 2013.
- [47] C. Zhu and W. Ahmad, "Emotion recognition from speech to improve human-robot interaction," in *2019 IEEE Intl Conf Dependable, Autonom. Secure Comput., Intl Conf Pervasive Intell. Comput., Intl Conf Cloud Big Data Comput., Intl Conf Cyber Sci. Technol. Congr. (DASC/PiCom/CBDCCom/CyberSciTech)*, 2019, pp. 370–375.
- [48] F. Noroozi, T. Sapiński, D. Kamińska, and G. Anbarjafari, "Vocal-based emotion recognition using random forests and decision tree," *Int. J. Speech Technol.*, vol. 20, no. 2, pp. 239–246, 2017.
- [49] Z. Peng, Z. Zhu, M. Unoki, J. Dang, and M. Akagi, "Speech emotion recognition using multichannel parallel convolutional recurrent neural networks based on gammatone auditory filterbank," in *2017 IEEE Asia-Pacific Signal Inf. Process. Assoc. Annu. Summit Conf. (APSIPA ASC)*, 2017, pp. 1750–1755.
- [50] T.-T. Wong, "Parametric methods for comparing the performance of two classification algorithms evaluated by k-fold cross validation on multiple data sets," *Pattern Recognit.*, vol. 65, pp. 97–107, 2017.