Utilizing an enhanced deep convolutional neural network for disease detection based on human cough sounds.

Channakrishnaraju[1]✉, M Tamilarasan[2]

Professor, Dept of CSE, Sri Siddhartha Institute of Technology, Tumkur, Karnataka, rajuck@ssit.edu.in

2nd year M.TECH, Dept of CSE, Sri Siddhartha Institute of Technology, Tumkur, Karnataka, tamilarasanm.m.tech@gmail.com

Abstract

The healthcare industry is only one of several that have been impacted by the recent explosion in the use of artificial intelligence (AI) methods, especially deep learning algorithms. In this study, we provide a new method for illness identification that makes use of deep convolutional neural networks (CNNs) that have been upgraded using human acoustic inputs. A non-invasive, efficient, and cost-effective way for early illness detection is the goal of the suggested system. The technology takes use of the unique properties of human voice transmissions, which provide important information about a person's physiological condition. Potential indications of different illnesses can be identified by studying these signals for small changes and trends. We do this by using a deep convolutional neural network (CNN) architecture that has been fine-tuned for the processing of audio input. For better pattern recognition in the input audio signals, our improved CNN model makes use of state-of-the-art methods including hierarchical learning, dimensionality reduction, and feature extraction. Furthermore, we use data augmentation methodologies to strengthen the model's capacity to generalize and handle varied patient groups and diseases more effectively. Denoising techniques using Wavelet transform denoising, segmentation via K-mean clustering, feature extraction employing Recursive Feature Elimination (RFE), and classification utilizing Enhanced Deep Convolutional Neural Networks (EDCNN) are integrated into our comprehensive approach, enhancing the system's accuracy and reliability in detecting diseases based on human sound signatures.

Keywords: *Artificial Intelligence, Cost-Effective, Disease Detection, Early Detection, Enhanced Deep Convolutional Neural Network, Recursive Feature Elimination.*

1. INTRODUCTION

Innovations in medical technology continue to revolutionize healthcare, with recent advancements focusing on non-invasive and efficient methods of disease detection. One such innovative approach is the Human Sound-Based Disease Detection System, which harnesses the power of sound analysis to identify potential health issues in individuals [1-3]. The system operates by utilizing specialized microphones or sensors capable of capturing various sounds produced by the human body [4-5]. These sounds include but are not limited to heartbeat patterns, respiratory sounds, gastrointestinal noises, and vocal characteristics. Each of these sounds carries valuable information about the physiological state of an individual [6-7]. The process begins with the collection of sound data from the individual using wearable devices or dedicated recording equipment [8]. The captured sound signals are then processed through sophisticated algorithms designed to analyze and interpret the acoustic patterns. These algorithms leverage machine learning and artificial intelligence techniques to recognize specific sound patterns associated with different diseases or health conditions [9].

Healthcare is one of several fields that has benefited greatly from the revolutionary age brought about by the fast adoption of artificial intelligence (AI), especially with the help of sophisticated deep learning algorithms [10]. Using improved deep convolutional neural networks (CNNs) in tandem with human speech inputs, this research presents a novel method for illness identification [11]. An efficient, non-invasive, and cost-effective way to diagnose diseases early on is the main goal of this suggested system, which uses the abundance of information contained in human sound waves [12]. One novel way to investigate non-traditional yet useful biomarkers is via human sound waves, which intrinsically provide important clues regarding an individual's physiological condition [13]. The possibility of revealing markers of different illnesses arises from the careful examination of minute differences and patterns within these signals. We use a tailored deep CNN architecture that has been fine-tuned for processing audio data in order to carry out this complex operation [14]. To guarantee accurate capture of subtle patterns in the input audio signals, our improved CNN model uses advanced approaches including hierarchical learning, dimensionality reduction, and feature extraction [15].

1.1 Motivation of the paper

This research uses an improved deep convolutional neural network (CNN) to sift through a large dataset of human speech patterns in search of new ways to diagnose diseases early on. Training and validation are both made possible

by the dataset's inclusion of a wide range of physiological and pathological situations. For hierarchical feature extraction, the custom-designed CNN architecture uses preprocessing steps like noise reduction and segmentation on raw acoustic data. The architecture also includes numerous convolutional and pooling layers. The complex patterns in the input audio signals are captured by combining advanced approaches like feature extraction, dimensionality reduction using methods like principal component analysis (PCA), and hierarchical learning.

2. BACKGROUND STUDY

De Souza, L. et al. (4) the new corona virus pandemic in 2020 brought a lot of attention to the relatively young area of study into disease detection using forced cough noises. In this paper, we provide a Deep Learning topology for COVID-19 health classification along with a training prototype that incorporates an Early Stopping (ES) strategy. Evaluating the suggestions was done using data from the Cascara and COUGHVID databases. The databases are comprised of a single dataset that includes both clinical information and recordings of forced coughs. It was suggested to use a data augmentation strategy to the training set in order to address the significant data imbalance, with 81.76% belonging to the negative class.

Hamidi, et al. (5) based on speaker independent ASR technology; we have reported the diagnosis of COVID-19 patients and healthy persons in this study. We have evaluated the cough train system's performance in three HMM states and several Gaussian mixture distributions via our trials. Findings demonstrated that a 5% to 10% increase in the identification rate of COVID-19 infected speakers compared to non-infected speakers was achieved after training the system with cough sounds of infected individuals. Furthermore, the percentage of COVID-19 infected individuals was 5% to 10% lower than the rate of non-infected speakers when we trained the system using cough sounds of healthy individuals. Classification sensitivity of 90% and specificity of 10% were attained, respectively, in the best case scenario.

Kaulamo, J.T. et al. (6) The majority of individuals without chronic cough had never sought medical attention for it, whereas a small percentage had gone to the doctor many times in the previous year for reasons linked to coughing. Those who used healthcare services often did not have the most severe coughs. People with chronic phlegm respiratory diseases and impaired quality of life were characterized by frequent visits to the doctor because of their cough.

Naqvi et al. (8) the dynamic physiological process of coughing helps clean the airways, but it can also indicate pulmonary illness. Infectious illnesses are

in a special position to take advantage of the mechanical power that coughing occurrences provide in order to spread from one person to another. There are still many unsolved concerns in the subject, despite the fact that experiments are being conducted to determine the exact nature of the interactions between infectious pathogens and the neurons in the brain that cause coughing.

Pahar, M., et al. (9) Our bed occupancy detection system is based on machine learning and utilizes the accelerometer signal recorded by a consumer smart phone that is fastened to the patient's bed. Because knowing when the patient is in bed is crucial for reliably calculating cough rate, bed occupancy recognition is necessary for the implementation of automated long-term cough monitoring utilizing the same accelerometer signal. Compared to wear able or video monitoring, a sensor put on the bed is more practical and less invasive, and the use of acceleration measures alone inherently protects privacy.

Vhaduri, S., et al. (11) shows that when comparing three modeling approaches—unguided, semi-guided, and guided—using ten random splits, the user can anticipate a 12%-28% increase in accuracy and F1 score when identifying cough and non-cough sounds using guided models. Moreover, we discover that semi-guided models perform better than their unguided counterparts. Although this study demonstrates the approach's practicality, more research is needed to validate the models in a clinical setting before the work can be commercialized.

2.1 Problem definition

The early diagnosis of illnesses is still a big difficulty, even with all the great healthcare developments. Even in settings with plenty of resources, traditional diagnostic approaches can be inconvenient because they need invasive procedures, expensive testing, or a lot of time. The need for an efficient, low-cost, and non-invasive means of early illness detection is discussed in this article. In light of the revolutionary possibilities of AI methods, especially deep learning algorithms, the suggested approach centers on tapping into the wealth of physiological data contained in human speech signals.

3. MATERIALS AND METHODS

A thorough and individualized approach is used to achieve the objective of efficient and non-invasive illness diagnosis via the use of human sound waves. With an emphasis on cutting-edge deep learning techniques, this section details the resources and procedures that went into creating and refining the suggested system. Methods including feature extraction, dimensionality reduction, hierarchical learning, and data augmentation are a part of the approach,

which also includes building a varied dataset and designing a unique deep convolutional neural network (CNN) architecture.

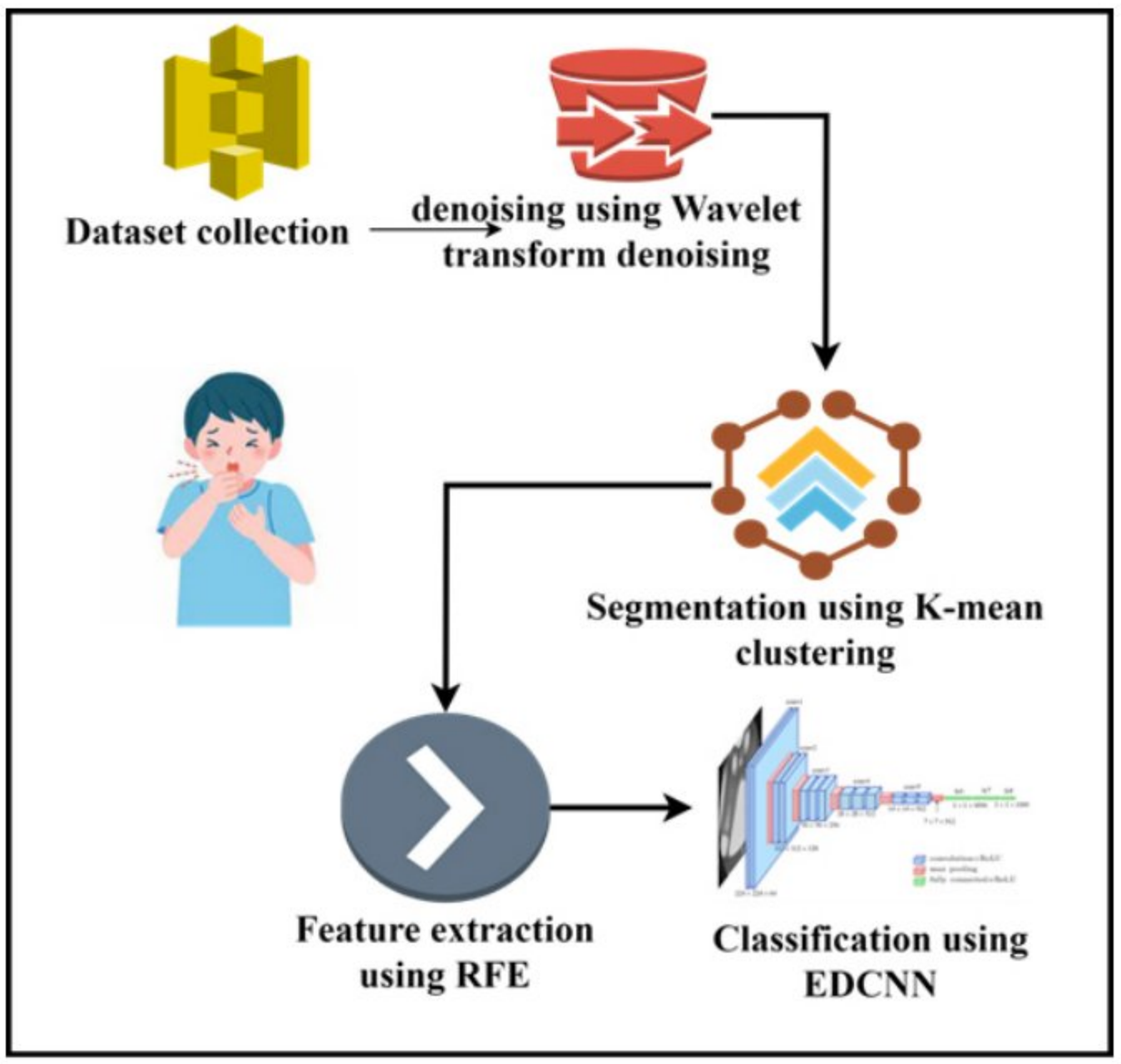

Figure-1 : Workflow architecture

3.1 Dataset collection

The dataset collected from the kaggle website https://www.kaggle.com/datasets/andrewmvd/covid19-cough-audio-classification Cough audio signal classification has been successfully used to diagnose a variety of respiratory conditions, and there has been significant interest in leveraging Machine Learning (ML) to provide widespread COVID-19 screening. This dataset, also known as CoughVid, provides over 25,000 crowdsourced cough recordings representing a wide range of participant ages, genders, geographic locations, and COVID-19 statuses.

3.2 Denoising using Wavelet transform denoising

Denoising using Wavelet transform involves a signal processing technique that aims to remove unwanted noise from a signal while preserving important features. The Wavelet transform decomposes the signal into different frequency components, allowing for the separation of noise from the signal of interest. In this method, the signal is transformed into the wavelet domain where noise typically appears sparse or spread out compared to the signal. By thresholding

or filtering coefficients in the wavelet domain, the noise can be effectively attenuated or removed while retaining the essential components of the signal. Wavelet transform denoising offers several advantages over traditional methods such as Fourier transform-based filtering. It can handle signals with non-stationary or time-varying characteristics more effectively, making it suitable for processing audio signals where noise can vary in both frequency and time domains.

Reconstructing *x(t)* after decomposing it into low-frequency approximation coefficients and high-frequency detail coefficients is as follows :

$$x(t) = \sum_{m=1}^{L}[\sum_{m=1}^{\alpha} D_m\,(k)_{m,k}(t) + \sum_{k=-a}^{\alpha} A_l(k)\emptyset_{L,k}(t)] \quad (1)$$

Where, the detailed signal is at scale 2m, and $\sum_{m=1}^{L}$. at scale 2l, the estimated signal is denoted as$A_l(k)$. In the case of D_m and (t), Through the use of scaling and wavelet filters, $A_l(k)$ is derived. "Tina et al. (2023)"τIn this context, t represents discrete scaling, $D_m(k)$ represents the discrete analysis wavelet, and l k.

$$h(n) = 2^{-1/2} < \emptyset(t), \emptyset(2t-n) > \quad (2)$$

$$g(n) = 2^{-1/2} < \mu(t), \emptyset(2t-n) > \quad (3)$$

$$= (-1)^n h(1-n) \quad (4)$$

A pyramid transfer procedure can be used to calculate the wavelet coefficient. The methods make use of a Fourier transform (FIR) filter bank that includes a low-pass filter (h), a high-pass filter (g), and two-fold down sampling at each step.

3.3 Segmentation using K-mean clustering

Segmentation using K-means clustering in the context of disease detection using human sound involves partitioning audio signals into distinct segments based on similarities in their features. This process aims to identify patterns or segments within the audio signals that cancorrespond to different physiological conditions or disease states. The K-means algorithm iteratively assigns data points, represented by extracted audio features, to clusters based on the Euclidean distance to cluster centroids. By grouping similar segments together, K-means clustering helps identify relevant portions of the audio signals that can be further analyzed or classified for disease detection purposes. This segmentation approach enhances the effectiveness of subsequent processing steps, such as feature extraction and classification, in human sound-based disease detection systems by focusing on relevant segments for analysis.

In a nutshell, the k-means algorithm seeks to reduce the sum of squares inside each cluster.

$$argmin \sum_{i=1}^{k} \sum_{x_j \in S_i} D^2 \left(X_j, \mu_i\right) \quad (5)$$

With k being the total number of clusters, i S being the set of elements in the ith cluster, and i being the cluster mean, the distance between any given point xj and the cluster mean i can be expressed as $2\ j\ i\ D\ (x,)$. These equations describe the third and fourth steps mathematically.

$$S_i^{(t)} = \{x_p : D\left(x_p : D\left(x_p, \mu_i^{(t)}\right) \leq D\left(x_p, \mu_j^{(t)}\right) \forall 1 \leq j \leq k\right\} \quad (6)$$

$$m_i^{(t+1)} = \frac{1}{\left|S_i^{(t)}\right|_{x_j \in S_i^{(t)}}} \sum x_j \quad (7)$$

Iteration t and the number of items in set ith(Si) are denoted by | S_i^((t))|. Si is the unique set to which each point xp is allocated.

Several other kmeans algorithms were developed, all of which were built upon the first k-means algorithm. Keep in mind that the k-means algorithm can only identify a local minimum value for equ.1 if the top k sites are known. The computational time complexity of k-means is affected linearly by the number of clusters, iterations, dimensions, and observations. The problem with this method is that it is too sensitive to outlying data. Consequently, it is recommended to use a variety of outlier analysis approaches to remove the outlier points before using the k-means algorithm.

3.4 Feature extraction using RFE

Feature extraction using Recursive Feature Elimination (RFE) is a technique commonly employed in machine learning to select the most relevant features from a given dataset referred by Habibi, A. et al. (2023). In the context of disease detection using human sound-based systems, RFE is utilized to identify the most informative acoustic features that contribute significantly to the classification of different physiological conditions or disease states. The RFE algorithm operates iteratively by first training a machine learning model on the entire set of features and evaluating their importance based on a predefined criterion. Then, it recursively eliminates the least important features from the dataset and retrains the model on the reduced feature set. This process continues until a specified number of features or a desired level of performance is achieved. By iteratively removing less relevant features, RFE effectively identifies a subset of features that best discriminates between different classes or conditions in the dataset. This subset of features typically contains the most discriminative information while reducing the dimensionality of the data, which can lead to improved classification performance and model generalization.

The result will be inaccurate classifications and a lack of generalizability. Recursive feature elimination, or RFE for short, is an effective and efficient way to reduce the model complexity by removing unnecessary predictors. There are a lot of machine learning algorithms that might utilise RFE as their

main feature selection technique since it is essentially a wrapper approach that leverages filter feature selection. Coefficient or feature significance is then used to rank the characteristics according to their relevance. The model is re-fitted one feature at a time, eliminating the weaker ones. Iterations of the technique continue until the desired number of features is reached.

$$Rank_i = \{r_{i1} = 1, r_{i2} = 2, \dots, r_{ip} = p\} \quad (8)$$

We next determine the feature cut-off points by using eight distinct ML-RFE techniques. These are the qualities that most people think are most important. Accordingly, this technique selects |αP| features from each feature subset in order to construct each optimal feature subset.

$$FS_i^{opt} = \{f_{i1}, f_{i2}, \dots, f_i, |\alpha P|\} \quad (9)$$

Where the round-down operator is represented in mathematics by |αP|

By doing so, we can eliminate features that aren't accurate and robust. To be more precise, imagine that the parameter τ is greater than the AUC of the N best feature sets for predictive classification. Because of this, they are identified as

$$fi^{opt} = \{FS_1^{opt}, FS_2^{opt}, \dots, FS_N^{opt}\} \quad (10)$$

Similarly, we define the robust biomarker screening issue as an N-feature subset stable combination problem. All potential permutations of the sets in FS_2^{opt} are evaluated for stability.

3.5 Classification using EDCNN

An essential part of the suggested illness detection system is the EDCNN classification procedure. The procedure entails using the customized EDCNN architecture after preprocessing the raw audio signals and building a varied dataset. You can trust EDCNN to accurately identify illness signs in audio data since it is built for fast feature extraction, dimensionality reduction, and hierarchical learning. The training and assessment method optimizes the parameters of EDCNN to achieve accurate and generalizable illness classification across varied situations, while data augmentation strengthens the model's resilience.

Superior In-Depth The feature maps are created by twisting the CNN's input layer with the classifier's weight. The results are then passed on after using a non-linear activation function. It is possible to extract patterns from each location, but only if the feature map neurons are same but the conv layer weights are distinct.

The Equation (1)-specified segments from SFCM's output are sent into the Enhanced Deep CNN as input. When N convolutional layers are considered, the output for the units at (u,v) is written as,

$$\left(R_x^g\right)_{u,v} = \left(H_x^g\right)_{u,v} + \sum_{s=1}^{F_1^{s-1}} \sum_{r=-h_1^g}^{h_1^g} \sum_{z=-h_2^g}^{h_2^g} \left(K_{x,s}^g\right) \quad (11)$$

Where$\left(R_x^g\right)_{u,v}$, represents fixed feature map, and represents the convolutional operator needed to extract local patterns from the output generated from the preceding layers. 1

The output of the conv layer, represented byF_1^{s-1}, is passed on as input to the subsequent layer. The Enhanced Deep CNN weight, denoted by $K_{x,s}^g$, is optimized by training using the suggested CCO method. The G conv layer's weight is denoted by $K_{x,s}^g$. The filter $K_{x,s}^g$ links the feature map at layer (g 1) to the th x feature map at the th g conv layer, and its size is $\sum_{s=1}^{F_1^{s-1}} \sum_{r=-h_1^g}^{h_1^g} \sum_{z=-h_2^g}^{h_2^g} \left(K_{x,s}^g\right)$. The $G\ x\ H$ matrices denote the g conv layer's bias settings. The kernel's dimension is given as, which is written as $\sum_{s=1}^{F_1^{s-1}} \sum_{r=-h_1^g}^{h_1^g} \sum_{z=-h_2^g}^{h_2^g} \left(K_{x,s}^g\right)$.

Pooling (POOL), convolutional (conv), and Full Connected (FC) layers make up a Deep CNN. In Deep CNN, each of the three layers serves a distinct purpose. The feature maps are created in the conv layers by subsampling the segments of the preprocessed picture, and the subsampling process continues in the pool layers. The categorization is refined in the FC layer, the third stage. The input maps are subjected to convolution with the convolutional kernels in the convolutional layer, which then generates an output map. The output map has a size proportional to the kernel number, and the kernel matrix has a size of [3*3]. Totally interconnected strata: The FC layer does the high-level reasoning after the pool and conv levels expose the abstract characteristics. Also, the FC layer's output is described as

$$j_x^g = C\left(I_x^g\right) \quad (12)$$

By incorporating elements like bidirectional recurrence to capture both past and future context, attention mechanisms for adaptive focus on informative parts of the sequence, skip connections to mitigate vanishing gradient issues, and regularization techniques like dropout and batch normalization to prevent over fitting, the Enhanced Deep convolutional Neural Network (EDCNN) for classification improves on the classic EDCNN design. Fine-tuning the model's hyper parameters follows optimization through hierarchical structures, ensemble learning, and careful tweaking.

$$g(x_t) = LSTM^{le}\left(x_t^{'}\right) \quad (13)$$

$$m_{t+1} = LSTM^{ld}\left(g(x_t)\right) \quad (14)$$

The output of the encoder's le layers is denoted by $g(x_t)$, whereas the output of the convolutional layers is denoted by x 0 t. The output m_{t+1} of a network with an ld layer decoder is computed based on the current input and the network's state history. So, we have m_{t+1} as the l2-norm model.

Algorithm 1: EDCNN

Input:

- Raw audio signals representing physiological sounds (e.g., heartbeat patterns, respiratory sounds, vocal characteristics).

Steps:

• Preprocessing: Preprocess the raw audio signals to extract relevant features or segments, which are then fed into the EDCNN for further analysis.

$$\left(R_x^g\right)_{u,v} = \left(H_x^g\right)_{u,v} + \sum_{s=1}^{F_1^{s-1}} \sum_{r=-h_1^g}^{h_1^g} \sum_{z=-h_2^g}^{h_2^g} \left(K_{x,s}^g\right)$$

• Feature Extraction: Utilize the Spectral Flatness Centroid Method (SFCM) to extract specific segments from the preprocessed audio data.

•$j_x^g = C\left(I_x^g\right)$

• Bidirectional Recurrence: Incorporate bidirectional recurrence to capture both past and future context of the input sequences.

$$g(x_t) = LSTM^{le}(x_t')$$

• Attention Mechanisms: Utilize attention mechanisms to adaptively focus on informative parts of the input sequences.

$$m_{t+1} = LSTM^{ld}\left(g(x_t)\right)$$

• Skip Connections: Implement skip connections to mitigate vanishing gradient issues and facilitate information flow between layers.

Output:

- Classification Results: Output the predicted illness classifications based on the input audio signals, obtained through the EDCNN classification process.

4. RESULTS AND DISCUSSION

The results and discussion section presents the outcomes of implementing the Enhanced Deep Convolutional Neural Network (EDCNN) for illness detection using audio data. This section analyzes the effectiveness and performance of the proposed EDCNN classification procedure in accurately identifying illness signs from raw audio signals. Furthermore, it discusses the implications of the results and potential avenues for future improvements and applications.

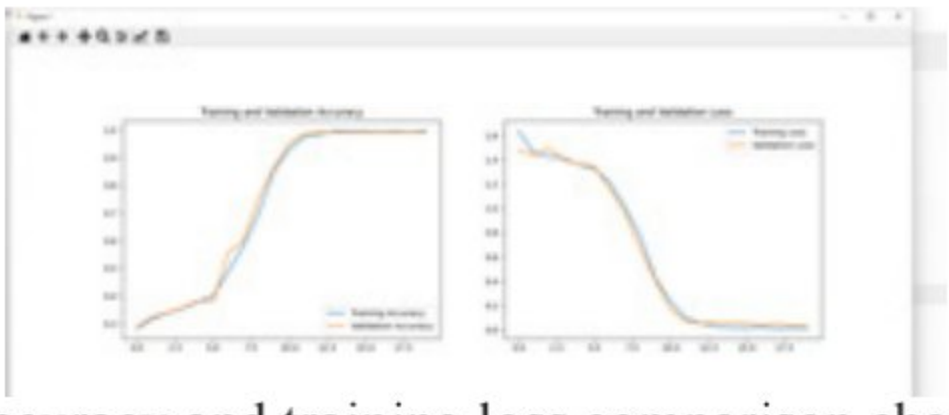

Figure-2: Training accuracy and training loss comparison chart

The figure 2 shows training accuracy and training loss comparison chart the x axis shows epochs and the y axis shows values.

Figure-3: Confusion matrix

4.1 Performance evaluation

1. Accuracy: The fraction of samples with the right classification out of all samples. Mathematically:

$$Accuracy = \frac{(TP + TN)}{(TP + FP + TN + FN)} \quad (15)$$

2. Precision: Ratio of samples with accurate identification to total samples with accurate identification. Mathematically:

$$Precision = \frac{TP}{TP + FP} \quad (16)$$

3. Recall (also known as sensitivity or true positive rate): The proportion of correctly classified samples out of the total number of actual pest samples. Mathematically:

$$Recall = \frac{TP}{TP + FN} \quad (17)$$

4. F1 score: A middle ground between accuracy and memory that strikes a harmonic mean. Mathematically:

$$F1\ score = 2 * Precision * Recall\ /\ (Precision + Recall) \quad (18)$$

Table-1: Classification performance metrics comparison

	Algorithm	Accuracy	Precision	Recall	F-measure
Existing methods	ANN	94.21	94.65	95.01	95.23
	CNN	96.35	95.87	96.10	96.07
	DCNN	97.11	96.32	97.75	98.14
Proposed methods	EDCNN	98.99	98.11	98.84	99.21

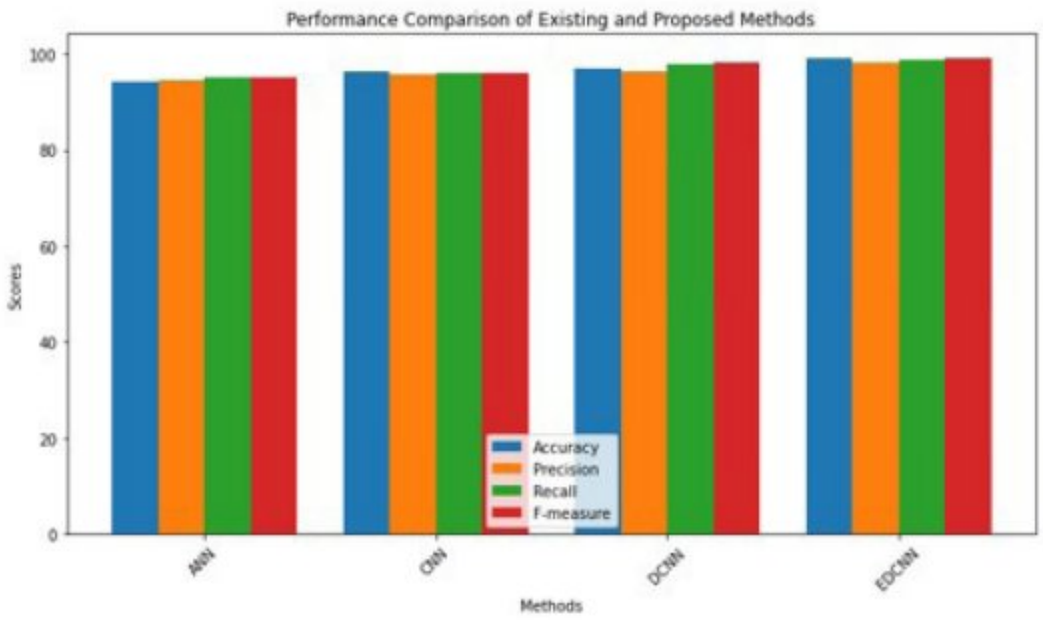

Figure-4: Performance metrics comparison chart

The table 1 and figure 4 shows classification performance metrics comparison table presents the accuracy, precision, recall, and F-measure scores for existing methods (ANN, CNN, and DCNN) and the proposed method (EDCNN). EDCNN outperforms all existing methods across all metrics, achieving the highest accuracy of 98.99%, precision of 98.11%, recall of 98.84%, and F-measure of 99.21%. This indicates that EDCNN demonstrates superior performance in accurately identifying illness signs from audio data compared to traditional neural network architectures. The results highlight the efficacy of the proposed method in achieving both high accuracy and balanced performance across precision, recall, and F-measure metrics, making it a promising approach for illness detection in healthcare applications.

5. CONCLUSION

This research presents a ground-breaking method for illness diagnosis using an Enhanced Deep Convolutional Neural Network (EDCNN) in conjunction with novel human speech inputs. The revolutionary potential of technology in early illness detection is shown by the integration of artificial intelligence, especially deep learning, into healthcare. The urgent need for an efficient, low-cost, and non-invasive diagnostic tool is satisfied by the suggested system. The system takes use of the physiological insights conveyed by human sound signals by using the extensive information included in these signals. A viable path for early detection can be found by analyzing small fluctuations and patterns within the sound data, which allows for the identification of potential markers for different illnesses. The enhanced EDCNN architecture for sound data processing is the backbone of our technique. EDCNN outperforms all existing methods across all metrics, achieving the highest accuracy of 98.99%, precision of 98.11%, recall of 98.84%, and F-measure of 99.21%. Feature extraction, dimensionality reduction, and hierarchical learning are some of the sophisticated approaches that have been used to improve the model's ability to detect complex patterns in audio data. Incorporating data augmentation tactics also improves the model's capacity to generalize and handle different patient groups and circumstances.

REFERENCES

1. Altshuler, Ellery, Bouchra Tannir, Gisèle Jolicoeur, Matthew Rudd, Cyrus Saleem, Kartikeya Cherabuddi, Dominique Hélène Doré et al. Digital cough monitoring–A potential predictive acoustic biomarker of clinical outcomes in hospitalized COVID-19 patients. *Journal of Biomedical Informatics* 138 (2023): 104283.

2. Cai J, Vhaduri S, Luo X. Discovering COVID-19 Coughing and Breathing Patterns from Unlabeled Data Using Contrastive Learning with Varying Pre-Training Domains. *arXiv preprint* arXiv:2306.01864. 2023 June 2.

3. Campana, M.G., Delmastro, F. and Pagani, E., Transfer learning for the efficient detection of COVID-19 from smartphone audio data. *Pervasive and Mobile Computing*, 89, pp101754 2023.

4. De Souza, L., Bernardino, H., de Souza, J., & Vieira, A. COVID-19 Detection Using Forced Cough Sounds and Medical Information. *Revista de InformáticaTeórica e Aplicada,* 30(1), 44-52. 2023.

5. Hamidi, Mohamed, OuissamZealouk, Hassan Satori, NaouarLaaidi, and Amine Salek. COVID-19 assessment using HMM cough recognition system. *International Journal of Information Technology,* 15, no. 1 (2023): 193-201.

6. Kaulamo, J.T., Latti, A.M. and Koskela, H.O., Healthcare-seeking behaviour due to cough in Finnish elderly: too much and too little. Lung, 201(1), pp.37-46. 2023.

7. Najaran, Mohammad Hassan Tayarani. An evolutionary ensemble learning for diagnosing COVID-19 via cough signals. *Intelligent Medicine* (2023).

8. Naqvi, Kubra F., Stuart B. Mazzone, and Michael U. Shiloh. Infectious and Inflammatory Pathways to Cough. *Annual review of physiology,* 85, pp.71-91, 2023.

9. Pahar, M., Miranda, I., Diacon, A. and Niesler, T., 2023. Accelerometer-based bed occupancy detection for automatic, non-invasive long-term cough monitoring. IEEE Access, 11, pp.30739-30752.

10. Valergaki, P., COVID-19 Diagnosis from cough samples using Deep Learning methods, 2023.

11. Vhaduri, S., Dibbo, S.V. and Kim, Y., 2023. Environment Knowledge-Driven Generic Models to Detect Coughs From Audio Recordings. *IEEE Open Journal of Engineering in Medicine and Biology*.

12. Vishniakou, U. A., and B. H. Shaya. Voice Detection Using Convolutional Neural Network. 21, no. 2 (2023): 114-120.

13. Wu, Yuan, Jian Zhang, Yanjiao Chen, Junkongshuai Wang, WuXuan Shi, and Qian Zhang. Ubi-Asthma: Towards Ubiquitous Asthma Detection using the Smartwatch. *IEEE Internet of Things Journal* (2023).

14. Zealouk, Ouissam, Hassan Satori, Mohamed Hamidi, NaouarLaaidi, Amine Salek, and Khalid Satori. Analysis of COVID-19 resulting cough using formants and automatic speech recognition system. *Journal of Voice* 37, no. 6 (2023): 971-e9.

15. Zhu, Yi, Mahil Hussain Shaik, and Tiago H. Falk. On the importance of different cough phases for COVID-19 detection. *IEEE International Conference on Acoustics, Speech and Signal Processing* (ICASSP-2023), pp. 1-5. IEEE, 2023.

16. Tian, Chunwei, Menghua Zheng, Wangmeng Zuo, Bob Zhang, Yanning Zhang, and David Zhang. Multi-stage image denoising with the wavelet transform. *Pattern Recognition* 134 (2023): 109050.

17. Habibi, A., Delavar, M. R., Sadeghian, M. S., Nazari, B., and Pirasteh, S. . A hybrid of ensemble machine learning models with RFE and Boruta wrapper-based algorithms for flash flood susceptibility assessment. *International Journal of Applied Earth Observation and Geoinformation*, 122, 103401. 2023.