COVID CLASSIFICATION FROM COUGH SOUNDS USING DEEP LEARNING BASED DEEP CONVOLUTIONAL AUTO ENCODERS

KEERTHIYA R
Ph.D. Research Scholar
Department of Computer Applications
St. Peter's Institute of Higher
Education and Research, Chennai,
Tamil Nadu, India
keerthiya27694@gmail.com

LATHA R
Head of the Department
Department of Computer Applications
St. Peter's Institute of Higher Education
and Research, Chennai,
Tamil Nadu, India
latharamavel@gmail.com

***Abstract*— In this study, the variations are analyzed in the auditory noise that can be detected in COVID-19 by making recordings of real people voices and listening to those recordings. However, because to the vast number of datasets that are now available, it can be challenging to locate COVID-19 patients through cough sounds. In this paper, A deep convolutional auto encoders (DBN)-based classification model designed to monitor the audio stream for cough sounds. The model bases its training on the datasets that were used for the training. Touse a relatively modest sample of patient self-recorded voices from a number of different time periods in order to train the model. These voices were taken from different points in the patient lives. The results of the simulation suggest that the strategy that has been suggested is superior to the methods.**

Keywords — *Cough Detection, Deep Learning, Data Augmentation, Covid-19, Deep Convolutional Auto Encoders*

I. INTRODUCTION

Coughing is a common condition that affects roughly 33% of the world population [1], and even the youngest of children have a possibility of experiencing it at some point in their lives. The majority of people, when faced with this situation, seek medical assistance as their first and foremost course of action. A severe cough, which is frequently brought on by a cold, typically disappears on its own after about two weeks have passed. On the other hand, a persistent cough is responsible for 10–38% of all inquiries for treatment for respiratory illnesses. The development of a chronic cough is associated with a number of factors, including but not limited to exposure to environmental contaminants, smoking cigarettes, and being exposed to secondhand smoke. There are many conditions, both respiratory and non- respiratory, that can lead to a persistent cough [2] [3].

It is imperative that special attention be paid because of the regularity with which it occurs and the enormity of the costs that are associated with it. Despite how widespread it is, cough is a difficult ailment to diagnose and assess correctly. This is due to the fact that cough symptoms can vary greatly [4]. Long before their reliability was increased objective techniques relied on manual assessment, which limited their deployment. This continued even after their dependability was enhanced. The situation remained the same well into the 21st century. When new ways of acquisition technology and processing technology arose in the early 2000s, automatic processing finally became a viable alternative [5]. Prior to this time, automatic processing was not a possibility at all. This method is frequently utilized in objective cough assessment due to the fact that listening to a person cough is a simple, non-invasive, and obvious strategy to determining the strength of a cough [6].

A clinician can utilize a patient cough as a diagnostic or prognostic indicator for a wide range of ailments, including pneumonia and asthma, to mention just two of those conditions. The cough of a patient can be used in either of these ways. Recent research has used AI-guided methods in a variety of different ways to investigate the classification of coughs for the goals of disease diagnosis and prognosis. This has allowed researchers to obtain a better grasp of how to classify coughs [7].

It is conceivable to build AI-guided tools and techniques for audio and speech processing in order to pre-screen COVID-19 recordings obtained from coughing by identifying and making use of acoustic biomarker features. This can be done in order to protect against the spread of the COVID-19 virus. Before we get into the specifics, let first go over some examples of where biomarkers and tools or techniques based on smartphones have been applied in the past. These examples will be taken from the world of smartphones [8].

During the course of the investigation into COVID-19, it has been recommended by a variety of authorities that methods which include speech and signal processing be utilized. It is helpful to use a dataset with a high degree of neuro motor synchrony across the breathing, phonation, and articulation subsystems of speech in order to detect COVID-19 in individuals who are asymptomatic as well as in individuals who have symptoms associated with the infection. This can be done in order to detect COVID-19 in individuals who have the infection but do not have any symptoms. When paired with technologies guided by AI, these characteristics of the biomarker have the potential to dramatically improve the accuracy with which forced cough COVID-19 is detected [9] [10].

This research only looked at a small group of patients, but if the findings are valid, it demonstrates that utilizing cough as a screening metric could assist in recognizing infections in their latter stages. According to the findings of this study, coughing is one of the symptoms that is mentioned the most often, and as a consequence, it is the symptom that occurs the second most commonly after fever. With such motivation, in this paper, develop a deep convolutional auto encoders to classify the cough sounds from the input audio dataset. The model classifies the cough sounds to predict if the sounds of the patient have covid or not.

II. RELATED WORKS

Dash et al. [11] made use of the MFCCs of the various speech signals. The researchers employed 570 participants from the Coswara dataset, which they referred to as database-1, and they asked each participant to send nine audio recordings to a range of sample groups. The researchers used the data from the Coswara dataset. They did not specify, however, whether the data had been independently confirmed by a laboratory or not. Following that, the authors made use of a database that was compiled through the use of crowdsourcing and had a total of 6,631 individuals. Out of these individuals, 235 were found to have a positive result for COVID-19. In order for the authors to derive the greatest possible benefit from the recordings, they took the necessary steps to ensure that the frequency range as well as the conversion scale were precisely calibrated. The authors constructed the manufactured data by using an adaptive synthetic sampling technique for unbalanced learning. This allowed for the production of more accurate results. They found that the accuracy of database-2 (database-1) cough sounds had improved to 0.86 (0.74) since the beginning of the study as a result of their research and came to this conclusion as a result of their researches.

Mouawad, et al. [12] did research on a symbolic recurrence quantification measure that was generated from MFCC features in order to detect COVID-19 instances using cough sounds and speech. This was done in order to detect cases of COVID-19 using cough sounds and speech. This was done in order for them to reach their objective successfully. They used a recurrence dynamic and a variable Markov model on recordings of people saying ah for extended periods of time in order to demonstrate that their model is effective at detecting the beginning of the disease in utterances of this kind. This was done in order to demonstrate that their model is effective at detecting the beginning of the disease in utterances of this kind. Only 32 patients and 20 speakers contributed their cough sounds and speech records, respectively, for the collection. There was a total of 1,488 speech recordings in the collection. The authors tried out a variety of data sampling procedures in order to make certain that the classification model did not have any sort of bias toward the class that was considered to be the majority. One of these procedures involved oversampling the minority class and under sampling the class that was considered to be the majority. They found that their accuracy was 0.91 and that their area under the curve (AUC) was 0.84 when they carried out the tests that they had planned.

Deep learning algorithms were applied by Shimon et al. [13]. The information that was obtained through the use of crowdsourcing was not readily available to the general public and was made up of a variety of recordings that were created by participants over the course of a number of days. In order to extract the auditory characteristic, the researchers made use of a wide variety of methods, such as the openSMILE toolkit and Librosa. After the classifiers had categorized each audio clip, the experimenters used a voting method called simple majority to decide whether or not the patient had been infected with COVID-19.

Despotovic et al. [14] utilized a number of deep learning methods, such as deep audio embeddings and wavelet scattering features. The methods described here were implemented by the researchers who carried out the study. Their dataset was generated using data obtained through crowdsourcing, with 1,103 individuals contributing two to five samples apiece. This number represents the total number of samples. These were utilized by the authors in order to compile a thorough dataset that featured 496 distinct cough samples. For the purpose of ensuring the accuracy of their findings, they carried out a fivefold cross-validation approach with the exception of one component. It would appear that accuracy, sensitivity, and specificity were all at their absolute peak when using the wavelet scattering characteristics as a boost.

Melek [15] assess whether coughing was a positive or negative indicator for COVID- 19, he utilized a number of different deep learning architectures. It is possible to make use of a wide variety of alternative architectures, some of which include polynomial-SVM, linear-LDA, and Euclidean-kNN, to name just a few of the available options. During the course of their investigation, they used a total of 121 datasets derived from Virufy as well as 59 datasets derived from NoCoCoDa.

Nessiem et al. [16] investigate the possibility of using deep learning models as a widespread, low-cost pre-testing method for identifying COVID-19 from audio recordings of breathing or coughing that were taken on mobile devices while connected to the internet. The recordings were made by coughing or breathing while the devices were connected to the internet. They started by gathering 1,427 user- submitted audio samples into a database, and then worked their way backwards through the information. Next, the researchers trained an ensemble of three convolutional neural networks to detect COVID-19 infection by evaluating breathing and coughing raw audio, spectrograms, and Mel-spectrograms.

Gokcen et al. [17] classified individuals according to whether or not they had COVID-19 infection using a deep neural network (DNN). Although the publication did not provide an exhaustive description of their neural network, it does suggest that they made use of a relatively dense network of neurons. The sample size of the study, which consisted of 822 coughs selected from the dataset that was publicly available at MIT, is relatively simple to identify. Even though it appears to be a rather straightforward neural network, it was nevertheless able to achieve an accuracy of 0.79 and a recall of 0.75.

Feng et al. [18] deployed a recurrent neural network (RNN) in order to differentiate between persons who are infected with COVID-19 and those who are not infected with the virus. This was done in order to determine which individuals carry the virus. The Coswara dataset was used for both the validation and training sets, but the Virufy dataset was used for the test set in the end. There were 1,433 people that took part in the study that was conducted by

Coswara. They used an 80/20 split in their work. The small sample size of the Virufy test set only included a total of 16 recordings, each of which was taken from one of seven different patients. Although it performed well when it was tested on the training and validation sets, when it was applied to the Virufy dataset, their model only achieved an accuracy of approximately 0.81 and an area under the curve of approximately 0.79. This is despite the fact that it performed well when it was tested.

Brown et al. [19] employed deep CNN that was augmented with numerous feature channels and data. The outcomes of their efforts were far superior to the outcomes of any previous attempts they had undertaken. Regrettably, it was not made apparent if the better performance was attributable to the utilization of a larger dataset, the implementation of the DCNN classifier, or whether all of these things combined contributed to the improved performance. Furthermore, the authors do not make it quite apparent whether or not their findings have been confirmed in a laboratory setting. They say that patients have been permitted inside the clinic, but they don't explain what that phrase means in further detail. When the positive and negative occurrences of COVID-19 were compared to one another, the accuracy was found to be 0.95.

To further clarify, diagnostic tests are intended to determine whether or not a disease is present, whereas screening tests look for early warning signs of disease. Tests that are diagnostic are also sometimes referred to as confirmatory tests. In a nutshell, we make use of tools that are directed by AI in order to carry out an intensive analysis of screening processes. This allows us to examine everything thoroughly. Methods that are directed by AI, which are one of the tools that are both the quickest and most accurate at detecting and screening COVID-19, based on data that has been clinically diagnosed and thoroughly observed by medical experts (cough sounds). The analysis of cough sounds is the primary focus of this study because coughing is one of the most common symptoms, it can be done in a relatively short amount of time and at a reasonable cost for a large number of participants, and it is being examined as part of this study. During screening, it possible that other pieces of information, such the sounds of the respiratory system and breathing, will be taken into consideration.

III. PROPOSED METHOD

This paper develops a deep convolutional auto encoders (DBN)-based classification model with the goal of recognizing coughs in an audio dataset.

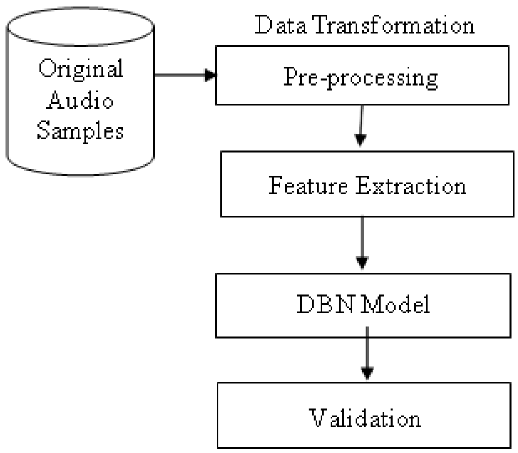

Fig.1. Proposed Model

3.1 Dataset description and preprocessing

During the course of our empirical research, we made use of the following four datasets: Coswara, Cambridge, Virufy, and Virufy+NoCoCoDa. A Hann window is applied to the samples of coughing after they have been resampled at a rate of 22.5 kHz.

CAMBRIDGE DATASET:

The University of Cambridge has developed a program that is able to monitor how a person reads a certain line, including how they talk, cough, and breathe while doing so. The program was used to create the program. For instance, in the Cambridge dataset, COVID-19-positive participants are separated from non-COVID-19-positive persons by splitting them into two groups: those with symptoms and those without symptoms. This allowed us to identify COVID-19-positive participants. Due to the fact that its creators at Cambridge University made it accessible to the general public in accordance with a legal arrangement that was one-to-one,looked into it. Because of this restriction, the dataset could only be accessed and exploited for the purpose of doing research within an academic institution such as a library or another form of similar establishment.

When utilizing this dataset, it is simple to distinguish between patients who have a positive test result for COVID-19 and those who do not have such a result. It would be unjust to treat everyone the same if they had a positive test result, especially if they did not have any warning signs, did not have a family history of the disease, and had never smoked. However, it would be fair to treat them the same if they had a positive test result. 141 of the cough samples proved positive for COVID-19, whereas the other 298 samples tested negative for the virus (those who had never smoked, had a clean medical history, and had no symptoms.

COSWARA DATASET

The Coswara dataset includes 1,134 negative cough samples and 185 positive cough samples for the purposes of both training and testing. The total number of negative cough samples in the dataset is 1,134.

VIRUFY DATASET

The first COVID-19 cough sound dataset that is made accessible to the general public is the Virufy COVID-19 open cough dataset. It was taken from the patient in a medical facility, and the patient informed consent was obtained before the operation was carried out. Additionally, the procedure was carried out under medical supervision and in accordance with standard operating protocols. The infection status of patients has been preprocessed and tagged as part of the scope of this dataset. A total of 121 sputum samples were obtained through the participation of 16 different persons (48 were positive, and the remaining were negative).

NOCOCODA DATASET

Patients who have tested positive for COVID-19 and who have been interviewed by the public media during coughing episodes are included in the NoCoCoDa data collection. It is possible that this hacking cough began at any moment before, during, or after the emergency situation. It even possible that it started before the emergency scenario that we're in now. After manually segmenting the interview,to identify 73 separate coughing episodes and categorize the coughing phases that were associated with each of those coughing episodes. This allowed us to better understand the patient coughing pattern. Both the NoCoCoDa dataset, which only contained COVID-19 samples, and the Virufy dataset were combined for use in this experiment. Both datasets were combined into one for your convenience (which had both COVID-19 samples and other samples). Sadly, none of the datasets that were accessible had any COVID-19-free samples in their entirety.

3.2 Feature Extraction Method

An audio signal is a representation of a sound that is made up of a series of integers that are spread out at regular intervals in time. An audio signal is also known as a time-frequency representation. This page has a lot of information, but sadly, it is not easy to understand what all of it means. It requires additional processing in the same manner as the ear and the brain do it. The signal is going to be described by making use of a small number of aspects that are meant to be representative of the entire thing.

Features in the frequency domain, which are also known as spectral characteristics, provide information about the spectral content of the signal, whereas features in the time domain, such as duration and amplitude, provide information about the temporal evolution of the signal. Spectral characteristics and features in the frequency domain both provide information about the signal. The information about the signal can be obtained from the frequency domain as well as the time domain.

Mel-Frequency spectral domain is formed by first translating the spectrum into the logarithm of its magnitude, and then doing the reverse calculation in order to return to the pseudo-time domain. This is done after the p-to-178 features have been discovered. During the process of putting the final results in order of importance, the authors discarded any characteristics that were not included in the final results as well as any qualities that were not connected to sound in any way. The total of 178 is not an exact count of the entire number of people; rather, it is more of an estimate than a count.

Each of the characteristics has been randomly placed into one of 17 categories based on the kind of information that it is designed to collect. Despite the fact that there are other classifications that could be just as appropriate, this offers a complete picture of the situation. Cepstral coefficients are still put to use in a substantial quantity of situations nowadays. The classification that is referred to as deep learning raw data includes a subset of characteristics that is known as custom features, and this subset is located within the classification. They are well suited for the rapidly developing deep learning industry, which can take advantage of their capabilities. Cepstral analysis is currently being examined as a possible option for cough analysis as a result of the success it has had when used to speech analysis.

The sample of the sound waveform that is used in feature extraction is taken at a frequency of 22 kilohertz, which is the standard frequency used by the industry for audio applications. As a consequence of this, it is possible to rely on the procedure being carried out in the same manner each and every time. Users are able to isolate five spectral properties of the sampled sound thanks to the librosa module for the Python programming language (Mel-Scaled Spectrogram, Mel-Frequency Cepstral Coefficients, Chromagram, Tonal Centroid, and Spectral Contrast).

One of the useful aspects of audio analysis is the mel-frequency cepstral coefficients, also known as MFCCs. In the course of study, these coefficients have been put to use to discern between dry and wet coughs. After a windowing operation has been completed, the fast fourier transform, abbreviated as FFT, is utilized so that the power spectrum of each frame may be determined during the process of MFCC feature extraction. After that, the power spectrum goes through the process of having a filtering network applied to it. The Mel scale serves as the basis for this network. In order to develop a mel- scaled filter, we need to find a solution for the following equation. In this equation, the independent variable will be denoted by the symbol f, which stands for the physical frequency. We are able to carry out the computations necessary to determine the MFCC coefficients.

$$f_{mel}=2595\log_{10}(1+f_{700}) \quad (1)$$

Following the transformation of the power spectrum to the logarithmic domain, the discrete cosine transform (DCT) can be implemented on the audio stream.

VOICE CLASSIFICATION USING DEEP AUTOENCODERS

Transfer learning enables individuals to draw from their existing stores of knowledge and experience in order to identify workable solutions when confronted with new challenges or in unfamiliar contexts. One of the most important architectural components for the process of transfer learning is known as the transformer. This component includes both an encoder and a decoder.

The encoder is responsible for converting the input representation (in the case of WSD, a sequence of input symbols) into an intermediate representation (a contextual embedding), and the decoder is responsible for converting the intermediate representation into the desired output representation (a sequence of target symbols). In the case of

WSD, the input representation is a sequence of input symbols. In the instance of WSD, the representation of the input is a sequence of the symbols that are being read in.

SELF-ATTENTION MECHANISM IN TRANSFER LEARNING

A self-attention mechanism is currently under development with the purpose of capturing the relevance of each sound in relation to the remaining sounds. This is being done with the intention of enhancing the capacity to transmit a diverse range of contextual meaning and providing a more meaningful contextual representation for WSD. The meaning score of each sound is kept up to date by adding the global score of the entire context to the attention score for each pair of sounds, which is calculated by the self-attention mechanism. This brings the total attention score for each pair of sounds to the total attention score for each pair of sounds. Because of this, the overall score for the attention that was devoted to each pair of sounds has now reached its highest conceivable value The sentential context vectors of the sound Wamb are represented by Camb, and the series of n sounds is designated by Cs=[W1,W2, Wamb, Wn]. Wamb∈R512 is located at the location where Camb is located. Self-attention has as its primary purpose the encoding of all Camb entities through the utilization of information regarding the global context.

In order to calculate the attention weights that will be utilized by Camb between the two sound vectors, we make use of the dot products that are generated by the query (represented by Qn) and the transposition of the key (represented by KnT). However, the values of these scalars Snn are subject to significant amounts of variation; hence, in order to normalize Camb to the range 0–1, we make use of the softmax function wnn, which

$$w_{nn} = Softmax[S_{nn}] = \frac{Q_n \cdot K_n^T}{\sqrt{d_k}} \qquad (2)$$

Where dk- Kvector dimension The contextualized sound embedding en for the sound vectors in Camb can be derived by applying the weighted sum of attention weights wnn and values V1, V2, V3,…, Vn in accordance with the equation that is displayed in Eq. (3)

$$e_n = \sum_{i=1}^{n} w_{nn} V_n \qquad (3)$$

Where

n- input vectors So, are the relevancies of all the sound with each other in the context, where is the relevance of all sound to describe the sound.

The contextual dependencies that are present among all sounds are represented by the set S=(S1,S2,S3,…,Sn), where s1 represents the set of all sound dependencies on the sound W1. The Sn values can vary quite a bit from one another, it is vital that the data be standardized. For the purpose of standardizing the range of values from 0 to 1, the function known as softmax (Sn) is utilized shown in Eq (4).

wn1=Softmax[Sn1],…,wnn=Softmax[Snn] (4)

The weights of all the contextual relevance with each other are denoted by w=(w1,w2,w3,…,wn), where w1 is the weight used to determine the contextual embedding for sound W1, w2 is the weight used to determine the contextual embedding for sound W2, and so on and so forth.

$$w_n = (w_{31}, w_{32}, w_{33}, \ldots, w_{nn}) \text{ where } \sum_{i=1}^{n} w_n = 1 \qquad (5)$$

As a result, the contextualized embedding(s) for the input sound vectors W1, W2, W3, W4,…Wn will be determined by using the weighted sum of all Vn.

eWn= wn1V1+wn2V2+wn3V3+wn4V4+⋯+wnnVn (6)

where, eWn influence [W1, W2,….Wn] and it is expressed as wnn

These do not convey the same direction that the original sounds did; rather, they imply a different direction. The surrounding sounds have a significant influence on the contextual representation eW1, eW2, eW3,…, eWn are of each sound W1, W2,….Wn, and this is true regardless of the score you have or the setting in which you are listening. This is the case regardless of the setting in which the sound is being experienced by the listener.

IV. RESULTS AND DISCUSSIONS

The validation features and the numerical findings have both been compiled to their final form.When aSingle study includes the findings of more than one experiment, we whittled it down to the ones that the authors of the study as well as ourselves felt best exemplified the overall findings of the study. Even if broad generalizations will be drawn, it is necessary to keep in mind that the outcomes of the research will not be directly comparable to one another. It is customarily necessary to have a substantial amount of training data available for machine learning classifiers before they can perform effectively.

However, there is still a constraint on the data quality because it is difficult to obtain enormous volumes of data and manually classify them. This is one of the reasons why there are limitations on data quality. This is only one of the many reasons why machine learning is gaining more and more attention these days. In addition to this, it is absolutely necessary to put the system through its paces utilizing data that was not covered during the training phase. In order to get the most out of the data that is being used, many cross-validation methods are utilized. Either collecting a whole new set of data for validation and excluding its use in training or reserving a portion of the data for validation before putting it to use in training is an element of the test data strategy. Neither of these options is recommended. The test data plan takes into consideration both of these potential courses of action.

The data is initially randomly split into a training subset and a testing subset using the Monte-Carlo approach before an average of the findings is presented. After that, this operation of segmentation and testing is carried out multiple times, typically anywhere from tens to hundreds of times. There are no limitations for the amount of time that should be spent on each individual subject while using this method, despite the fact that it makes it simpler to test a large sample size. The data are first randomly segmented into k folds of equal size as the first step in the k-fold methodology. The system is next trained on folds that are one fewer than the total number of folds, and it is ultimately tested on the final fold. After that, the folds are rotated in such a way that each fold is tested exactly once, and then the results of the averaged tests are supplied. After that, the results of the tests are presented. The final technique is termed the leave-one-out method, and it involves training the system on all of the data except for one sample before testing it on that one- sample subset. This method gets its name from how it trains the system on all of the data except for one sample. This process is performed as many times as necessary until each sample has been utilized as a test sample precisely once. In this scenario, the k-fold cross-validation method might be applied as a testing strategy.

The purpose of this study was to investigate the feasibility of using cough sounds as diagnostic tools for COVID-19 and to do so, cough recordings were evaluated. After this, the classifiers are put through a DAE ranking in order to show that our technique is effective at classification when applied to the Cambridge dataset. Following this, to give proof that the feature selection technique known as recursive feature elimination (RFE) performs magnificently on each and every dataset that was utilized in our investigation and that our strategy is successful at classification when it is applied to the Cambridge dataset.

The method of feature selection known as recursive feature elimination (RFE) performs exceptionally well on each and every dataset that was utilized in our investigation. The authors of this piece of work are the ones who should be held responsible for the strategy that was developed in it. As a consequence of this, to put this methodology through its paces by utilizing a number of datasets, and afterwardcompared the outcomes to the industry benchmark

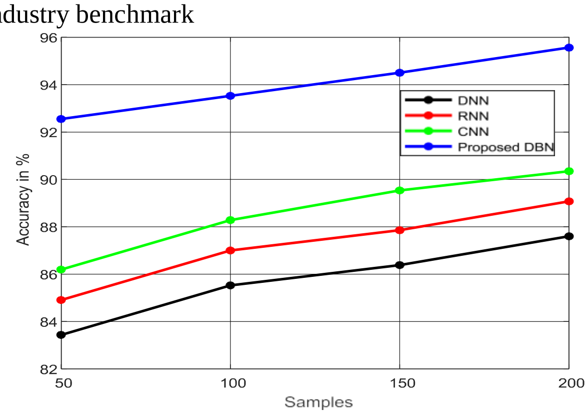

Fig.2. Accuracy

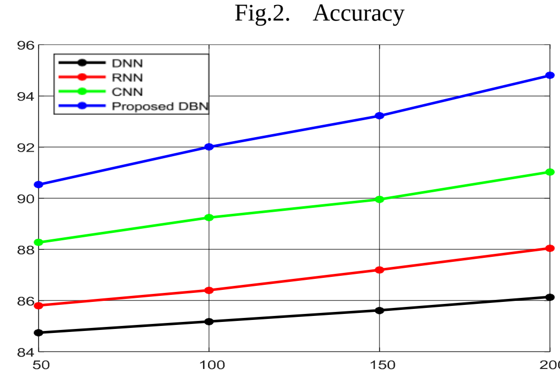

Fig.3. Precision

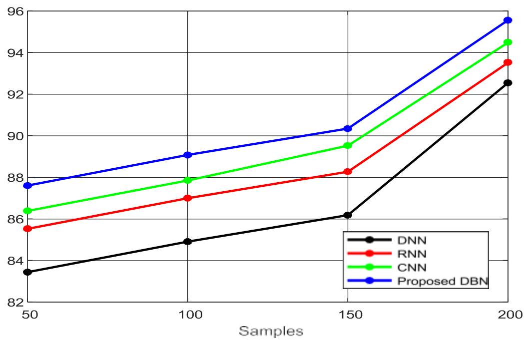

Fig.4. Recall

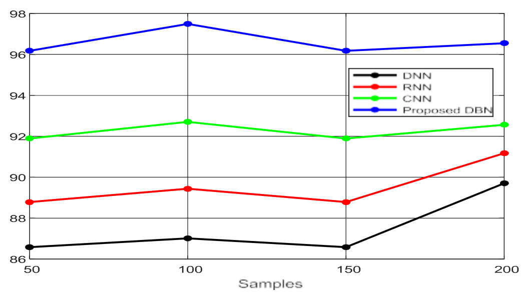

Fig.5. F-Measure

Figures 2–5 present the results of the simulation, which reveal that the suggested model is superior to the baseline in terms of accuracy, precision, recall, and F-measure. The results of the simulation are displayed in these figures. The DAE performs extremely well when applied to the classification problem that is posed by the dataset, and it generates test results that are superior to those produced by other competing algorithms.

V. CONCLUSIONS

For the purpose of recognizing coughs in real-time audio streams, to trained a DAE- based classification model utilizing a variety of open-source datasets. The diagnosis of coughs was our primary focus. Only a very tiny part of the noises that were caught by the clinician at a variety of time periods are used to train the model in the early phases of the training process. This is because there are so many different time periods that the clinician recorded the noises at. According to the results of the simulation, the methodology that has been suggested is superior to the methodologies that are currently deemed to be state-of-the-art in terms of their level of accuracy.

References

[1] Lella, K. K., & Pja, A. (2022). Automatic diagnosisof COVID-19 disease using deep convolutionalneural network with multi-feature channel from respiratory sound data: cough, voice, and breath. Alexandria Engineering Journal, 61(2), 1319-1334.

[2] Khadidos, A., Khadidos, A. O., Kannan, S., Natarajan, Y., Mohanty, S. N., & Tsaramirsis, G. (2020). Analysis of covid-19 infections on a ct image using deepsense model. Frontiers in public health, 8, 599550.

[3] Lella, K. K., & Pja, A. (2021). Automatic COVID-19 disease diagnosis using 1D convolutional neural network and augmentation with human respiratory sound based on parameters: cough, breath, and voice. AIMS public health, 8(2), 240.

[4] Kannan, S., Yuvaraj, N., Idrees, B. A., Arulprakash, P., Ranganathan, V., Udayakumar, E., & Dhinakar, P. (2021). Analysis of convolutional recurrent neural network classifier for COVID-19 symptoms over computerised tomography image. International Journal of Computer Applications in Technology, 66(3-4), 427-432.

[5] Maheshwari, V., Mahmood, M. R., Sravanthi, S., Arivazhagan, N., ParimalaGandhi, A., Srihari, K., ... & Sundramurthy, V. P. (2021). Nanotechnology- based sensitive biosensors for COVID-19 prediction using fuzzy logic control. Journal of Nanomaterials, 2021, 1-8.

[6] Kumar, S., Gupta, S. K., Kumar, V., Kumar, M., Chaube, M. K., & Naik, N. S. (2022). Ensemble multimodal deep learning for early diagnosis and accurate classification of COVID-19. Computers and Electrical Engineering, 103, 108396.

[7] Abdel-Jaber, H., Devassy, D., Al Salam, A., Hidaytallah, L., & El-Amir, M. (2022). A review of deep learning algorithms and their applications in healthcare. Algorithms, 15(2), 71.

[8] Hamdi, S., Oussalah, M., Moussaoui, A., & Saidi, M. (2022). Attention-based hybrid CNN-LSTM and spectral data augmentation for COVID-19 diagnosis from cough sound. Journal of Intelligent Information Systems, 59(2), 367-389.

[9] Amin, J., Anjum, M. A., Sharif, M., Rehman, A., Saba, T., & Zahra, R. (2022). Microscopic segmentation and classification of COVID-19 infection with ensemble convolutional neural network. Microscopy research and technique, 85(1), 385-397.

[10] Chahar, S., & Roy, P. K. (2021). Covid-19: A comprehensive review of learning models. Archives of Computational Methods in Engineering, 1-26.

[11] Dash, T. K., Mishra, S., Panda, G., & Satapathy, S.C. (2021). Detection of COVID-19 from speech signal using bio-inspired based cepstral features. Pattern Recognition, 117, 107999. Mouawad, P., Dubnov, T., & Dubnov, S. (2021). Robust detection of COVID-19 in cough sounds: using recurrence dynamics and variable Markov model. SN Computer Science, 2(1), 34.

[12] Shimon, C., Shafat, G., Dangoor, I., & Ben-Shitrit, A. (2021). Artificial intelligence enabled preliminary diagnosis for COVID-19 from voice cues and questionnaires. The Journal of the Acoustical Society of America, 149(2), 1120-1124.

[13] Despotovic, V., Ismael, M., Cornil, M., Mc Call, R., & Fagherazzi, G. (2021). Detection of COVID-19 from voice, cough and breathing patterns: Dataset and preliminary results. Computers in Biology and Medicine, 138, 104944.

[14] Melek, M. (2021). Diagnosis of COVID-19 and non- COVID-19 patients by classifying only a single cough sound. Neural Computing and Applications, 33(24), 17621-17632.

[15] Nessiem, M. A., Mohamed, M. M., Coppock, H., Gaskell, A., & Schuller, B. W. (2021, June). Detecting COVID-19 from breathing and coughing sounds using deep neural networks. In 2021 IEEE 34th international symposium on computer-based medical systems (CBMS) (pp. 183-188). IEEE.

[16] Gökcen, A., Karadağ, B., Riva, C., & Boyacı, A. (2021).Artificial intelligence–basedCOVID-19detection using cough records

[17] Feng, K., He, F., Steinmann, J., & Demirkiran, I. (2021, March). Deep-learning based approach to identify COVID-19. In SoutheastCon 2021 (pp. 1-4). IEEE

[18] Brown, C., Chauhan, J., Grammenos, A., Han, J., Hasthanasombat, A., Spathis, D., ... & Mascolo, C. (2020). Exploring automatic diagnosis of COVID-19from crowdsourced respiratory sound data. arXiv preprint arXiv:2006.05919