International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

A Survey on the Identification of COVID-19 from Cough Audio Sounds Using Deep Learning

[1]V Aravind, [2]V Harish Kumar, [3]S V Chandra Sekhar, [4]V Sravani, [5]V Vineela

[1,2,3,4,5] *Department of Computer Science and Engineering, GMR Institute of Technology, Rajam, Andhra Pradesh, India - 535127*

A B S T R A C T

The human generation is facing the threat of dangerous disease COVID-19 from 2019 to till now. Though the effect of disease is reducing now-a-days, complete cure for the disease has not been found till now. Detection of such dangerous diseases should be done in a faster way for effective diagnosis and better results. This disease can be detected through many ways like antigen and RT-PCR test which were proven to be effective but they are time consuming. Researchers have been found that this disease can also be detected through the cough sound of the person. Detection of COVID-19 from the cough audio sample can be done in a very rapid and reliable way using various machine learning and deep learning techniques. Audio sample is pre-processed first and various useful features can be extracted based on the frequency patterns of the audio sample. This frequency patterns can be represented in an image called spectrograms which represents the audio sample. These images are very effective and these can be used as inputs for training on the deep learning models and features can be extracted from images by using various pre-trained models. After deep learning algorithms are applied and performance is measured using various performance metrics. This project is on applying the above-mentioned techniques for detection of Covid in the given audio sample.

Keywords: Deep Learning, spectrograms, cough audio samples, feature extraction, neural representation learning.

1. Introduction

By the time of September, 2022, there have been 62 crores confirmed cases of COVID-19, pandemic caused by the coronavirus SARS-CoV-2 reported to World Health Organization (WHO) the severity of the infection and the associated fatality rates around the world are increasing at an alarming rate. This virus gives rise to serious respiratory infections with really high death rate and creates serious threats to people. In fact, common symptoms of COVID-19 include headache, fever, cough, and fatigue. Other symptoms include shortness of breath, muscle and sore throat, and sometimes diarrhea and vomiting. Among these symptoms, shortness of breath and dry cough caused by respiratory failure are the major causes of death. Unfortunately, Coronavirus sometimes appears to affect other essential organs, namely the heart, brain and lungs.

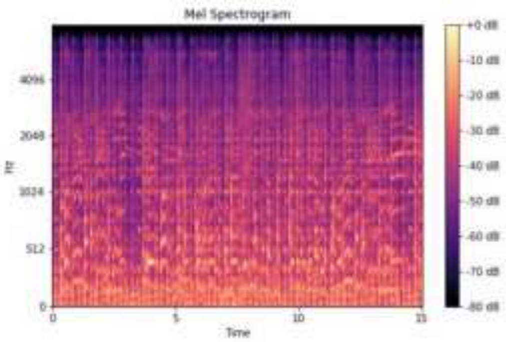

Fig 1. Audio to image representation

Considering the cough sound, basically the (audio file) as the key element and applying Deep Learning algorithms COVID19 has been detected by only on the basis of cough sound. This project has taken the Coswara Dataset as the ideal dataset to work with this project. The audio file has contained both positive and negative audio samples (1000 audio files) are taken into consideration and worked upon. A pre-processed audio sample is used to extract a variety of relevant information based on its frequency patterns. A spectrogram, a visual representation of the audio sample, can show these frequency patterns. These photographs are very useful and can be used as training data for deep learning models. Various pre-trained models can be used to extract characteristics from the images. Performance is assessed using a variety of performance metrics following the application of machine learning algorithms. This project focuses on using the aforementioned methods to find Covid in the provided audio clip. Now, the application of the Deep learning algorithms is done to find out how accurately the model can predict the audio sample to detect Covid-19 disease.

2. Literature Survey

Here we have gathered several periodicals that have conducted research on our connected work, which is based on the rice plant, and we have separately summarized each work as shown below.

A considerable amount of literature has been published on a machine learning strategy that combines signal processing and noise removal techniques with an ensemble of tuned deep learning networks to swiftly detect COVID-19 using audio recordings made on consumer devices, enabling COVID identification on coughs. These three sizable cough datasets with COVID-19 positive samples were publically available: the EPFL COUGHVID dataset, Coswara, and Covid19-Cough. The datasets used in this study can be divided into three types: prior crowdsourcing, curated clinical, and newly gathered data. The research employed a modified light-weight CNN architecture that added a second channel for log-frequency positional encoding to the Mel-spectrogram image-like input. researchers also used an ensemble of DCNN and a bagged ensemble of Gradient Boosted Trees and 10-fold cross-validation using publicly available datasets.A comparative study by Cough recordings were used to measure the performance of CNN, GB, and Ensembles (Variants I and II), whereas data from breath and vocalization were used to measure the performance of GB-breath and GB-vocalization, respectively. And found that CNN and GB cough sub-ensemble setup, with higher weights allocated to them, produces the best results.Detailed examination of covid outlines the preprocessing and diagnosis pipeline implementation in a mobile application created for quick COVID screening. The majority of respondents felt that a cough will be heard in a particular recording is represented by the distribution of model outputs, with 9.8% of recordings being rejected as lacking a cough. To give users who would record with the odd coughs some wiggle room, the threshold for the model was set at 0.25 accept.[1]

The approach of the hybrid architecture for COVID-19 identification and diagnosis using several ML algorithms from cough audio signals is proposed in this research. The updates to the prior paper are included in this one. They employed a small dataset in the earlier articles, which makes overfitting an issue.The issue has grown in importance in light of recent diagnosis tool for Covid19 based on respiratory, cough, and speech sounds is created using Coswara-data from the Indian Institute of Science (IISc) Bangalore.Logistic Regression, Linear Discriminant Analysis, K-Nearest Neighbors, Gaussian Naive Bayes, and Support Vector Machine are kind of machine learning approaches that were used in this survey.The best strategies for the management With characteristics like the Fitness function and Crossover approaches, the Genetic Algorithm (GA), a subclass of evolutionary algorithms, is utilised to produce superior solutions to optimization and search issues.The machine learning classifiers' performance metrics that were performed are accuracy, recall, and F-measure.The GA-ML technique helps the suggested framework attain accuracy of 92.19% for LR, 94.32% for the LDA, 92.19% for CART, 91.48% for the NB, and 93.61% for SVM respectively.[2]

It has been demonstrated that inorder to classify COVID-19 and other respiratory sound disease symptoms including those of asthma, pertussis, and bronchitis, a lightweight convolutional neural network (CNN) with modified-mel-frequency cepstral coefficient (M-MFCC) is proposed in this research.The developments are made by using the pulmonary COVID-19 sounds crowdsourced dataset, which comprises of 256 audio files for feature analysis, was obtained from Cambridge University with mutual consent for research purposes. This collection of 256 audio sound features came from about fifty COVID-19-infected patients.The main target is to extract detailed features from human respiratory sound data, the model has been created using two feature extraction techniques, namely MMFCC (Modified Multifrequency Cepstral Coefficients) and EGFCC (Enhanced Gamma-tone Frequency Cepstral Coefficients).By the help of the improved gamma-tone frequency filter banks and modified Mel-frequency feature extraction channels, the lightweight CNN categorises the human respiratory audio sounds based on the various respiratory diseases (Normal Flu, Asthma, Pertussis, Negative_ COVID-19, Positive_ COVID-19, Bronchitis, and Healthy human respiratory sound features).Based on audio sound metrics like frequency, loudness, air volume, and cough peak flow rate, the model detected deep features.It has been conclusively shown that Accuracy and F1-score are the performance indicators. The proposed model has a class COVID f1-score of 93.48 and an accuracy of 92.32.[3]

Several attempts have been made for creating the classification model, a deep learning-based architecture is proposed that combines long-short term memory, convolution neural networks, and an attention mechanism.The architecture is made up of several elements. Recordings from the COUGHVID audio dataset were first sent to a pre-processing module. Then, to increase the training set and address the issue of class imbalance, two levels of data augmentation have been applied to both the audio signal and the spectral data (Mel-spectogram augmentation).This paper attempts to show that inorder to produce binary classification results, the resulting Mel-spectogram features are input to a fresh attention-based hybrid CNN-LTSM model.A small number of those interviewed indicated that the large-scale, publicly-accessible COUGHVID crowdsourcing dataset from the Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland, contains 27550 recordings, including 1155 positive cases.Three blocks make up the CNN-LSTM architecture. The first block makes use of a CNN architecture, which takes form (39883)-added mel-spectrograms as input. The convolution layers then extract the features that are the most pertinent and instructive. The deep features with a high temporal correlation are chosen to be provided to the attention block in the second block in order to capture more valuable patterns. Attention-LSTM feature maps are passed to the LSTM block in the first block. A straightforward fully linked layer is utilised for feature learning and classification in the third block.Accuracy, sensitivity, precision, and f1-score are the performance measures. The suggested model achieved 91.13% testing accuracy and 90.93% sensitivity, which is superior than LSTM and CNN individually[4]

This paper contests the claim that, researchers look into the use of deep learning models as a widely used, low-cost pre-testing method for COVID-19 identification using audio recordings of breathing or coughing made online or with mobile devices.The database, which consists of 1427 audio clips totaling 3.93 hours and was compiled by the University of Cambridge utilising an online form and the "COVID-19 sounds" Android app, was used in the experiments.This account seeks to Adapt a group of convolutional neural networks to classify whether a speaker is COVID-19-infected or not using spectrograms, raw audio of breathing and coughing, and other inputs.Raw , a mixture of spectrograms with hop lengths of S = 8 ms and S = 16 ms, and audio are the three input types that are employed. N convolution blocks are utilised for each of them, proceeded by a time reduction block, feature concatenation, and fully connected layers for classification.The proposed models' best results on the held-out speaker independent test partition were a

UAR score of 74.9% and an AUC score of 80.7%.The results are being limited by the quantity of data that is currently available, which might make it impossible to use even larger models, which is where deep learning models frequently excel.[5]

This paper presents a variety of models that performed admirably when put to the test against the largest evaluation dataset currently available in scholarly literature.Data from several studies have identified to ensure minimal bias during model training. IATOS, COUGHVID, and Coswara, are three publically available datasets that were integrated. 20,072 cough recordings are included in COUGHVID, and over 2000 cough samples were tagged in the Coswara dataset after it was processed by Virufy.The two main processes that the cough audio data goes through to provide easily usable data for model construction are standardisation and preprocessing.Three models such as the support vector machine (SVM), the self-supervised learning with transformers (SSL), and the convolution neural networks(CNN) are compared in this study Support vector machines (SVMs) are used to classify data utilising handcrafted features such as spectral centroid, spectral roll-off, root mean square, Mel-frequency cepstral coefficients (MFCC), delta-MFCC, and double delta MFCC (SVM).Researchers assess how well a self-supervised learning (SSL) and CNN technique compare to a standard SVM model. As anticipated, we find that the SSL and CNN models are better at differentiating cough signals than the simpler SVM model.AUC was 0.75, sensitivity was 0.14, and specificity was 0.96 for the SVM model. With a specificity of 0.51; a sensitivity of 0.89; and an AUC of 0.807 on the validation data, the self-supervised learning model produced findings that were significantly more balanced. With a specificity of 0.74, sensitivity of 0.77, and AUC of 0.802, the CNN also performed comparably.[6]

The aim of this study is to develop a method for automatically detecting cough during audio recordings of talks in order to diagnose COVID-19.The five key components of the technique are COVID-19 diagnosis, sound feature extraction, cough detection, and classification.SVM, KNN, and RNN are used as the machine learning and deep learning models which can successfully identify COVID-19 from audio recordings.Data from two open-source sources were used for feature extraction and model training. The first database is Coswara, which currently offers open source data from a total of 1433 individuals. The second dataset is called Virufy, and it is very accurate as the data was collected in a medical setting under the direction of professionals who were aware of patient permission and who followed standard operating procedures (SOP).To extract the sound activity in conversational recordings, the silent sections must first be removed. In order to extract coughs from the discussion, several signal characteristics must first be collected, grouped using a KNN model, and then an RNN model trained to identify if the cough is brought on by COVID-19 must be employed.The recurrent neural network (RNN) model excels in order to achieve best results. Despite the fact that the test set and training set come from different databases, it nevertheless achieves an accuracy of 81.25 percent (AUC) of 0.79.[7]

This study discusses a DNN-based investigation that uses cough sound samples to identify COVID-19 patients early.The proposed approach consists of three main steps.The first is removing acoustic features from the cough sound samples, second is constructing a feature vector, and third is building the feature vector, using a deep neural network.The existing research uses Three different kinds of acoustic feature vectors are considered in this study: (a) time-domain, (b) frequency-domain, and (c) mixed-domain (i.e., a combination of features in both time-domain and frequency-domain).The Virufy database's cough sound samples are utilised in this work. The database includes the crowdsourced and clinical data.The suggested system is implemented using a modified version of the DNN model . Three hidden layers make up the DNN employed in the network. 20 nodes composed in each hidden layer and as well as 500 input nodes for the matrix input are present in the network. Using the frequency-domain feature vector, DNN achieves training accuracy of 100%, validation accuracy of 98.50%, and testing accuracy of 97.50%. It can be said that the frequency-domain feature vector has a greater testing accuracy than the time-domain feature vector.[8]

In this analysis, the importance of speech signals is focused. Early COVID-19 virus detection and diagnostics are being handled by the Recurrent Neural Network but the use of LSTM also comes into the picture which deals with the acoustic features of the patients' voices, and respiration. Speech samples were taken from 20 COVID-19 patients and 60 healthy speakers (40 men and 20 women) (12 male and 8 female). Every participant was required to record a sample of their voice, breathing, and coughing sounds.Using the programme PRAAT, researchers manually preprocessed (clearing of silence) all user samples.The RNN is mostly used to predict a future data sequence using samples from the past data, and LSTM is continued because the gradient is disappearing.It has been shown that Spectral Centroid (SC), Spectral Roll-off (SR), Zero-Crossing Rate (ZCR), Mel-Frequency Cepstral Coefficients (MFCC), and MFCC and 2 MFCC are characteristics that have been extracted.Reasearchers uses voices, breath noises, and cough sounds to achieve the results. In which the comparison of these three results reveals that breathing sound has the highest accuracy when compared to the other two noises, reaching up to 98.2%.The main weakness of this study is that the amount of data gathered is rather little, and neither healthy individuals nor other patients with various respiratory ailments make up the control group.[9]

In this work, researchers present a novel, fully automated algorithm framework for COVID-19 detection and cough extraction. This framework combines sound processing and machine learning techniques.The audio recordings from the Biovitals® Sentinel smartphone app, which included real coughs from confirmed COVID-19 patients as well as widely accessible public datasets of different sound recordings, were used to evaluate the algorithm performances. To enhance the functionality and applicability of the model, the Virufy and Coswara datasets were also added.The CNN framework is chosen in the cough detection model since it has been extensively used in the literature for audio signal processing.The design of the questionnaire was performed by the "spectral gating" strategy for audio noise reduction which uses an acoustic noise reduction method.The collected audio segments of different lengths are transformed into Mel-frequency cepstral coefficients to create the input samples for this cough detection classifier (MFCCs).The majority of respondents felt that this algorithm is capable of managing crappy audio captured using a smartphone in a noisy environment.The strong evidence of the proposed method completes the COVID19 classification task with a mean cross-validation AUROC of 98.1% and the cough extraction task with a duration Area Under Receiver Operating Characteristic curve (AUROC) of 98.6%.[10]

In this paper, researchers suggested a machine learning-based method for separating COVID-19 patients from non-COVID-19 patients by evaluating just one cough sound.The open cough data set which was collected has a total of 121 single cough records in the virufy COVID-19 and 73 single cough sounds in the NoCoCoDa.Using the MFCCs approach, the features were extracted from cough noises and categorised using seven different classifiers. The accuracy rate measure was obtained using the LOO-CV technique in order to choose values for the hyperparameters used in the feature extraction and classification operations.The detailed examination of each classifier performed a unique feature selection procedure based on the SFS approach.The seven distinct classifiers that were employed are Polynomial-SVM, RBF-SVM, Linear-LDA, Quadratic-LDA, Euclidean-kNN, Chebychev-kNN, and PLSR. Results revealed that the Euclidean-kNN classifier, which had an accuracy rating of 0.9833, had successfully completed this process.[11]

In a randomised controlled study on the use of signal processing and audio recording on cellphones. This research was based on the detection of COVID-19. One benefit of these studies is that they may be carried out anywhere, at any time, and with minimal effort.The dataset was gathered in a hospital or clinic under the direction of doctors or other medical staff. The cough sound data that was mentioned was recorded using the "Virufy" mobile app at Stanford University requests and made accessible on "GitHub..“Each variety of subject's cough time was estimated to be 1640 milliseconds (ms) in a data pool of 121 segmentations, and the study's sampling frequency was set at 48,000. To increase the reliability of measures two methods, STFT and mel-frequency cepstral coefficients, were employed as feature extraction strategies (MFCC).Inorder to recognise and categorise COVID-19 cough, the support vector machine (SVM) method was applied to the analysed signals.The approach with the highest success rate was determined to be the combination of the RBF kernel SVM with the MFCC feature extraction algorithm.[12]

In order to address the coronavirus infection crisis and the shortage of medical experts plaguing healthcare systems, this study proposes a deep learning-based method for the differential diagnosis of coronavirus disease that may be used in clinics. In this work, we combine a recurrent network with a convolutional neural network as an encoder and an attention mechanism.From April 2020 to October 2020, the information was gathered in hospitals in Russia, Belarus, and Kazakhstan using the Acoustery mobile application. A data record of about 3000 coughing samples were taken.The recommended model is attention mechanism which is used for analyzing the significance of an audio recording's fragment and its contribution to the coronavirus infection diagnosis.The classification accuracy of the model is 85%. Precision and recall measures had respective values of 78.5% and 73%.It has been conclusively shown that there is an use of two multiclass classifiers by the researchers. One of them employs Mel spectrograms as an input, while the other uses feature extraction based on MFCC (Mel Frequency Cepstral Coefficients) and PCA (Principal Component Analysis).The Strong evidence of the proposed RCNN model over the aforementioned technique is that it uses spectrograms of records as input data, as opposed to multidimensional matrices of frequency features[13]

This work was designed as an experimental investigation into the effectiveness of bottleneck feature extraction and transfer learning in identifying COVID-19 from audio recordings of speech, breath, and cough. This kind of screening can be used on low-cost consumer devices, like a smartphone, and is non-contact.The datasets consists of cough, sneeze, speech, and other noises but lack COVID-19 records. the areas under the receiver operating characteristic (ROC AUC) of 0.98, 0.94, and 0.92 of coughs, breaths, and speech are present .The three deep neural network models used in this study for pre-training are the CNN, LSTM, and Resnet50.The majority of respondents shows that a Resnet50 classifier trained using this transfer learning method performs optimally or almost optimally across all datasets, obtaining areas under the receiver operating characteristics.The COVID-19 cough, breath, and speech audio categorization can be improved with deep transfer learning and bottleneck feature extraction, leading to automatic COVID-19 detection with a better and more reliable overall performance. Results also demonstrate that using transfer learning and extracting bottleneck features increased performance, but also helped identify bottleneck features using the bigger datasets without COVID-19 labels.[14]

In this study, they compared the diagnostic performance of chest CT with the industry gold standard RT-PCR, conducting a systematic analysis of chest (CT) images to understand the changes in the lung during COVID-19 recovery as well as by the chest CT features of COVID-19 patients.Inorder to segment abnormalities in COVID-19 Chest CT images.Anam-Net, a suggested lightweight CNN based on anamorphic depth embedding is proposed.The study set out to Anam-Net's advantage over other models is its low computational complexity, which requires 50% less training time and 7 times less parameters than the next-best performing network.This study undergoes, a quick, fully automated method for segmenting abnormal lung tissue and normal lung tissue in chest CT images of COVID-19 patients with anomalies is proposed.The suggested approach successfully segmented the anomalies in all COVID19 chest CT images, and Anam-Net performed better in terms of sensitivity (sensitivity is 97% in this case), precision, accuracy, and Dice score.Therefore, the suggested Anam-Net is simply deployable in mobile platforms to give a quick assessment of the abnormalities in COVID-19 chest CT images because the deployment in mobile as well as embedded hardware platforms confirmed that it is well suited for the point of care settings.[15]

This study looks into the feasibility of determining the severity of the Covid-19 infection in persons who test positive for the virus, in addition to developing a test for Covid-19 recognition using cough sounds. In this work, the performance of a primary Covid-19 screening technique based purely on the cough sound is evaluated using 8,380 clinically validated samples and laboratory molecular tests (2,339 Covid-19 positive and 6,041 Covid-19 negative) from accredited laboratories.The suggested generic method is an Empirical Mode Decomposition (EMD)-based algorithm for cough sound recognition with subsequent classification based on a tensor of audio sonograms and a deep artificial neural network classifier with convolutional layers named "DeepCough."Three different sonographs were employed by researchers: Based on the audio recordings, the frequencies are extracted using 1) Mel-frequency Cepstral Coefficients (MFCCs), 2) Mel-scaled spectrogram (MelSpec), and 3) Linear Predictive Coding Spectrum (LPCS) coefficients.The article focuses mostly on DeepCough3D's Statistical Performance Measures. When used with DeepCough2D, AutoML is a comprehensive model meta-learning system that combines Bayesian optimisation in a series of s. In this review, the detection performance of a primary Covid-19 screening tool based only on the cough sound is evaluated using 8,380 clinically validated samples with laboratory molecular testing (2,339 Covid-19 positive and 6,041 Covid-19 negative) under quantitative RT-PCR (qRT-PCR) from certified laboratories.DeepCough3D's performance achieves an AUC of 98.80 0.83 and a sensitivity of 96:43% 1:85% in comparison to the claimed sensitivity (91% 10%) of accelerated serology tests in the view of saliva.[16]

3. Relevant Study

Analysis of different techniques to build a system for effective detection of COVID-19 from audio using cough audio samples by using following images:

3.1 Spectrogram

A spectrogram displays the strength of a signal over time at a waveform's various frequencies. Spectrograms can be two-dimensional graphs with a third variable represented by colors or three-dimensional graphs with a fourth color variable. To generate a spectrogram, a time-domain signal is divided into shorter segments of equal length. Then, the fast Fourier transform (FFT) is applied to each segment. The spectrogram is a plot of the spectrum on each segment. The Frame Count parameter determines the number of FFTs used to create the spectrogram and, as a result, the amount of the overall time signal that is split into independent FFTs.

3.2 Mel Spectrogram

The Mel Scale is a logarithmic transformation of a signal's frequency. The core idea of this transformation is that sounds of equal distance on the Mel Scale are perceived to be of equal distance to humans. It is actually much harder for humans to be able to differentiate between higher frequencies, and easier for lower frequencies. So, even though the distance between the two sets of sounds are the same, our perception of the distance is not. Mel Spectrograms are spectrograms that visualize sounds on the Mel scale as opposed to the frequency domain

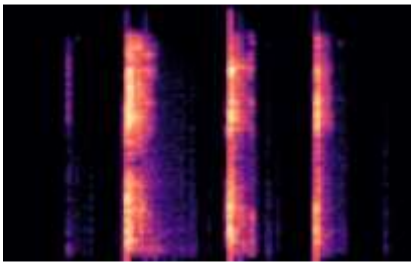

Fig 2. Mel Spectrogram

3.3 Data Acquisition

Data is the primary element to perform any kind of analysis for a particular problem. The audio files dataset which consists of cough, breath and speech audio samples were available open source and are widely used by many researches. Few of the datasets which are mainly used are Coswara, Virufy, COUGHVID etc. These datasets were considered more and were proven to be performed with good analysis.

3.4 Audio Processing

Audio file is loaded into librosa as a float point time series. Audio will be automatically resampled to the given rate i.e 22050. This is called sampling rate.it means number of samples taken per second. Generally, sampling rate is considered high values because the more number of samples in a sound, the more accurate is the digital representation. It will give the one-dimension floating point array of audio time series with the considered sample rate. The length of the output array depends upon the length of the audio file (duration) considered while loading the sample. Later this output array and sample rates are used to produce a graphical visual spectrums.

3.5 Deep Learning

Artificial intelligence is evolving and expanding it's art of learning to all branches of the current technical world. The latest developed computational resources are helping to achieve more efficient results in this AI Technology. Machine learning is a sub-branch of AI, it uses self-learning approach to derive meaningful insights from presented data without manual rules. Deep learning is a type of machine learning that uses neural networks to build more complex models. At the earlier stages of AI they have been limited in terms of computational resources. But with latest advancements we can build more complex and large computational models by using deep learning to process large data effectively. As the name suggests AI Neural networks imitates the human brain nerve mechanism to learn from presented data. Deep learning has many variations based on the given input data such as images, text, numerical etc. Convolutional Neural Networks are used for processing image data and their usage has been applied to solve many computer vision problems. CNN process the input data by applying many convolution functions and filters to extract features from the given input images. Most of the research work was performed using CNN models for spectrograms processing. Pre-trained models were considered like VGG-16, ResNet and combination of LSTM-CNN architectures were also considered.

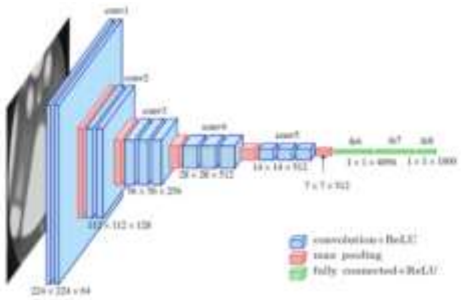

Fig 3. Convolutional Neural Network

Feature extraction was also performed for extracting the useful features from the spectrograms and melspectrograms. Since, the size of spectrograms are large and also the computation of those requires more time and processing units, some fixed number of features were extracted from each image. Each feature is a numerical value hence, these features can be scaled and later provided to machine learning models for analysis which takes less time and processing power compared to image processing. Autoencoders can be used to extract the features. It consists of a encoder and decoder architecture and a bottle-neck layer is present in between them. We can get the features at the bottle-neck layer which is the compressed representation of image information in a numerical form.

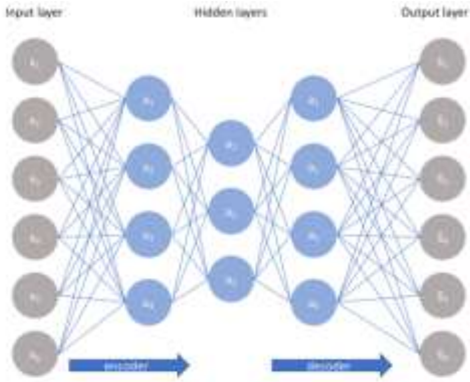

Fig 4. Autoencoders

4. Conclusion and Future scope:

Though the spreading of COVID-19 disease is reducing now-a-days, there are high chances for spreading of disease again in future days. Dry cough is identified as a major symptom for identification of COVID-19 and it is really helpful for diagnosing of disease. Machine learning and deep learning techniques can be applied for faster identification of disease compared to existing chemical methods. Most of the used models are convolutional neural networks and combination of LSTM with them. Pre-trained models were considered in analyzing the image data. Since, the size of current available datasets are small, the performance can be further increased since the dataset is constantly updating. These models can also be applied to other diseases other than COVID which can be diagnosed using audio or cough samples. Other digital image representations like spectrograms, short-term fourier transform images can also be applied for analyzing the performance and identification of disease.

References

[1] Ponomarchuk, A., Burenko, I., Malkin, E., Nazarov, I., Kokh, V., Avetisian, M., & Zhukov, L. (2022). Project Achoo: a practical model and application for COVID-19 detection from recordings

[2] 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.

[3] Haritaoglu, E. D., Rasmussen, N., Tan, D. C., Xiao, J., Chaudhari, G., Rajput, A., ... & Pilanci, M. (2022). Using Deep Learning with Large Aggregated Datasets for COVID-19 Classification from Cough. *arXiv preprint arXiv:2201.01669*.

[4] 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

[5] Islam, R., Abdel-Raheem, E., & Tarique, M. (2022). A study of using cough sounds and deep neural networks for the early detection of COVID-19. *Biomedical Engineering Advances*, *3*, 100025.

[6] Hassan, A., Shahin, I., & Alsabek, M. B. (2020, November). Covid-19 detection system using recurrent neural networks. In *2020 International conference on communications, computing, cybersecurity, and informatics (CCCI)* (pp. 1-5). IEEE

[7] Zhang, X., Pettinati, M., Jalali, A., Rajput, K. S., & Selvaraj, N. (2021, November). Novel covid-19 screening using cough recordings of a mobile patient monitoring system. In *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)* (pp. 2353-2357). IEEE.

[8] 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

[9] Melek Manshouri, N. (2022). Identifying COVID-19 by using spectral analysis of cough recordings: a distinctive classification study. *Cognitive Neurodynamics*, *16*(1), 239-253.

[10] Hemdan, E.ED., El-Shafai, W. & Sayed, A. CR19: a framework for preliminary detection of COVID-19 in cough audio signals using machine learning algorithms for automated medical diagnosis applications. *J Ambient Intell Human Comput* (2022).

[11] Al-Dhlan, K.A. An adaptive speech signal processing for COVID-19 detection using deep learning approach. *Int J Speech Technol* 25, 641–649 (2022).

[12] Kranthi Kumar L, Alphonse PJA. COVID-19 disease diagnosis with light-weight CNN using modified MFCC and enhanced GFCC from human respiratory sounds. Eur Phys J Spec Top. 2022 Jan 24:1-18.

[13] Hamdi, S., Oussalah, M., Moussaoui, A. *et al.* Attention-based hybrid CNN-LSTM and spectral data augmentation for COVID-19 diagnosis from cough sound. *J Intell Inf Syst* (2022).

[14] Mitrofanova, A. Y., Mikhaylov, D., Shaznaev, I., Chumanskaia, V., & Saveliev, V. (2021). Acoustery System for Differential Diagnosing of Coronavirus COVID-19 Disease. *IEEE Open Journal of Engineering in Medicine and Biology*, *2*, 299-303.

[15] Pahar, M., Klopper, M., Warren, R., & Niesler, T. (2022). COVID-19 detection in cough, breath and speech using deep transfer learning and bottleneck features. *Computers in Biology and Medicine*, *141*, 105153.

[16] Paluru, N., Dayal, A., Jenssen, H. B., Sakinis, T., Cenkeramaddi, L. R., Prakash, J., & Yalavarthy, P. K. (2021). Anam-Net: Anamorphic depth embedding-based lightweight CNN for segmentation of anomalies in COVID-19 chest CT images. *IEEE Transactions on Neural Networks and Learning Systems*, *32*(3), 932-946.