Diagnosis of COVID-19 from CT Images and Respiratory Sound Signals Using Deep Learning Strategies

S. Maheswaran, G. Sivapriya, P. Gowri, N. Indhumathi, and R. D. Gomathi

Introduction

Coronavirus was first originated in Wuhan, China, in 2019. A 55-year-old man from China's Hubei region may have been the first to get COVID-19, a sickness caused by a novel coronavirus that is sweeping the globe. According to the South China Morning Post, the case dates back to November 17, 2019. That's more than a month sooner than physicians in Wuhan, China's Hubei province, reported cases at the end of December 2019. Authorities thought the illness was spread by something sold at a municipal wet market at the time. However, it is now obvious that during the early stages of what is becoming a pandemic, some affected persons had no access to the market. This includes one of the first instances, which occurred on December 1, 2019, in a person who had no connection to the seafood market, according to researchers who published their findings in the journal *The Lancet* on January 20.

Symptoms of COVID

The most frequent symptoms are:

- Fever and cough.
- Tiredness
- Loss of flavor or odor

   Symptoms that are less common:

- Throat pain

S. Maheswaran (✉) · G. Sivapriya · P. Gowri · N. Indhumathi
Department of Electronics and Communication Engineering, Kongu Engineering College, Erode, TN, India

R. D. Gomathi
Department of English, Kongu Engineering College, Erode, TN, India

- Headache
- Pains and aches
- Diarrhea
- A cutaneous rash or discoloration on the fingers or toes
- Eyes that are red colored or inflamed

  Severe symptoms:

- Breathing difficulties
- Discomfort in the chest

Economical Impact on the World

The global economy has been impacted in a variety of ways since the COVID-19 epidemic began in March 2020. Poorer nations have suffered the most, while wealthier ones, despite their superior resources, have experienced their own issues. This article examines the influence of COVID-19 in various parts of the world.

First is dividing 171 countries into three groups based on per capita income: low, middle, and high income. Second is looking at health statistics to see how badly these countries were struck by the virus. Then, third is by comparing the International Monetary Fund's (IMF) pre-pandemic economic expectations for 2020 with their actual values and generated estimates for the pandemic's influence on growth and important economic policy variables [1].
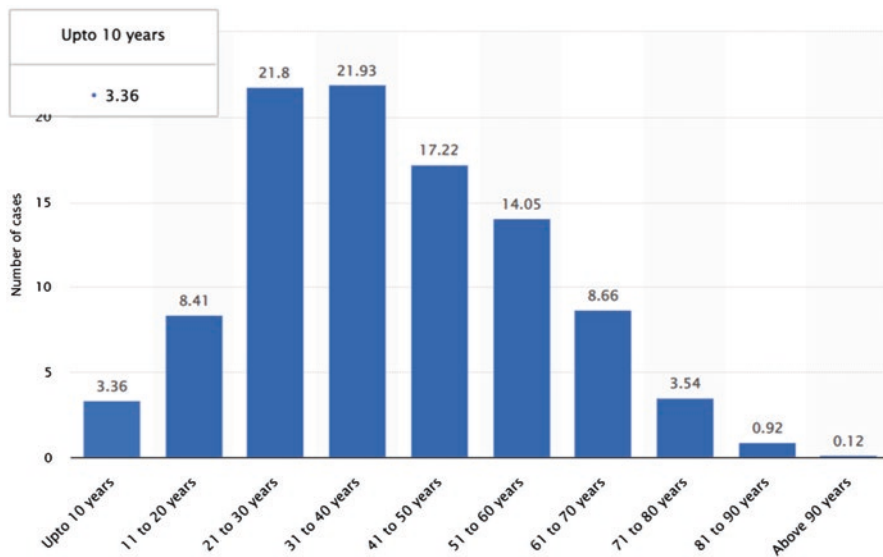

Fig. 1 COVID-19-positive cases in India with age group

The low- and high-income categories each account for 25% of the world's countries, while the middle-income group accounts for 50%. In 2019, the average income per capita in the middle-income category was more than five times that of the low-income group. It was over 20 times higher in high-income countries. Figure 1 shows the COVID-19-positive cases in India with age group.

Economical Impact of India

India has been devastated by the pandemic, particularly during the virus's second wave in the spring of 2021. Despite the fact that the huge drop in GDP is the greatest in the country's history, the economic harm sustained by the poorest households may be understated. From April to June 2020, India's GDP plummeted by a stunning 24.4%. The economy declined by 7.4 in the second quarter of the 2020/2021 fiscal year, according to the most current national income predictions (July to September 2020).

The third and fourth quarters of the recovery (October 2020 to March 2021) were still slow, with GDP increasing by 0.5 and 1.6%, respectively. This indicates that India's total rate of contraction for the fiscal year 2020/2021 was 7.3%. Only four times since independence has India's national GDP declined before 2020 – in 1958, 1966, 1973, and 1980, with the 1980 loss being the most significant. This predicts that the 2020/2021 fiscal year will be the worst in the country's history in terms of economic shrinkage, significantly worse than the worldwide average. The details are presented in Fig. 2. The reduction is principally responsible for reversing the global inequality trend, which had been dropping for three decades but has recently begun to rise again [2–4].
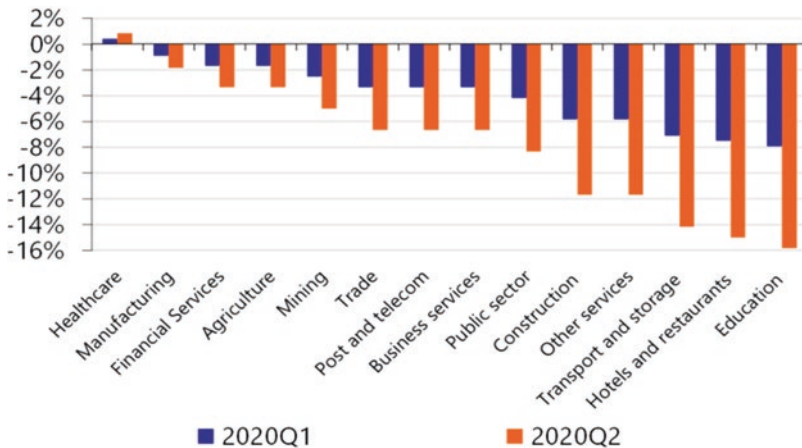

Fig. 2 Economy in different countries

Some Positives of COVID-19

Permits for a proper work and life balance: Many firms, located in large urban areas, and people were historically burdened with long travels, tiresome traffics, or long lines for public transportation. Employees working at home, particularly those with whole families quarantined, get the option to spend time with family members.

Budget-friendly, with improved cost management: Industries can shift their budgets and save money on running expenditures. Expenses are being considered in more realistic and business-oriented manner, notwithstanding the global financial disaster. Workplace supplies, equipment, and leasing costs might be used to aid the existing monetary condition.

Productivity and attention are increasing: Despite the fact that there are distractions at home, the survey concluded that overall attention at work has improved. Stress is minimized when external irritants are eliminated, and projects are finished more resourcefully and correctly. Reduced pressure levels as a result of fewer external irritants permit activities to be completed faster and with fewer mistakes.

Some Drawbacks of Covid

Quarantining: Even while several IT personnel characterize themselves as introverts, the research discovered that they are not truly introverted. There has been no brighter spotlight on this than during quarantine and seclusion. Isolation has had a significant impact on the emotions of employees and executives, and watercooler chatter has proven necessary, even for individuals who self-describe or are thought to be unsociable.

Incompatible with a home office: Not everyone has a great workspace in which to work remotely, and it can be difficult to create an efficient and well-organized workspace. They are unable to create an environment favorable to complete focus, which has a direct impact on staff efficiency. When it comes to production, no man is an island. Autonomy does not work for everybody, and not everybody is capable of professionally self-organizing their job [5].

Electronic communications are prone to being misrepresented: Aside from missing direct physical discussions and meetings, all of those polled were concerned about the essential digital communication methods, such as e-mail, social media, and the possibility for communication gaps. Figure 3 shows the total cases and deaths till 2021.
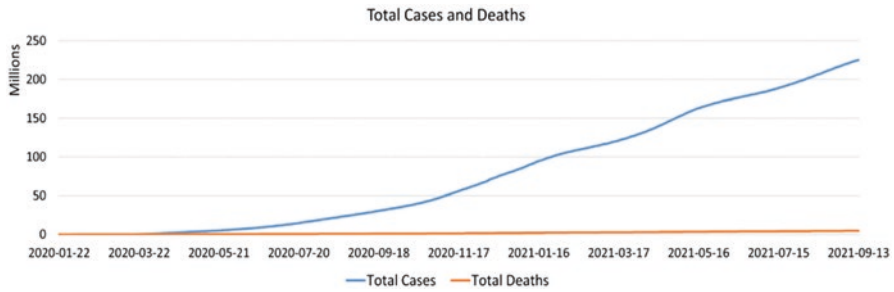
Total Cases and Deaths

Fig. 3 Total cases and deaths till 2021

Impact on Students

The worldwide lockdowns caused by the global pandemic have had a negative impact on several crucial areas, one of which is education. Because of the epidemic, schools, colleges, and universities were forced to close unexpectedly, exposing pupils to online learning. Changes in classroom to digital learning during the pandemic disrupted the learning of students in low-income regions all across the world. Families that couldn't afford smartphones, Wi-Fi, computers, or laptops had undergone serious depression.

The parents, many of whom lacked the necessary qualifications to become home educators, were forced to take on the task on the spur of the moment. The three primary hurdles to online learning for schoolchildren are poor Internet access, insufficient data, and a lack of resources. The major concern here is whether kids are genuinely learning anything when the process shifts to digital learning and virtual classrooms. Changes in learning techniques have recommended us to shift to approaches that have never been used before, allowing us the opportunity to alter and gain exposure. Initially, the institutions were perplexed, since they had no notion how to continue, but as they established the digital infrastructure, the study pattern began to settle. During the epidemic, most students favored open and distant learning modes, because they foster self-learning and provide opportunity to learn from varied resources as well as personalized learning according to their requirements. Because of the epidemic, most recruiting efforts were limited. Students' placements were also impacted. Many graduating students missed crucial career prospects, and many students and professionals were forced to return home from abroad because to the epidemic, interrupting their work.

This chapter gives detailed description on the chosen datasets, preprocessing techniques adopted, models used to train the network, and different classification method. Finally, different classifiers are compared in terms of various performance measures.

Dataset Description

COVID-CT dataset contains of 349 CT images, which are reported as COVID-19 positive, and these images are collected from 216 patients with different age groups. The non-COVID-CT scan images were considered as negative samples, which are taken from the MedPix database, the LUNA database, and the Radiopaedia website.

SARS-CoV-2 CT scan gives a collection of large datasets with 1252 CT images taken from COVID-19-positive tested people and 1200+ CT images, which are taken from negative tested COVID-19 samples affected with some different pulmonary diseases and shown in Fig. 4. The detailed number of patients with both genders is given in Fig. 2.

Coswara is prepared for the diagnosis of COVID-19 with cough and respiratory sounds and shown in Table 1. There are 1079 healthy subjects and 92 COVID-19-affected subjects available in this dataset. The age group of the subjects included in this dataset is 20–50 years.

Sarcos dataset contains 26 subjects tested negative for COVID-19 and 18 subjects tested positive for COVID-19. This dataset has higher female subjects than the male. The sampling rate for recorded audio sounds is 44.1 kHz.

Proposed Methodology

A new framework was proposed for diagnosing the COVID-19 using CT images and breathing sounds. The entire network is designed to predict the class as normal, COVID-19, bacterial pneumonia, and viral pneumonia using the multiclass classification network MLP. The dataset used for the respiratory sounds are taken from
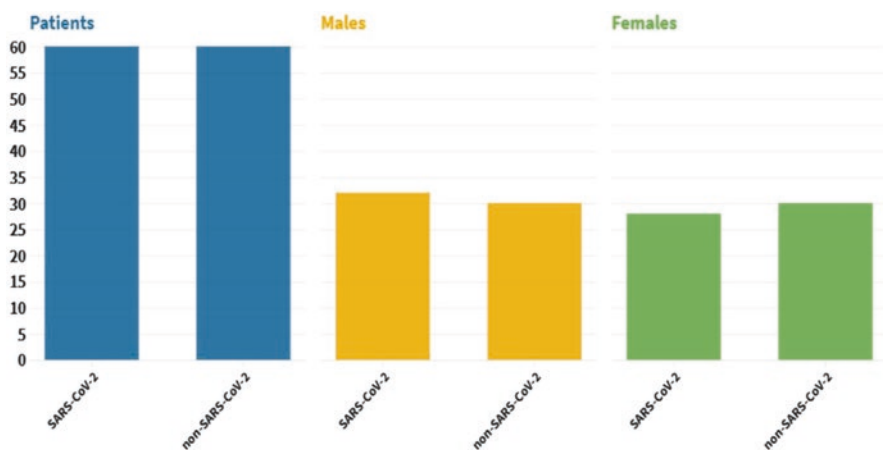

Fig. 4 Number of samples available in SARS-CoV-2 CT-scan

Table 1 Sarcos and Coswara dataset summary

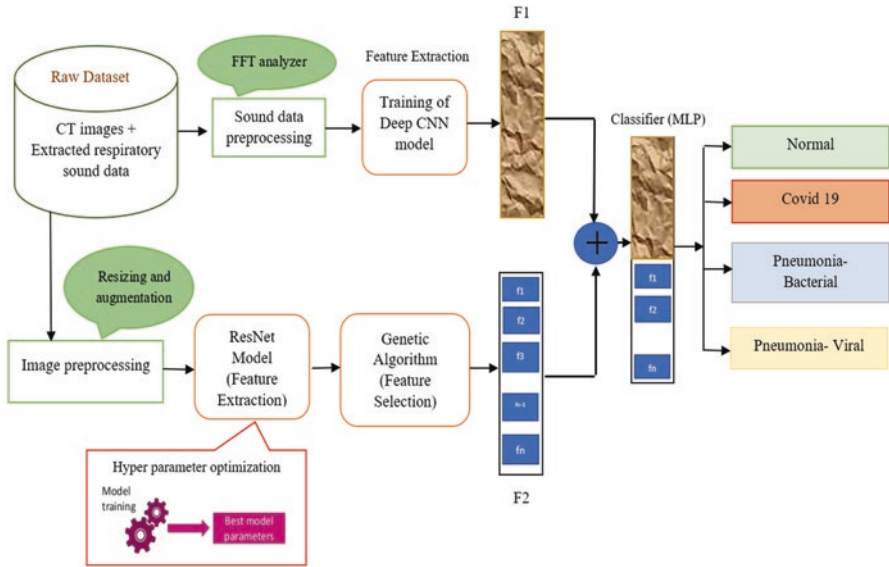| Dataset name	Label type	Size	Audio average per subject	Total recorded audio
Coswara	Positive	92	2.77 s	4.24 min
	Negative	1079	3.26 s	0.98 h
Sarcos	Positive	18	2.91 s	0.87 min
	Negative	26	3.63 s	1.57 min

Fig. 5 Block diagram of proposed model for COVID-19 detection

Coswara dataset and Sarcos dataset containing 92 COVID-19-positive samples, 1079 healthy samples and 19 positive, 26 healthy samples respectively. Features like MFCCs, ZCR, log energies, and Kurtosis are needed to be extracted for identifying dry/wet coughs, variability present in the signal, prevalence of higher amplitudes, and for increasing the performance in audio classification. All these features can be extracted with the deep CNN architecture with the series of convolution, pooling, and ReLU layers. Finally, the classification is done with a multilayer perceptron (MLP) classifier. In parallel to this, the diagnosis of the disease is improved by analyzing the CT images. Two publicly available datasets COVID-CT and SARS-CoV-2 CT scan are used for training the proposed network. Various architectures are proposed and used for the classification of images in the literature, but still ResNet-50 helps in providing the promising results. Figure 5 depicts the overview of the proposed framework.

The proposed framework has two modules: (i) respiratory sound analysis framework and (ii) CT image analysis framework. These modules exhibit the workflow for data gathering, data preprocessing, and the development of the deep learning

model (deep CNN+MLP). In respiratory sound analysis framework, the gathered audio signals are converted to spectrogram video using FFT analyzer. The resulting video signal are used as dataset for the deep CNN model for feature extraction. In image analysis framework, the acquired CT images are preprocessed and used to train the ResNet model for feature extraction. As the number of layers in ResNet is high, the features extracted will be more, and to select the desirable features, the genetic algorithm is opted. Obtained features from both the frameworks are concatenated and fed to the multiclass classifier. This framework can be implemented in real time, to reduce the number of existing expensive confirmatory tests.

Dataset Preprocessing

Preprocessing is first done to enhance the quality of the image, so that it becomes easier for the network to train faster. Noise and other irrelevant information are discarded in this stage and preserving all the essential information. To reduce the computation cost, the input CT image is first resized to 224 × 224 pixels with the help of OpenCv and converted into NumPy array. Histogram equalization is performed next to adjust the intensity level of each pixel in the image. It is used when the images lack the shadows and highlights.

Data Augmentation

Augmentation is preferred to increase the volume of the input dataset. This makes the network to train with the different versions of the image, and it increases the overall model performance. Random transformations, like image zooming, resizing, shearing, shifting, and rotation at different angles, are done to overcome the insufficiencies of the dataset. Vertical and horizontal flips can also be done where the vertical flip is equal to rotating the image by 180°.

Dataset balancing: Positive COVID-19 samples are not sufficient in the Coswara and Sarcos datasets. To create an equal number of samples in the training process, SMOTE technique of data balancing is adopted here, which helps in overcoming the problem of class imbalance. SMOTE helps in providing the better results compared to the extended versions like borderline-SMOTE. Five different COVID-19-positive samples are chosen randomly for every available sample in the dataset, and Euclidean distance is determined and represented as x_{nn}.

Then, the new positive samples are created with the equation below:

$$X_{\text{SMOTE}} = x + u\left(x_{nn} - x\right)$$

where u is the multiplication faction distributed in the range (0,1).

Feature Extraction

Figure 6 shows the different features extracted from the preprocessed respiratory and cough dataset.

MFCC

In MFCC, initially the signal is filtered and applied Fourier transform, then the frequencies are warped in mel spectrum. In the next step logarithmic scale is applied to the mel spectrum output followed by Discrete Cosine Transform is applied. The formula to determine the mel frequency is

$$mel(f) = 2595 x \log(1 + f / 700)$$

Here, $mel(f)$ represents the frequency in mels. The block diagram below represents the overall process carried out in calculating the MFCC. Figure 7 shows the complete process involved in finding MFCC.

The MFCC function is calculated with the formula below:

$$C_n = \sum_{n=1}^{k} (\log S_k) \cos\left[n\left(k - \frac{1}{2}\right)\frac{\pi}{k} \right]$$

The coefficient of mel spectrum is represented by k, output from filter bank is S_k, and C_n represents the output MFCC coefficient.
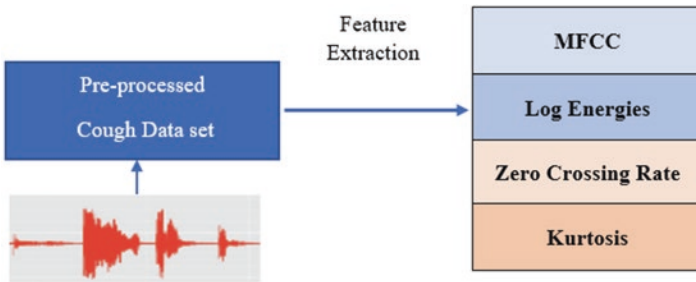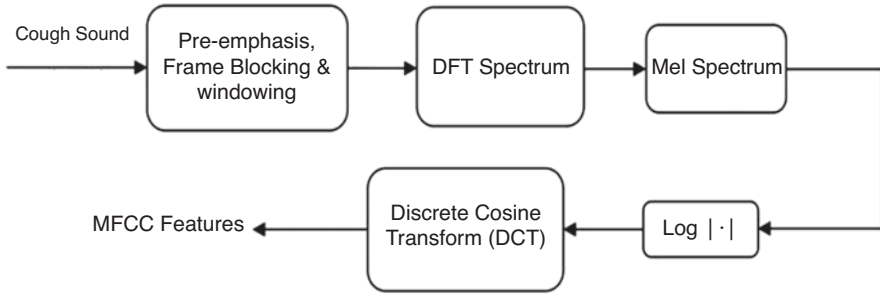

Fig. 6 Feature extraction

Fig. 7 Complete process involved in finding MFCC

Zero-Crossing Rate

The zero-crossing rate (ZCR) is maximum times the wave moves from positive sign to the negative sign. Voice signals oscillation will be very slow like it takes a 100 cross per second for a frequency of 100 Hz and a non-voice signal takes barely 3000 crossing for the same frequency. The zero-crossing rate is calculated with the formula

$$zcr = \frac{1}{T-1} \sum_{t=1}^{T-1} 1_{R<0} \left(s_t s_{t-1} \right)$$

T is the length of the signal s and $1_{R<0}$ is an indicator function.

Kurtosis

It is a signal's fourth-order moment which will measure the peak values associated with the given input audio signals.

$$Kurtosis = \frac{E\left(y_i\left(t \right) - \mu \right)^4}{\sigma^4}$$

ResNet-50 Model

ResNet-50 (residual network) is a modified version of traditional convolutional neural network. It has nearly 50 deep layers with 26 million hyperparameters, and the network was introduced by [5–8].

When the network layer goes deeper increases, the performance of the network may reduce because of the vanishing gradient problem. Vanishing gradient is the main reason for the degradation in the accuracy of the models. Residual networks are introduced to solve this problem with the help of skip connection. Skip connection prevents the gradient from vanishing, and it also makes the higher layers to work better than the lower layer by providing identity connections.
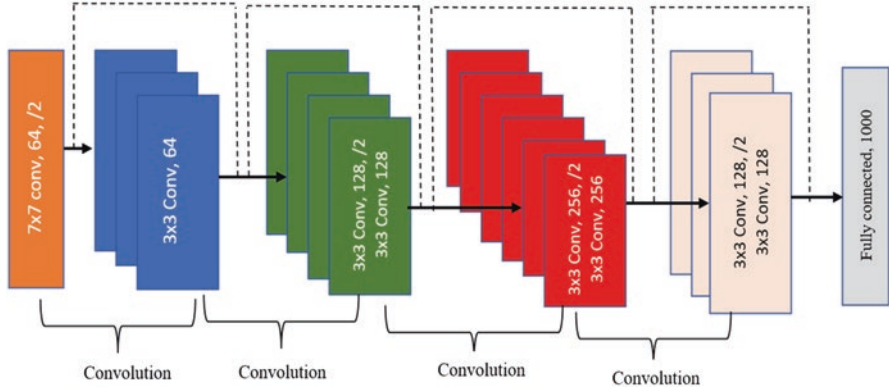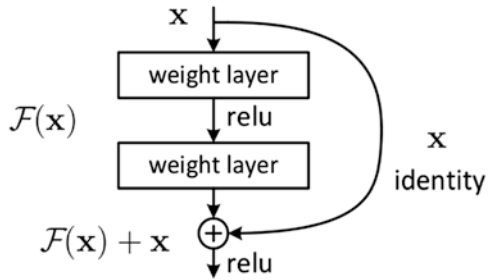
Fig. 8 ResNet-50 model

Fig. 9 Representation of skip connection in residual networks

Figures 8 and 9 shows the residual network with identity connection. It is represented mathematically as $y = F(x, W_i) + x$. Here, y is the output of the residual network, x is the input given to the network, and $F(x, W_i)$ is the residual function. The figure shows the architecture of ResNet-50, which has four blocks with different numbers of convolution layers. The input image of size $224 \times 224 \times 3$ is given to the input layer, and ResNet initially performs 7×7 convolution with 3×3 of kernel size. In the first block, the three convolution operation is carried out, where the kernel size is set as 64, 64, and 128, respectively, for the three layers. The dashed line in the architecture represents the identity connection from one block to the next block. As the stride size is fixed as two in the following blocks, the input's height and width gets reduced to half, whereas the channel width will be doubled [9–11].

Feature Selection

Genetic algorithm (GA) is mostly preferred for feature selection; it is a stochastic method for producing the optimal result with the help of biological evaluation. GA makes use of population by evaluating the fitness function. The population with the

higher fitness values are carried out for the next genetic process, where the remaining solutions will be removed. These chromosomes carried out for the next stage undergo the process of mutation and crossover to generate new populations. This entire process will be carried forward until till it converges to a solution [12]. Fitness is the most important parameter to be considered in genetic algorithm. The solution of the fitness function determines whether the population continues for the next generation process. Deep learning models, like CNN and ResNet, tend to extract most of the features from the input image. Genetic algorithm is used to select the best features extracted from ResNet-50 architecture. GA is preferred because of its evolutionary process through which it can offer the best solution to classify the CT scanned images as positive or negative [13–16].

- For every input CT images, the features are extracted with the help of proposed method.
- Individuals are represented as zero or one in the population; this indicates the presence of desired attributes in the individual. The size of an individual population depends on the number of features extracted from the image. Figure 10 shows an example for selecting an individual.

Roulette method [8] is used for selecting the parent pair, which involves in the mating process and produces the next new generation. One-point crossover (A. H. Wright) is used for selecting a crossover point for each parent pair. The offspring generated first with the genes present at the right side of the first crossover points and so on.

$$x_{[i]} \leftarrow \begin{cases} P1_{[i]} & if \quad i < \gamma \\ P2_{[i]} & if \quad i \geq \gamma < t \end{cases}$$

Here, x is the child vector, i represents the position index between parent and child, t denotes the parent's size, and γ denotes the randomly chosen crossover points.
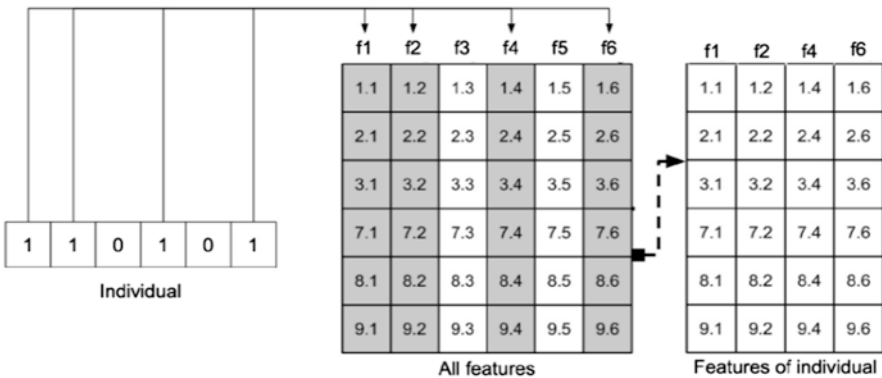

Fig. 10 Creation of individual features

Bitwise Mutation

Bitwise mutation [8] is the mostly preferred operator in case of binary encoding. Each gene is considered individually, and it makes each bit to be inverted within a minimum probability. Elitism is a method used here to generate new population by taking the next individuals from the existing generation. The cycle will continue until it reaches the stopping criteria.

LSTM

Long short-term memory (S. Hochreiter) is introduced for avoiding the problem of vanishing gradient. LSTM contains a cell state to store and covert the input memory cell to an output state. The LSTM architecture shown below has forget, output, input, and update gates. A received information that is sent to the next neuron is decided by the input gate; the information that needs to be forgotten is decided by the forget gate with the help of memory units. These four units interact together and work in a specified manner, where it takes all the long-term inputs and short-term inputs at a given time stamp. The mathematical representation of the input gate is represented as $i_t = \sigma(W_i * [h_{t-1}, x_t] + b_i)$.

The information to be neglected by the forget gate is mathematically given as $f_t = \sigma(W_f * [h_{t-1}, x_t] + b_f)$. The cell states updated with the update gate are represented as $c_t = \tanh(W_c * [h_{t-1}, x_t] + b_c)$, $c_t = f_t * c_{t-1} + i_t * c_t$. The output gate is updated with the equation $o_t = \sigma(W_o * [h_{t-1}, x_t] + b_o)$ and $h_t = o_t * \tanh(c_t)$.

Classification

Classification is a method of grouping given data based on the information found in a dataset, including annotations for which a category has been determined [17]. The CT images were classified as normal, COVID-19, bacterial pneumonia, and viral pneumonia. Cases using several classifiers include a random forest (RF) approach, a multilayer perceptron (MLP), LASSO, elastic net (ENet), a support vector machine (SVM), and an eXtreme Gradient Boosting (XGBoost) algorithm, with the default parameters in each classifier. The classification is verified for the findings using four generally used statistical assessment metrics found in the various literature: Acc (A), recall (R), precision (P), and F-score (F) [18, 19].

Performance Metrics

We can't assert that when a model has a higher level of accuracy, it makes the ideal forecast. The number of true positive (TP), false positive (FP), true negative (TN), and false negative (FN) values defined in the following formula are used to determine the four measures.

$$A = \frac{TP + TN}{TP + TN + FP + FN}$$

$$R = \frac{TP}{TP + FN}$$

$$P = \frac{TP}{TP + FP}$$

$$F = 2 * \frac{R * P}{R + P}$$

Further, the kappa index (K) measures the level of agreement between the proposed methodology's results and the experienced way of ground truth labeling. The area under the receiver operating characteristic (AUROC) curve indicates how successfully the classifier can discriminate between classes based on true positive against false positive numbers. The ratio of number of TP identified to the actual total of TP and FP is measured by the area under the precision-recall curve (AUPRC).

The closer these validation measures' values are to 1, the better the classifier can discriminate between the four distinct class pictures [20]. The suggested model's performance at all categorization thresholds is shown by ROC curve. The area under the ROC curve integrated from (0, 0) to (100, 100) is known as AUC (1, 1). It calculates the total of all potential categorization thresholds. With a range of 0 to 1, AUC is a 100% incorrect categorization. It is appealing for two motives: first, it is scale invariant, which means it evaluates the model's performance irrespective of the magnitude of the absolute values obtained, and second, it is classification threshold invariant, which represents that it evaluates the performance of the model, regardless of the threshold fixed to obtain the various categories used.

Eighty-five characteristics from the lung regions and the cough signal were assessed, and the statistically significant characteristics were those that had a p 0.05 in the univariate analysis. Only ten characteristics were preserved and utilized for COVID-19 detection after this stage. COVID-19 detection was a four-class classification job in the experiment. The classifiers are trained and verified using a tenfold cross-validation procedure based on the attributes obtained from the feature selection techniques [21]. Finally, an external testing dataset is used to put the model to the test. The various classifiers used in the process are analyzed, and the
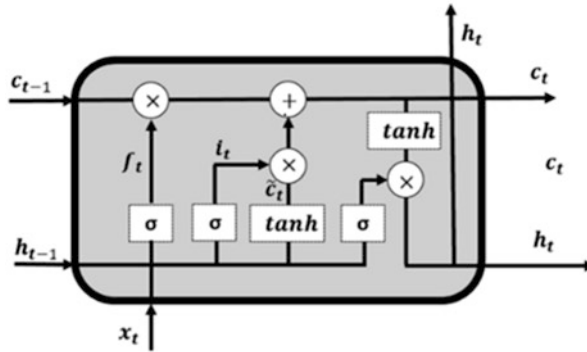
Fig. 11 Architecture of LSTM

tree-based classifier models random forest and XGBoost are found to outperform linear regression-based models in all the metrics. The comparison of the performance of the different model is given in Fig. 11.

Random Forest

RF is a tree-based regression/classification model, in which random samples are chosen from a set of data. The random forest algorithm will then be used to generate a decision tree for each sample. Voting will be done for each expected outcome for each decision tree. Finally, choose the prediction outcome with the highest votes as the final prediction outcome. To classify individuals as positive or negative to COVID-19, a RF classifier comprising of 250 trees was utilized. To split the node, completely grown and unpruned trees are employed, with the Gini index as the criteria. At each split, the features are permuted at random [22].

XGBoost

Gradient boosting is a machine learning-based regression and classification technique that produces a final output based on a group of low prediction decision tree models. The model architecture is framed stage-by-stage similar to other tree-based models such as RF or other boosting approaches; also it broadens the scale of the output by allowing optimization of any differentiable loss function. XGBoost is one of the gradient boosting implementations, but what sets it apart is that it employs a more regularized model formalization to control overfitting, resulting in an improved performance and a reduction in overfitting.

Support Vector Machine

Support vector machine is a frequently applied for classification or regression model that finds lines or boundaries, called hyperplane, using the extreme vectors to correctly identify the training dataset. Then, it chooses the line or boundary with the greatest distance from the nearest data points from those lines or boundaries.

LASSO, Ridge, and Elastic Net Regression

When the model coefficient is very high, the training dataset will be over-fitted. Regularization, which penalizes the bigger coefficient, can be used to solve these difficulties. Ridge regression is a modification of applying a penalty equal to the square of the coefficient values, often known as the L2-norm. Alternatively in the Lasso model, the loss function is changed with the addition of a penalty equal to the magnitude of the coefficient. Elastic net is a combination of both the ridge and LASSO model.

Some respiratory illnesses, like COVID-19, are thought to have a natural defense mechanism in the form of cough. Existing subjective clinical approaches to cough sound analysis were hampered by the human audible hearing range. As illustrated in this work, exploring noninvasive diagnostic options considerably above the audible frequency range (i.e., 48,000 Hz) employed for sample data can overcome this constraint. Signal processing-based techniques face extra hurdles due to the nonstationary properties of cough sound samples. Cough patterns also vary in human beings with the same medical condition. Cough characteristics that are closely related to intensity levels in the temporal domain can differ for the same disease. Figure 12 shows performance score achieved with different classifiers.

When an abnormal occurrence is present, the cough sound is analyzed by the basic frequency and the harmonics component. The instability in the cough sound that makes up the multiples of the fundamental frequencies is caused by airway
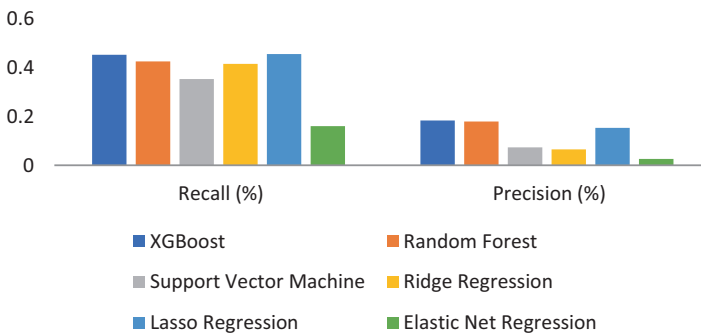

Fig. 12 Performance score achieved with different classifiers

compression. The system that is able to capture various features in the spatial and frequency domain of the cough signal for the complete cycle can accurately predict the variations in the cough due COVID-19 or pneumonia.

Because publicly available databases only include COVID-19-positive and COVID-19-negative examples, this research focuses on distinguishing COVID-19 and other pneumonia-related irregularities in cough sound and lung images from healthy/controls data. The proposed method, on the other hand, may be able to distinguish problematic cough sounds from various pulmonary/respiratory disorders, such as COVID-19, viral pneumonia, and bacterial pneumonia. Cough noises' pathophysiology and acoustic properties can provide valuable information in the frequency domain that can be used to characterize them for multiclass classification tasks. Pneumonia-related diseases make the patient's airways to become irritated and constricted [23, 24]. To demonstrate the distinctiveness of COVID-19, pneumonia, and normal samples, some of their frequency domain features are analyzed. When compared to COVID-19 and samples of asthma cough, the spectral entropy of the pneumonia sample is substantially greater for the majority of the frames. For the three respiratory illnesses stated, other features, such as spectrum flux, MFCC, and feature harmonics, are likewise nonidentical. Using a greater number of MFCCs consistently improves the performance. The study concludes that the machine learning classifiers are extracting the information that is not commonly sensible to human listeners, since the spectral frequency resolution utilized to compute the multidimensional MFCCs exceeds the audible property of the human auditory system, and these internal features are extracted and suitably selected by the time series-based deep learning LSTM model. Fig. 13 shows variation in the frequency domain features for the complete cycle: (a) spectral entropy, (b) spectral flux, (c) MFCC coefficient, and (d) feature harmonics for different cases of cough samples

When compared to previous studies, the work shows encouraging results in the categorization of CT scans into normal, COVID-19, bacterial pneumonia, and viral pneumonia.

Our findings show that analysis based on the spectral components of the cough signal may be utilized to discriminate COVID-19-positive persons and other pneumonia-related cough sound from noninfected people. To enable an effective classification, the collection of features taken from the pictures must be represented and adequate, to avoid producing mistakes in the classification stage. In light of this, a key component of the suggested technique is the use of a GA to choose the most significant traits. The GA and LSTM method used for feature selection improved the classification results and also considerably reduced the dimensionality of the feature used for classification, making the classification process more responsive. Figure 14 shows the comparison of the performance of various ML-based classifiers, and Table 2 gives the details about the comparison of different classifiers.

The suggested architecture had sophisticated capabilities for categorizing CT scans and cough sounds into COVID-19, pneumonia viral, pneumonia bacterial, and normal instances, and it may be utilized as a COVID-19 diagnostic or screening tool.

Fig. 13 Variation in the frequency domain features for the complete cycle: (**a**) spectral entropy, (**b**) spectral flux, (**c**) MFCC coefficient, (**d**) feature harmonics for different cases of cough samples
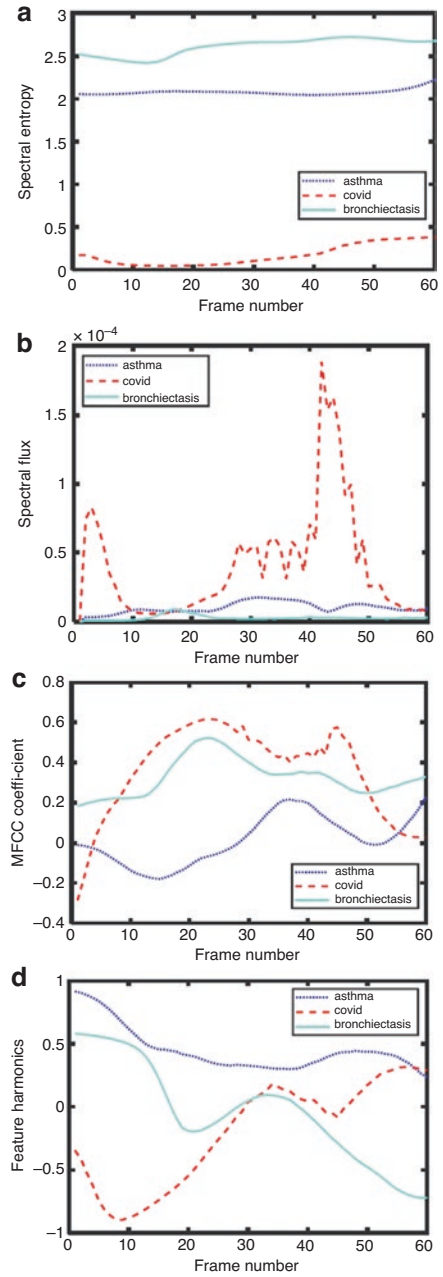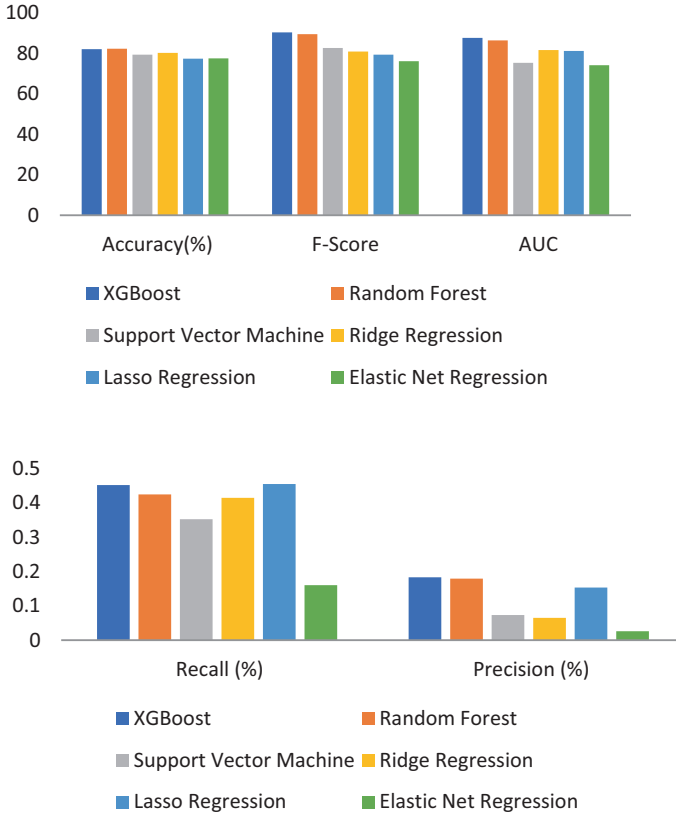
Fig. 14 Comparison of the performance of various ML-based classifiers

Table 2 Comparison of different classifiers

Experiment classifier	Accuracy (%)	Recall (%)	Precision (%)	F-score	AUC
XGBoost	81.9	0.451	0.183	90.2	87.5
Random forest	82.1	0.424	0.179	89.3	86.2
Support vector machine	79.2	0,352	0.073	82.5	75.2
Ridge regression	80.1	0.414	0.065	80.7	81.5
Lasso regression	77.2	0.454	0.153	79.2	81.0
Elastic net regression	97.2	0.26	0.026	0.026	98.7

Conclusion

This article proposes a new method of collecting multi-characteristic dataset of cough and lung images from people diagnosed with COVID-19 and two types of pneumonia along with noninfected individuals. The system studies the main features that contribute to COVID-19 for early results of COVID-19 detection from

cough signal and also using the lung images obtained at the later stage, highlighting the importance of using the combined feature of audio signatures and images to detect COVID-19 symptoms. The proposed literature suggests that extracting the features using deep learning algorithms, LSTM, and genetic algorithm for feature selection and machine learning algorithm is suitable for a COVID-19 detection task; also, we go a step further and provide an in-depth analysis of the most useful spatial and spectral characteristics of sound, with the goal of analyzing the phenomena that changes the acoustic characteristics of COVID-19 coughs. Machine learning approaches are useful not just for distinguishing COVID-19 cases from other pneumonia patients but also for assisting doctors in tracking and predicting their patients' prognosis and treatment outcomes. As a result, studies conducted based on the cough sounds; chest x-rays; lung images, along with laboratory data; and other features, like demographic along with usage of deep learning and machine learning algorithms for various process like feature extraction and classification, will result in early diagnosis of the disease.

References

1. E. Soares, P. Angelov, S. Biaso, M.H. Froes, D.K. Abe, SARS-CoV-2 CT-scan dataset: A large dataset of real patients CT scans for SARS-CoV-2 identification. *MedRxiv*. https://doi.org/10.1101/2020.04.24.20078584

2. A.M. Ayalew, A. OlalekanSalau, et al., Detection and classification of COVID-19 disease from X-ray images using convolutional neural networks and histogram of oriented gradients. Biomed. Signal Process. Control **74** (2022). https://doi.org/10.1016/j.bspc.2022.103530

3. S. Aydın, H.M. Saraoğlu, S. Kara, Log energy entropy-based EEG classification with multi-layer neural networks in seizure. Ann. Biomed. Eng. **37**(12), 2626 (2009)

4. R. Bachu, S. Kopparthi, B. Adapa, B.D. Barkana, Voiced/unvoiced decision for speech signals based on zero-crossing rate and energy, in *Advanced Techniques in Computing Sciences and Software Engineering*, (Springer, 2010), pp. 279–282

5. K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2016), pp. 770–778

6. M. La Salvia, G. Secco, et al., Deep learning and lung ultrasound for Covid-19 pneumonia detection and severity classification. Comput. Biol. Med. **136** (2021). https://doi.org/10.1016/j.compbiomed.2021.104742

7. H. Wright, Genetic algorithms for real parameter optimization, in *Foundations of Genetic Algorithms*, (vol. 1, Elsevier, 1991), pp. 205–218

8. A.E. Eiben, J.E. Smith, et al., *Introduction to Evolutionary Computing*, vol 53 (Springer, 2003)

9. S. Hochreiter, J. Schmidhuber, Long short-term memory. Neural Comput. **9**(8), 1735–1780 (1997). https://doi.org/10.1162/neco.1997.9.8.1735

10. K.E. ArunKumar, V. Dinesh, et al., Comparative analysis of gated recurrent units (GRU), long short-term memory (LSTM) cells, autoregressive integrated moving average (ARIMA), seasonal autoregressive integrated moving average (SARIMA) for forecasting COVID-19 trends. Alex. Eng. J. **61** (2022). https://doi.org/10.1016/j.aej.2022.01.011

11. E.D. Carvalho et al., An approach to the classification of COVID-19 based on CT scans using convolutional features and genetic algorithms. Comput. Biol. Med. **136**, 104744 (2021)

12. P. Anandhanathan, P. Gopalan, Comparison of machine learning algorithm for COVID-19 death risk prediction (2021)

13. R. Islam, E. Abdel-Raheem, M. Tarique, A study of using cough sounds and deep neural networks for the early detection of Covid-19. Adv. Biomed. Eng. **3**, 100025 (2022)
14. V. Despotovic et al., Detection of COVID-19 from voice, cough and breathing patterns: Dataset and preliminary results. Comput. Biol. Med. **138**, 104944 (2021)
15. WHO et al., Summary of probable SARS cases with onset of illness from 1 November 2002 to 31 July 2003. http://www.who. int/csr/sars/ country /table2004_04_21/en/index. html (2003)
16. R. Miyata, N. Tanuma, M. Hayashi, T. Imamura, J.-I. Takanashi, R. Nagata, A. Okumura, H. Kashii, S. Tomita, S. Kumada, et al., Oxidative stress in patients with clinically mild encephalitis/encephalopathy with a reversible splenial lesion (MERS). Brain Dev. **34**(2), 124–127 (2012)
17. F. Pan, T. Ye, P. Sun, S. Gui, B. Liang, L. Li, D. Zheng, J. Wang, R.L. Hesketh, L. Yang, et al., Time course of lung changes on chest ct during recovery from 2019 novel coronavirus (covid-19) pneumonia. Radiology **295**(3) (2020). https://doi.org/10.1148/radiol.2020200370
18. H. Ritchie, E. Mathieu, L. Rodes-Guirao, C. Appel, C. Giattino, E. Ortiz-Ospina, J. Hasell, B. Macdonald, D. Beltekian, M. Roser, Coronavirus pandemic (covid-19), Our World in Data (2020). https://ourworldindata.org/coronavirus
19. A. Halder, B. Datta, Covid-19 detection from lung ct-scan images using transfer learning approach. Mach. Learn. Sci. Technol. **2** (2021). https://doi.org/10.1088/2632-2153/abf22c
20. G. Soldati et al., Proposal for international standardization of the use of lung ultrasound for patients with COVID-19: A simple, quantitative, reproducible method. J. Ultrasound Med. (2020). https://doi.org/10.1002/jum.15285
21. M.J. Fiala, A brief review of lung ultrasonography in COVID-19: Is it useful? Ann. Emerg. Med. (2020). https://doi.org/10.1016/j.annemergmed.2020.03.033
22. T. Ai et al., Correlation of chest CT and RT-PCR testing in coronavirus disease 2019 (COVID-19) in China: A report of 1014 cases. Radiology, 200642 (2020). https://doi.org/10.1148/radiol.2020200642
23. D. Wang, B. Hu, C. Hu, F. Zhu, X. Liu, J. Zhang, B. Wang, H. Xiang, Z. Cheng, Y. Xiong, et al., Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus–infected pneumonia in Wuhan, China. J. Am. Med. Assoc. **323**(11), 1061–1069 (2020)
24. A. Carfì, R. Bernabei, F. Landi, et al., Persistent symptoms in patients after acute COVID-19. J. Am. Med. Assoc. **324**(6), 603–605 (2020)