

RESEARCH

Open Access



A deep convolutional neural network approach using medical image classification

Mohammad Mousavi¹ and Soodeh Hosseini^{1*}

Abstract

The epidemic diseases such as COVID-19 are rapidly spreading all around the world. The diagnosis of epidemic at initial stage is of high importance to provide medical care to and recovery of infected people as well as protecting the uninfected population. In this paper, an automatic COVID-19 detection model using respiratory sound and medical image based on internet of health things (IoHT) is proposed. In this model, primarily to screen those people having suspected Coronavirus disease, the sound of coughing used to detect healthy people and those suffering from COVID-19, which finally obtained an accuracy of 94.999%. This approach not only expedites diagnosis and enhances accuracy but also facilitates swift screening in public places using simple equipment. Then, in the second step, in order to help radiologists to interpret medical images as best as possible, we use three pre-trained convolutional neural network models InceptionResNetV2, InceptionV3 and EfficientNetB4 and two data sets of chest radiology medical images, and CT Scan in a three-class classification. Utilizing transfer learning and pre-existing knowledge in these models leads to notable improvements in disease diagnosis and identification compared to traditional techniques. Finally, the best result obtained for CT-Scan images belonging to InceptionResNetV2 architecture with 99.414% accuracy and for radiology images related to InceptionV3 and EfficientNetB4 architectures with the accuracy is 96.943%. Therefore, the proposed model can help radiology specialists to confirm the initial assessments of the COVID-19 disease.

Keywords Deep learning, Convolutional neural networks, Medical image classification, Internet of health things

Introduction

Acute respiratory syndrome of the Corona virus-2 (SARS-Cov-2) is the virus resulting in COVID-19 which is a respiratory viral and disease. From May 2021 COVID-19 has been called a pandemic by the World Health Organization [1, 2]. Since COVID-19 is a contagious and easily transmissible disease, it has affected the lives of billions of people around the world. Early and accurate diagnosis of COVID-19 is very important to control the spread of the disease and reduce its mortality.

There is also a shortage of health workers to care for all patients. Therefore, it is very important to develop an automated intelligent method that provides immediate and high-accuracy results and essentially enables testing anywhere and anytime. This can be provided by the Internet of Health Things and the retrieved data can be analyzed using artificial intelligence techniques for diagnosis. Even if medical imaging centers are established in remote areas, availability of radiologists remains as a problem. Developing or undeveloped countries are struggling to improve their diagnostic capabilities; because current methods such as RT-PCR require expensive kits on-site testing, and these types of kits are not always easy to obtain. Hence, an easily accessible remote diagnosis model is essential for immediate screening and diagnosis of infected cases. According to the statistics published by

*Correspondence:

Soodeh Hosseini
so_hosseini@uk.ac.ir

¹ Department of Computer Science, Faculty of Mathematics and Computer, Shahid Bahonar University of Kerman, Kerman, Iran



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

world Health Organization (WHO) so far (April 2022), the total number of confirmed cases of COVID-19 all around the world has been more than 500 million persons. To deal with the pandemic, researchers are looking for a wide range of technologies like Internet of Health Things, artificial intelligence and metadata that can assist overcoming the challenges posed by COVID-19. The Internet of Health Things is an expanding ecosystem that integrates a variety of electronic devices and physical objects capable of exchanging information to communicate, collect and exchange data. The Medical Internet of Health Things, plays a fundamental role in the healthcare sector and increases the accuracy, reliability and efficiency of electronic devices.

In addition to RT-PCR test, several artificial intelligence-based methods have recently proposed that use chest CT-Scan [3–6] and X-ray [7–9] to detect visual indicators of COVID-19 viral infection. Meanwhile, to use RT-PCR, CT-Scan and X-ray for diagnosis, it is necessary to visit well-equipped clinical centers. Since the mentioned test protocol requires the presence of people from the treatment staff, there is a greater risk of infection due to the high possibility of infection. To limit the exponential growth of COVID-19 cases, one solution is to design a model that can perform biological tests without involving many people.

The main purpose of this paper is to propose an automatic detection model for cases of COVID-19 using a proposed neural network based on cough sound for screening and then three pre-trained convolutional neural networks with the help of transfer learning technique in diagnosing COVID-19 based on radiological images and CT-Scan of the chest. The proposed model includes two parts of collecting information with the help of devices equipped with Internet of Health Things technology and sending it to the information repository, and then processing information and extracting knowledge.

The main contributions of this research are as follows:

- 1) An IoHT-based model of deep learning models for automatic diagnosis of COVID-19 patients is proposed.
- 2) A deep learning algorithm is utilized for the detection of individuals with COVID-19. It leverages audio data, such as cough sounds, to conduct preliminary screening of individuals suspected of being infected. This approach enhances the speed and accuracy of diagnosis and enables screening in public settings with minimal equipment.
- 3) The selection of core features, such as spectral Roll-off, spectral bandwidth, spectral centroid, RMS, Chromagram, MFCC, and zero crossing rate, has strategically leveraged the effective representation of

the sound spectrum. These features deliver pivotal insights into the sound spectrum, thereby empowering the model to recognize audio patterns with high efficacy. Their proven track record in speech processing and pattern recognition applications signifies the pivotal role these features play in extracting meaningful information from the sound spectrum.

- 4) The strategy exploits transfer learning capabilities of three pre-trained convolutional neural network models to differentiate between cases of covid-19 induced pneumonia and healthy cases. These advanced models demonstrate a notable enhancement in disease diagnosis and identification compared to conventional models, leveraging transfer learning and pre-existing knowledge.
- 5) In order to enhance the performance of the pre-trained convolutional neural network model, data augmentation techniques have been employed. In cases where training data is limited or unbalanced within each class, data amplification provides the ability to offer more diverse information to the model, preventing overfitting. This approach enables the model to recognize intricate patterns within the data, ultimately leading to improved performance.
- 6) Detailed analysis of model performance has been done. To do so, confusion matrix for each model has been presented in addition to usage made of the main criteria.

Related work

An intelligent health care model equipped with Internet of Things has been introduced by Ahmed et al. [1] to identify and classify chest X-ray images in to three coronavirus, pneumonia, and healthy classes. In the first stage, after pre-processing, the data enhancement operation is applied to increase the diversity of the data set, then the data is divided into two training and testing sets, and two pre-trained architectures, VGG19 and InceptionV3, are using for classification. The total of 4500 chest X-ray images have used in their research and finally best accuracy has been recorded as 97%. In [10], the author proposed a deep learning model in the IOT platform to diagnose people with the disease of COVID-19 using chest CT-Scan images. This model includes several components including CT-Scan images received from mobile CT-scanners to establishment of dataset on cloud computing, online model training and obtaining final results. In recommended model, validation would be done via three pre-trained DenseNet121, ReseNet50V2, and Xception networks. Then predictions obtained would be integrated and evaluated for final classification. A recommended convolutional neural network have used by Thakur and Kumar [6] to classify chest X-ray and

computed tomography scan. Two-class and multi-class classification scenarios have used by them. To do so, 11,095 images have used. Finally, best accuracy in two-class classification has been reported to be 99.6%; and, it was 98.2% in multiclass classification. X-ray and CT-Scan images have used by Elpeltagy and Sallam [7] to predict cases infected by COVID-19 based on InceptionV3, Visual Geometry Group Network, DenseNet, AlexNet, GoogleNet, and ReseNet in addition to a recommended model. Highest accuracy obtained using X-ray images has been 97.7% and those from CT images has been 97.1%. An IoT-based model has been recommended by Loey and Mirjalili [8] in order for coughing classification of people suffering from COVID-19. First stage of the model is feature extraction and the sound would be turned to image based on scalogram. At second stage, feature extraction and classification would be done based on deep learning models including ReseNet, MobileNet, GoogleNet, and NasNet that are the most applied models of transfer learning. Audio dataset including 3325 sound of coughing (one second each) have used; and, finally the best result reported was 94.0% from ReseNet18. Chest X-ray images have used to classify people infected by COVID-19, via deep learning and through usage made of transfer learning. The paper has been dealt with five models based on pre-trained convolutional neural networks (AlexNet, VGG16, ReseNet50, ReseNet101, and ReseNet152) on 185 X-ray images including four classes. Considering low numbers of images in dataset, data augmentation process including 90, 180, and 270° image rotation in different axes have used. Also, brightness increase has used to improve classification performance. Best results would be obtained when pre-trained ReseNet 152 architecture would be taught through bigger data groups in average number of training courses through Nadam optimization function. To classify COVID-19 and healthy chest X-ray images deep learning-based approaches i.e. deep feature extraction and fine-tuning of pre-trained convolutional neural networks and end-to-end training of a developed convolutional neural network model have used by Ismael and Şengür [11]. The model includes 21 layers such as convolution, maximum integration, fully connected layers, and final classification layer along with batch normalization and ReLU layers. Here, dataset including 180 COVID-19 images and 200 healthy images have been applied and for deep feature extraction, pre-trained convolutional neural networks (VGG16, VGG19, ReseNet101, ReseNet50, and ReseNet18) have used. Finally deep features extracted from ReseNet50 model through SVM have obtained 94.7% accuracy which is the highest score among all the results obtained. A diagnostic method for COVID-19 through deep convolutional neural network based on EfficientNet-B4 has been made

by Marques et al. [12] to classify chest X-ray images. The best and highest accuracy in two-class classification is reported with an accuracy of 99.51%.

In continuation, Table 1 shows works done in relation to COVID-19 cases diagnosed through convolutional neural network.

Zhao et al. [23] introduced an innovative motif-aware MDA prediction model called MotifMDA, which integrates diverse high- and low-order structural information. Yang et al. [24] developed a new fuzzy-based deep attributed graph clustering model that performs the task in a fully unsupervised and end-to-end manner, eliminating the need for traditional clustering methods. Additionally, Zhao et al. [25] presented a new graph representation learning model aimed at drug repositioning, combining both higher and lower-order biological information. Furthermore, a novel computational approach was proposed for predicting lncRNA-miRNA interactions (LMIs) utilizing neighborhood-level structural representation [26].

Methodology

In this section, the proposed model is described. In addition to high accuracy in diagnosing cases of COVID-19, the proposed model minimizes human contact with the suspect, which also reduces the rate of transmission of the virus to healthy people. Two types of audio data and chest medical images are used in this model. Data transfer to the data repository in the proposed model is done by Internet of Health Things. In the first stage, a 30-s sample of the cough sound of people is divided into two classes, healthy and sick, through a binary classification with the help of a 5-layer neural network. In the next step, by using three convolutional neural network models, with the help of two types of medical images, CXR and CT-Scan, susceptible people are divided into three classes: COVID-19 patients, pneumonia patients, and healthy. In the following, we explain the details of the proposed model and the data have used.

Audio data

Screening people suspected of having COVID-19 infection before referring to medical centers is of high importance to protect healthy people and to reduce work pressure on health care personnel especially at peak of pandemic. For this purpose, we use two audio datasets to distinguish cases of COVID-19 from healthy individuals. The numbers of data to train neural network is of high importance. Moreover, numbers of data in classes of dataset have to be proportionate to each other. The first dataset [27] includes 19 cases of coughing sound belonging to those people suffering from COVID-19 and 21 cases of coughing by healthy people. In addition, to

Table 1 The summary of studies in terms of COVID-19 disease identification

Author	Technique	Dataset	Data type	Result
Ahmed et al. [1] (2022)	VGG19, InceptionV3	4500 CXR images	CXR	%97
Deb et al. [2] (2022)	VGGNet, GoogleNet, DenseNet, NASNet, ResNet, ReseNext, Ensemble model	J.P. Cohen[13], chest-xray-COVID19- pneumonia dataset, private dataset from MGM medical college and hospital	CXR	Private Dataset: Binary:%95.6 Multi:%93.4 Public Dataset: Binary:%98.5 Multi: %88.9
Lella and Pja [3] (2022)	Proposed DCNN, VGGNet	Collected the COVID-19 sound dataset from Cambridge university	Sound	%95.4
Chowdhury et al. [4] (2022)	Ensemble-base MCDM	Cambridge[14], Coswara[15], virufy[16], NoCoCoDa[17]	Sound	%95 using extra-trees classifier
Shorfuzzaman [5] (2021)	ReseNet50V2, DenseNet121, Xception	2482 CT images	CT Scan	%96.5
Thakur and Kumar [6] (2021)	Proposed CNN	11,095 medical images	CXR and CT Scan	Binary: %99.6 Multiclass: %98.2
Elpeltagy and Sallam [7] (2021)	ReseNet50, ReseNet53, GoogleNet, AlexNet, DenseNet201, VGGNet, InceptionV3, ReseNet50 + SVM, Proposed CNN	[18]	CXR and CT Scan	%97.7 for CXR, %97.1 for CT
Loey and Mirjalili [8] (2021)	ReseNet, GoogleNet, MobileNet, NasNet	3325 cough sounds	sound	%94.9 by ReseNet18
Lorencin et al. [9] (2021)	AlexNet, VGG16, ReseNet50, ReseNet101, ReseNet152	Data collection from Kragujevac	CXR	%95 by ReseNet152
Ismael and Şengür [11] (2021)	ReseNet18, ReseNet50, ReseNet101, VGG16, VGG19 + SVM	[19–21]	CXR	%94.7 by ResNet50 and SVM classifier with the Linear kernel
Marques et al. [12] (2020)	EfficientNet-B4	[13, 20]	CXR	Binary: %99.5 Multiclass: %96.7
Ahmed et al. [22] (2020)	Faster RCNN + ResNet101	11,000 CXR images	CXR	%98
Ouyang et al. [10] (2020)	Dual sampling using 3DResNet34	4982 CT scan images	CT scan	%94.4

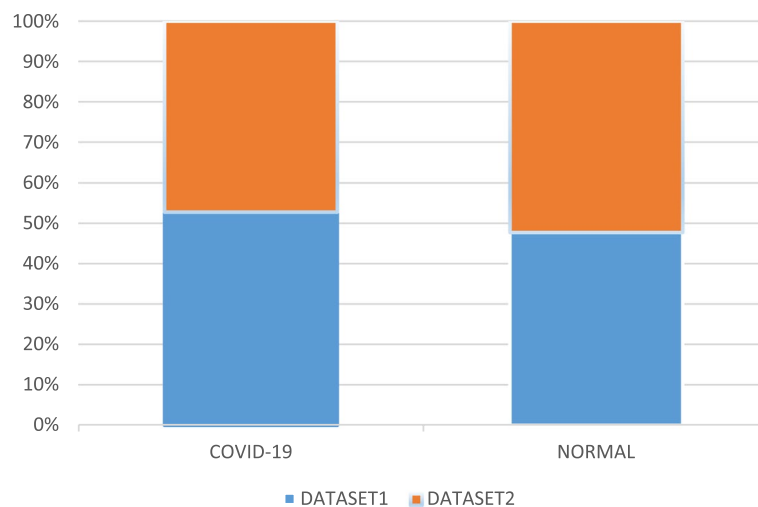


Fig. 1 Audio dataset structure

increase the number of data sets from the second dataset [28], which includes 17 audio data belonging to people with COVID-19 and 733 audio data belonging to healthy

people, all the data of COVID-19 along with 23 data of the healthy class are added to the data set. Figure 1 shows distribution of audio samples of coughing sound use in

the research. Each of coughing samples would be re-sampled with 225 kHz sampling rate.

Feature extraction

Each audio signal contains many features. Nevertheless, it is necessary to extract the characteristics that are relevant to the problem that it aims to solve. The process by which feature are extracted for use in analysis is called feature extraction. Audio wave considered in feature extraction process with standard frequency (22 kHz sampling rate) would be sampled to assure data uniformity. Seven spectral features of the sampled sound are extracted using the Librosa library [29] from Python, which are as follows:

Chromatogram

A chromatogram can be computed from one wave spectrum or exponent. One of the main characteristics of chromatogram is picturing harmonic and melodic features of music; while, they are sensitive to changes of sound. The chromatogram would be computed via Short-time Fourier transform (STFT).

RMS

Root Mean Square is a measurement tool measuring loudness of an audio tune within a range of almost 300 ms.

An audio signal can have the two values of negative and positive domains. If arithmetic mean of a sinusoidal wave would be considered; positive values would be compensated by negative values and zero would be the result. This is where RMS level can be useful domain can be considered as a criterion for signal strength based on signal magnitude, regardless of useful domain being negative or positive. The magnitude is calculated by squaring each sample, then the average signal is calculated and finally the square root operation follows.

Spectral centroid

The feature shows where center of the mass is located for a sound and would be computed as weighted average for available frequencies in the sound.

Spectral bandwidth

Bandwidth refers to maximum amount of data transferable by a device within one unit of time. The more the bandwidth would be, the more data can be sent. The Frequency (f) is the number of waves happening in one signal during one second. The Frequency is measured by Hertz (Hz). The Period is the time duration of a wave to be completed ($T = 1/f$). If maximum frequency would be considered as f (max) and minimum frequency would be f (min), the formula to calculate bandwidth would be as follows:

$$B = [f(\text{max}) - f(\text{min})] \quad (1)$$

Spectral rolloff

This feature is a signal shape size and shows the frequency below determined percentage of the whole spectral energy.

Zero crossing rate

This is the rate of changes of signs during one signal. In fact, it is the rate within which signal would be changes from positive to negative or vice versa. The feature is heavily use in speech recognition and music information retrieval.

Mel-Frequency Cepstral Coefficients (MFCC)

In MFCC feature extraction, fast Fourier transform (FFT) would be implemented to find exponent spectrum of each model. After that, MEL scale would be used to process filter bank in exponent spectrum. The general shape of a spectrum is briefly described by MFCC.

Medical images data

Researchers have used various public datasets such as X-ray images [7–9] and CT-Scan images to find cases of COVID-19. According to research performed, diagnosis of cases infected by coronavirus would be taken place with more accuracy according to CT-Scan images compared to those of X-ray. However, the patient receives lower dosage of X-rays compared to that of CT-Scan. This can be ignored in low number of tests; however, it is of high importance as far as pregnant women and children are concerned. Here, two datasets of X-ray and CT-Scan images have used to provide the possibility of using various types of medical images for the proposed model.

CT-Scan dataset

We combine two datasets of CT-Scan [29, 30] have been collected from various online sources. As it is observed in Fig. 2, 5203 images are belonging to COVID-19 class and 2418 images to healthy class in this dataset. Also, for better performance of proposed model in real world, 2618 images belonging to community-acquired pneumonia class have been added to the first dataset. In total, used 10,239 CT-Scan visual data.

X-ray dataset

X-ray images have been collected by Kumar [21] from several other sources. Images have been classified in to three classes of patients suffering from COVID-19, those suffering from Pneumonia not COVID-19, and healthy ones; and, all sizes of all images have been changed to 256*256. In total, used 5228 chest X-ray images and data

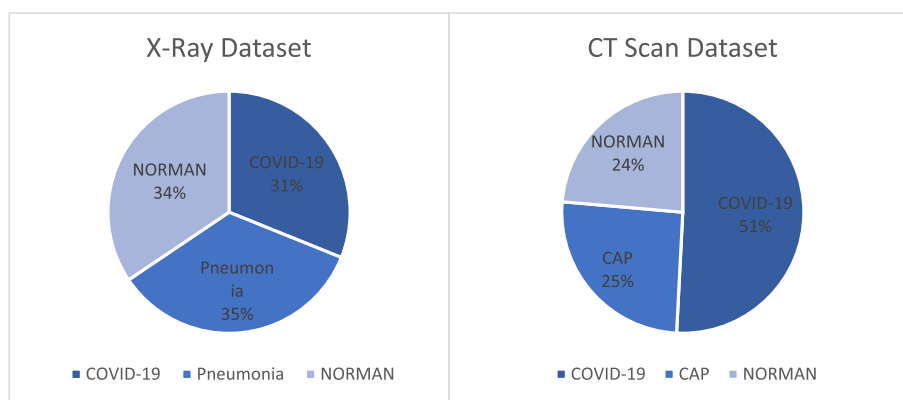


Fig. 2 Structure of medical image datasets

distribution of X-ray dataset is provided in Fig. 2. This dataset has some files with the same names and there is no other information available.

Data augmentation

Large number of training data is one of deep learning requirements. In this respect, data augmentation process have used to increase classification performance. The process has been performed to the aim of artificial increase of training datasets [31]; and, test dataset remain fixed. A set of various geometrical operations have used to increase dataset. Here, horizontal flip of images and image magnification have used in Fig. 3.

Proposed model

In this section, the proposed model is explained based on recommended Internet of Health Things. IoHT Supervising huge numbers of patients in hospitals or homes would become possible by Internet of Health Things. In addition to the main data used in the research, there is a possibility for transfer of other biometric data of patients to the main databank so that knowledge would be extracted and analyzed with no medical staff prone to infection. Research has been performed by [32] regarding usage made of Internet of Health Things to diagnose fever. A low-cost Internet of Health Thing model has been recommended, uploading automatically data resulted through wireless communication via smart phones in a global network. Hence, the test results in every place in the world would be available immediately. Such model of Internet of Health Things is a very important tool for physicians to confront infectious diseases.

In this paper, an automatic diagnostic model for people suffering from COVID-19 has been introduced based on Internet of Health Things. First, data have been sent to databank through IoHT. Then using machine learning algorithms, data have been processed to be extracted.

In the first step, to screen people suspected of COVID-19, one neural network model has used through usage made of the sound of coughing in a two-class classification of healthy and infected people. Then, three pre-trained convolutional neural network model have used through transfer learning so that COVID-19 infected cases, chest X-ray and CT-Scan images have used. In one multi-class classification, main goal is and correct diagnosis of images and their classification in to three classes of (Coronavirus) patients, pneumonia (non-COVID-19) patients, and healthy people. The reason for using three-class classification of medical images is helping radiologists to prioritize COVID-19 patients so that more outbreak of the disease would be prevented, and patients would be more effectively treated and the community remains safe. The work process is shown in Fig. 4.

Internet of health things

Nowadays, various applications are covered up by Internet of Health Things such as transportation, smart cities, supervision, health care, and etc. For example, IoHT in health care industry can play an important role in remote supervision at hospitals, especially at home for elderly people suffering from chronic diseases. Also, by using IoHT and automatic diagnostic models, people suffering from COVID-19 can be diagnosed at initial stages so that more irregular spread of the disease would be prevented. By using the technology, in future, health care models will experience main effects such as reduction of response time to diagnose anomalies, high quality care, low hospitalization cost, and high life expectancy.

During COVID-19 pandemic, artificial intelligence and IoHT have been more taken into consideration in the field of health care, where screening and diagnostic methods can be performed more conveniently. Thermal imaging and supervising social distance are also from those performances mainly considered at screening

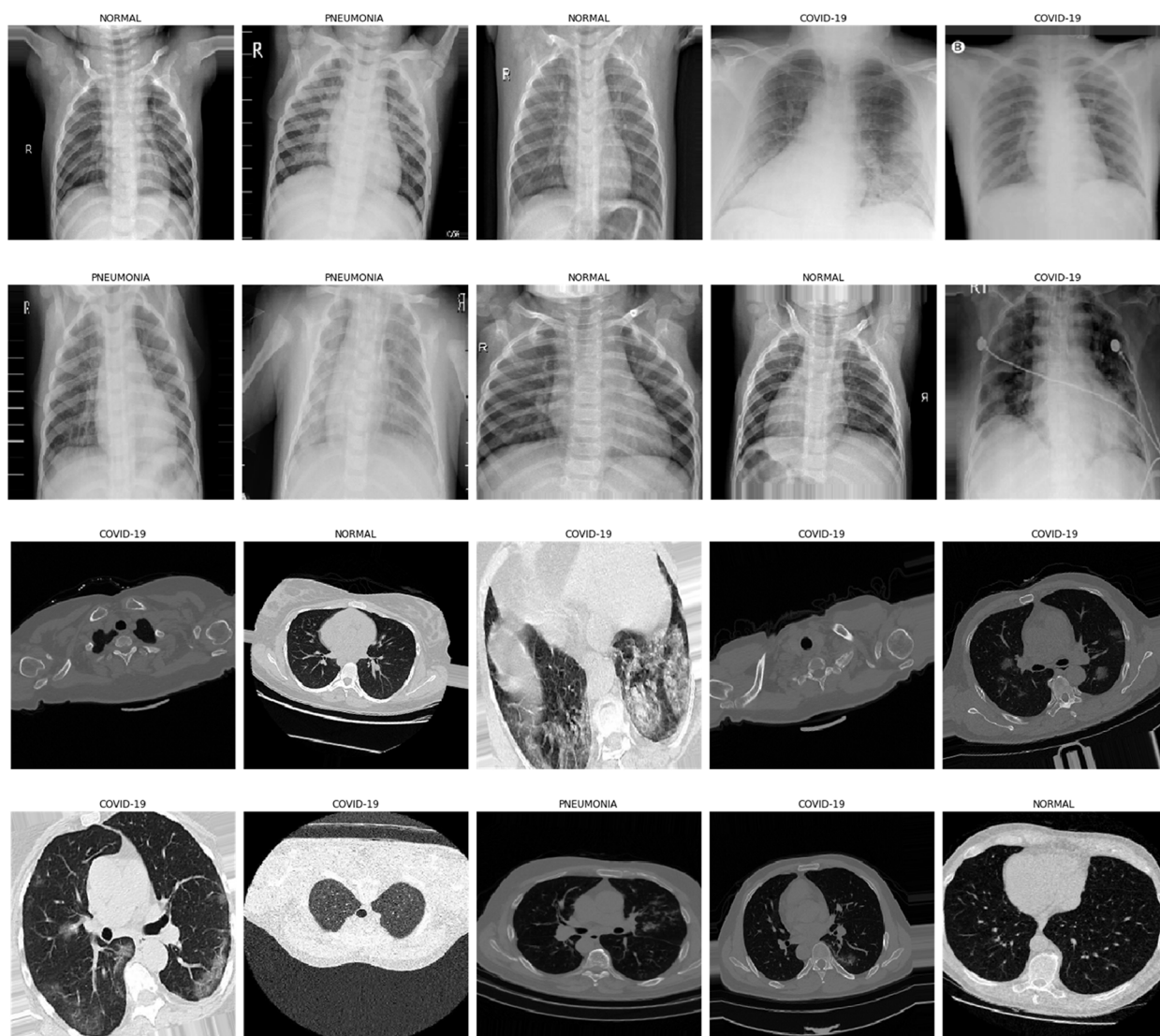


Fig. 3 Examples of images after applying the data augmentation process

stage. In fact, using these devices is aimed at detection of body temperature, screening people wearing masks, and controlling social distance. Finally, the Internet of Health Things can help reduce costs and reduce infrastructure complexity [31].

The rapid evolution and acceptance of IoHT especially during pandemic may create more security concerns. Therefore, the main challenge is protecting privacy of important and sensitive medical data. Numerous invasions, threats and risks and dangers can affect different layers of IoHT architecture. Hence, an IoHT ecosystem has to be safe, using strict privacy protocols [33]. IoHT devices with low security are one of the most effective channels for cybercriminals to disclose the clients' data through communication flows [34]. Hence, medical care

centers should develop risk assessment guidelines to ensure data protection.

Respiratory sound classification

Due to the advancement of artificial intelligence techniques in voice and signal processing, as well as the development of voice programs based on machine learning, we have proposed a model based on machine learning that can detect cases of COVID-19 using cough samples and two-class classification. The goal is to screen suspected cases of COVID-19 with the help of the Internet of Health Things and mobile phones, so that people can record a short sample of their cough for 30 s using their mobile microphone and upload processing in the program environment. In addition to preventing unnecessary visits to

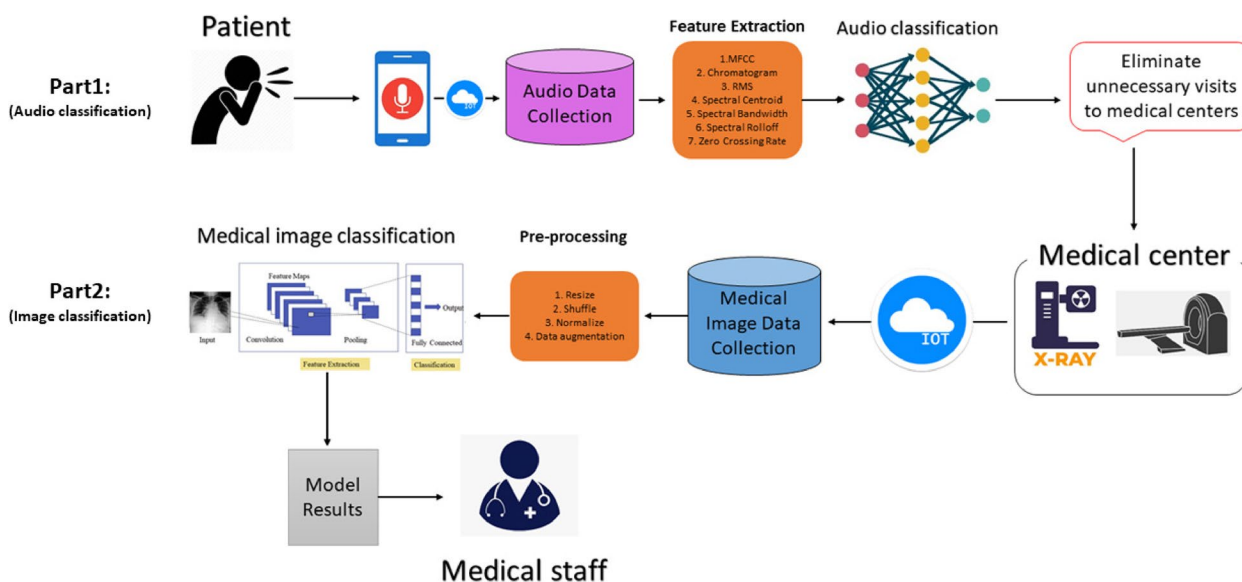


Fig. 4 The workflow of the proposed model

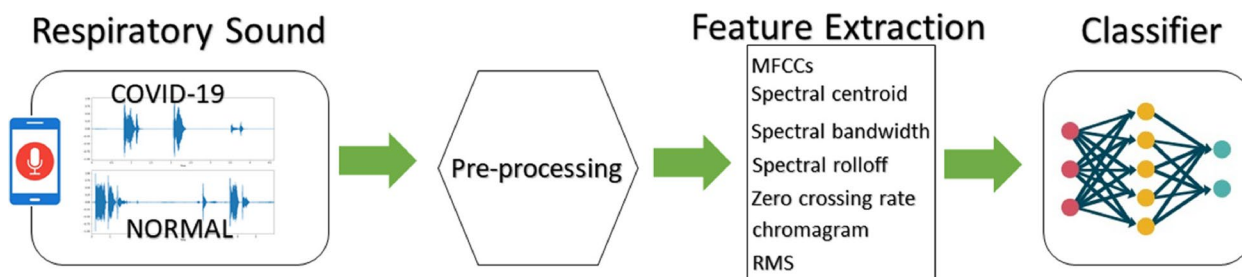


Fig. 5 The process of the proposed model of audio data classification

medical centers and reducing the pressure on the medical model, especially during the peak of the epidemic. This also protects the healthy community that is not infected with COVID-19. The process of the proposed audio data classification model can be seen in Fig. 5. We use audio features including MFCCs, spectral centroid, spectral bandwidth, spectral roll-off, zero crossing rate and RMS and trained the model. We also apply an end-to-end learnable deep neural network model with 186,976 parameters, which consists of 5 fully connected layers using ReLU activation function. In order to avoid overfitting, a dropout layer (0.5) after each connected layer is considered. The optimization function of the model is Adam with an initial learning rate of 0.001. Also, in order to learn the network better, a decreasing learning rate is applied. The number and types of layers used in the proposed model can be seen in Table 2.

Medical image classification

Convolutional neural network is the most reliable deep learning algorithm [11], and it is used to process huge

Table 2 Layer types and parameters used in network

Layer (type)	Output Shape	Param #
dense_1 (input layer)	(None, 512)	13,824
dropout_1	(None, 512)	0
dense_2	(None, 256)	131,328
dropout_2	(None, 256)	0
dense_3	(None, 128)	32,896
dropout_3	(None, 128)	0
dense_4	(None, 64)	8256
dropout_4	(None, 64)	0
dense_5	(None, 10)	650
dropout_5	(None, 10)	0
dense_6	(None, 2)	22
Total params		186,972
Trainable params		186,972
Non-trainable params		0

amount of data with no need for manual extraction of features. The architecture of a convolutional neural network is divided into two sections of learning the feature and

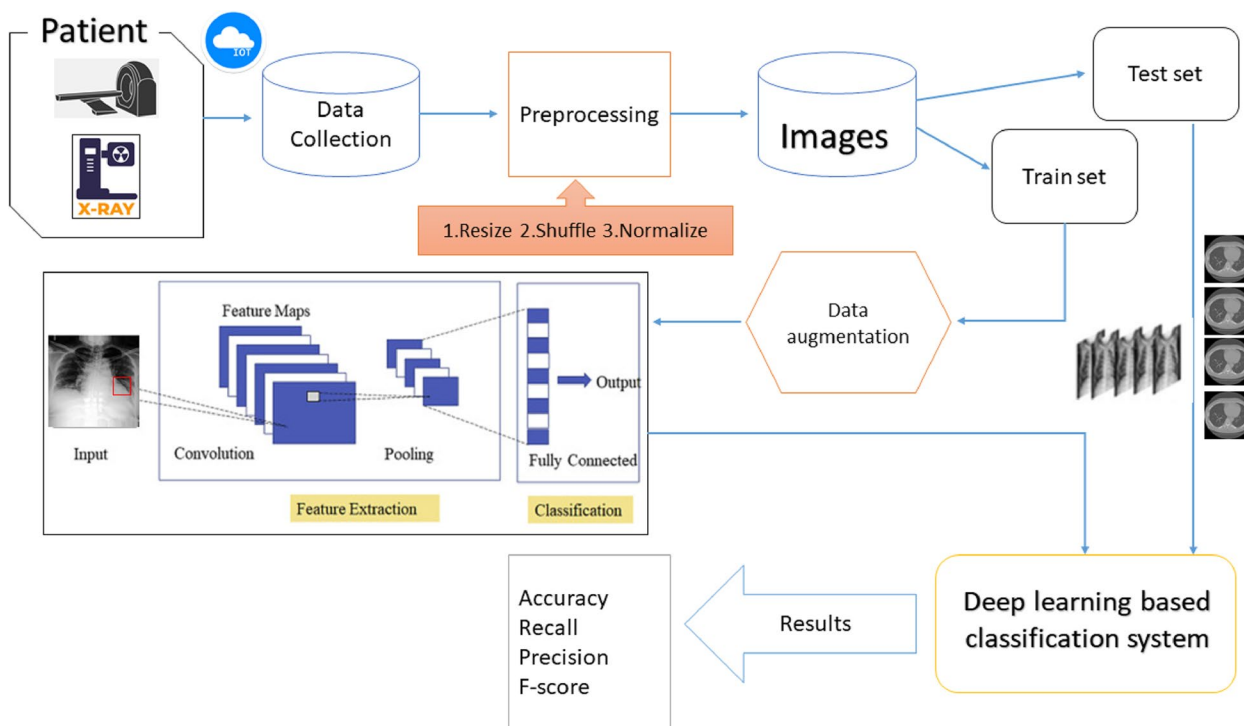


Fig. 6 The process of the proposed model for classifying medical images

classification. In general, these networks are formed of three types of layers: convolutional and pooling layers to extract features and fully connected layer to classify them. The schematic view of the proposed model can be seen in Fig. 6.

The recommended architecture includes four stages. First, people suspected of having COVID-19 who go to medical centers for chest imaging are sent to the dataset in real-time for processing using devices equipped with Internet of Health Things technology. Then, the images of the dataset are pre-processed and in the next step, the training process is done by convolutional neural network. At this stage, three EfficientNet-B4, InceptionV3, and InceptionResNetV2 architectures would be used through transfer learning technique to extract best features. Also, at the end of each architecture, a GlobalAveragePooling2D layer and then three fully connected layers are used for data classification. Also, in order to avoid overfitting, we use the Dropout layer (0.5). At the final stage, recommended networks would be examined by images of test set. In addition, ReLU and sigmoid activation functions have been applied in the network and the weights are generated using the Adam optimizer with initial learning rate of 0.003. Considering the training procedure, learning rate decay also have used. During the training process, there is a point where the output of the model does not improve, for this purpose, the

early stopping technique have used based on the lowest amount of validation error. The structure of the proposed models can be seen in Table 3.

Experimental results

In this section, we evaluate the performance of the proposed method, which is explained in the previous section. First, we describe the evaluation criteria used. Then, the results obtained to diagnose COVID-19 via coughing sound data would be presented. Finally, classification performance of the three proposed models in two datasets of X-ray and CT-Scan medical images would be explained and compared with each other.

Evaluation criteria

We use three standard evaluation criteria: Precision (2), Recall (3) and F-score (4). The cost function formula is also given in Formula 5. Other formulas for evaluation criteria are as follows:¹

- TP = True Positive
- TN = True Negative
- FP = False Positive
- FN = False Negative

¹ Categorical Cross Entropy

Table 3 Layer types and parameters used in networks

Layer (type)	Output Shape	Param #
EfficientNet_B4	(None, 7, 7, 1792)	17,673,823
global_average_pooling2d	(None, 1792)	0
Flatten	(None, 1792)	0
dropout_1	(None, 1792)	0
dense_1	(None, 1024)	1,836,032
dense_2	(None, 1024)	1,049,600
dropout_2	(None, 1024)	0
dense_3	(None, 3)	3075
Total params		20,562,560
Trainable params		2,888,737
non-trainable params		17,673,823
Inception_v3	(None, 5, 5, 2048)	21,802,784
global_average_pooling2d	(None, 2048)	0
Flatten	(None, 2048)	0
dropout_1	(None, 2048)	0
dense_1	(None, 1024)	2,098,176
dense_2	(None, 1024)	1,049,600
dropout_2	(None, 1024)	0
dense_3	(None, 3)	3075
Total params		24,953,665
Trainable params		3,150,881
non-trainable params		21,802,784
Inception_ResNet_V2	(None, 5, 5, 1536)	54,336,736
global_average_pooling2d	(None, 1536)	0
Flatten	(None, 1536)	0
dropout_1	(None, 1536)	0
dense_1	(None, 1024)	1,573,888
dense_2	(None, 1024)	1,049,600
dropout_2	(None, 1024)	0
dense_3	(None, 3)	3075
Total params		56,963,329
Trainable params		2,626,593
non-trainable params		54,336,736

$$Precision = \frac{(TP)}{(TP + FP)} \quad (2)$$

$$Recall = \frac{(TP)}{(TP + FN)} \quad (3)$$

$$F - score = \frac{2 * (precision * recall)}{(precision + recall)} \quad (4)$$

$$CCE = - \sum_{k=0}^{m-1} y_k \log(\hat{y}_k) \quad (5)$$

Audio classification results

We train the proposed model on the allocated data set and evaluate it using the evaluation criteria of the previous section. First, seven different types of features as have been extracted from all data: zero crossing rate, spectral roll-off, spectral bandwidth, spectral centroid, MFCCs, and RMS. Then, %75 of data has used to train the network and %25 remaining percent have used to test network performance. Training has been performed in 200 steps and using a reduced learning rate decay and with consideration of model performance during training in addition to early stopping based on minimum validation loss. The values of loss and accuracy for training and validation sets during network training can be seen in Fig. 7. Also, the values of Precision, Recall and F-score for each class are shown in Table 4. Figure 8 shows the results in form of confusion matrix. As can be observed in Fig. 8, the model has accurately predicted 19 out of 20 cases of the test set, and the accuracy of the test set is 94.999%. Based on the obtained results, screening susceptible people via artificial intelligence and deep learning algorithms and usage made of respiratory audio data is of relatively high accuracy and sensitivity. This can help reducing the pressure on medical staff especially at peak of pandemic and protecting the people not infected yet.

Medical image classification result

The Following criteria have used to evaluate the results obtained from the models: Precision (formula 2), Recall (formula 3), and F-score (formula 4). Three types of data are available in this classification: 1- patients suffering from COVID-19; 2- CAP (community-acquired pneumonia) patients; and, 3- healthy people. Three convolutional neural networks called InceptionV3, EfficientNet-B4, and InceptionResNetV2 have been trained on two datasets of X-ray and CT-Scan images with the help of transfer learning technique. To optimize the network performance, data augmentation technique also has been applied after data pre-processing. Respectively %70, %20, and %10 of data has used to train, test and to validate the network. Training has taken place through 30 steps and using learning rate decay with consideration of the model performance during training along with early stopping based on minimum validation loss value. To prevent over fitting, generalization technique has used as well. Moreover, to achieve best accuracy, hyper-parameters tuning has used on

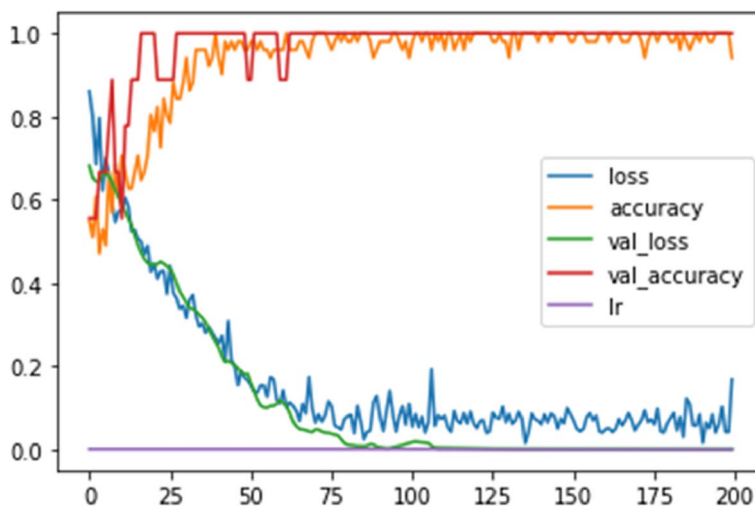


Fig. 7 Loss and accuracy values for training and validation sets during training

Table 4 Class-wise performance results for respiratory sound dataset

Class	Precision	Recall	F-score
COVID-19	0.90	1.00	0.95
NORMAL	1.00	0.91	0.95

models. Loss and accuracy values for test and validation sets during network training can be seen in Fig. 9 for CT-Scan dataset and in Fig. 10 for X-ray dataset. Also, the results obtained in general can be observed in Table 5 and for each class can be observed in Table 6. Best performance for X-ray dataset is jointly related to

EfficientNet-B4 and InceptionV3 with %96.943 accuracy. Although the InceptionV3 model is very fast and converges after 10 epochs of training, the best precision belongs to the efficientNetB4 model with 98%, and the best recall and f1-score belong to the InceptionV3 model with 97%.

The InceptionResnetV2 model demonstrates superior performance in diagnosing individuals with COVID-19 and healthy patients. Notably, the accuracy of all three models reached 100% when diagnosing COVID-19 patients. Furthermore, when identifying pneumonia cases, the EfficientNet-B04 model exhibited exceptional performance with a 99% accuracy and recall, and an impressive 96% f1-score.

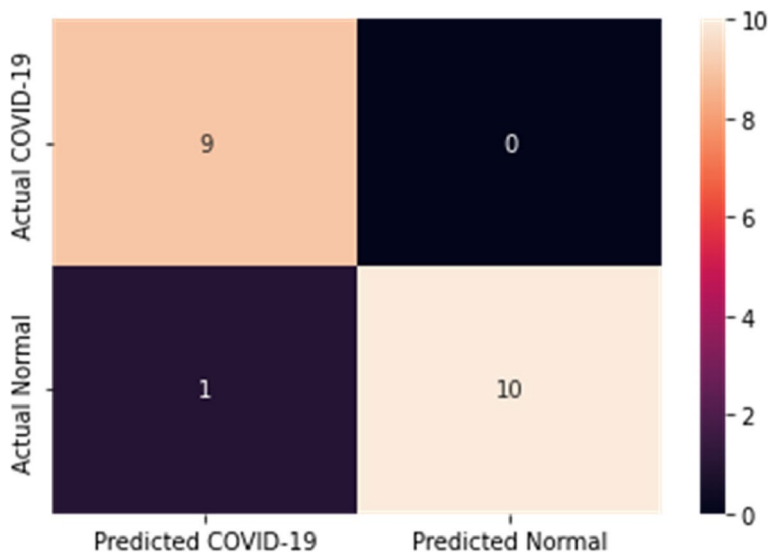


Fig. 8 The confusion matrix of the respiratory sound dataset

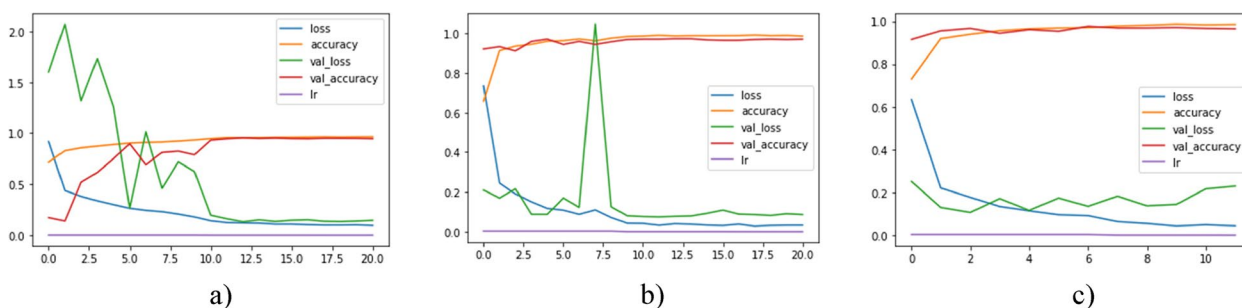


Fig. 9 Loss and accuracy values for training and validation sets during training using X-ray datasets: **a** EfficientNet-B4 network, **b** InceptionResNetV2 network, **c** InceptionV3 network

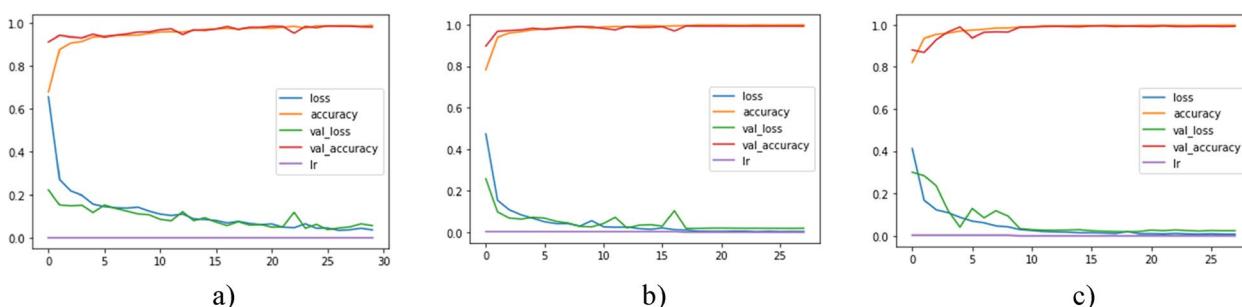


Fig. 10 Loss and accuracy values for training and validation sets during training using CT scan datasets: **a** EfficientNet-B4 network, **b** InceptionResNetV2 network, **c** InceptionV3 network

For the CT-Scan dataset, both the Inception and InceptionResNetV2 models reached convergence faster and completed training in a shorter period. Despite having the same precision, recall, and f1-score values, the highest accuracy is associated with the InceptionResNetV2 architecture with an accuracy of 99.414%, even though all three models detected CAP image accurately. This shows strong efficiency of recommended model to help medical staff to identify infected case at initial stages of the disease. However, based on the findings presented in Table 6, InceptionResNetV2 and InceptionV3 models demonstrated superior performance for COVID-19

Table 5 Performance results obtained from EfficientNet-B4, InceptionResNetV2 and InceptionV3 using X-RAY and CT SCAN datasets

Data type	model	Precision	Recall	F-score	Accuracy
X-ray	EfficientNet-B4	0.98	0.96	0.96	96.943
	Inception-ResNetV2	0.96	0.96	0.96	96.790
	InceptionV3	0.97	0.97	0.97	96.943
CT scan	EfficientNet-B4	0.99	0.99	0.99	99.073
	Inception-ResNetV2	0.99	0.99	0.99	99.414
	InceptionV3	0.99	0.99	0.99	99.366

Table 6 Class-wise performance results for all the studied models using X-RAY and CT SCAN datasets

Data type	model	Class	Precision	Recall	F-score
X-ray	EfficientNet-B4	COVID-19	1.00	0.98	0.99
		Pneumonia	0.99	0.99	0.96
		Normal	0.93	0.93	0.96
	Inception-ResNetV2	COVID-19	1.00	0.99	0.99
		Pneumonia	0.90	0.89	0.94
		Normal	0.99	1.00	0.95
InceptionV3	COVID-19	1.00	0.98	0.99	
	Pneumonia	0.94	0.94	0.96	
	Normal	0.98	0.98	0.96	
CT scan	EfficientNet-B4	COVID-19	0.99	0.99	0.99
		CAP	1.00	1.00	1.00
		Normal	0.98	0.99	0.98
	Inception-ResNetV2	COVID-19	0.99	1.00	0.99
		CAP	1.00	1.00	1.00
		Normal	0.99	0.98	0.99
InceptionV3	COVID-19	0.99	1.00	0.99	
	CAP	1.00	1.00	1.00	
	Normal	0.99	0.97	0.98	

patients, outperforming other models. Additionally, InceptionResNetV2 and EfficientNet-B04 models exhibited higher accuracy for healthy individuals.

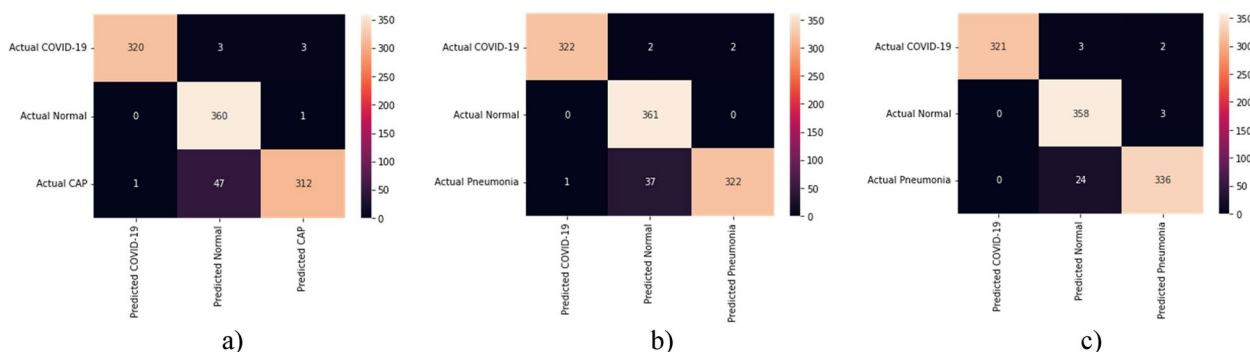


Fig. 11 X-ray dataset confusion matrix: **a** EfficientNet-B4 network, **b** InceptionResNetV2 network, **c** InceptionV3 network

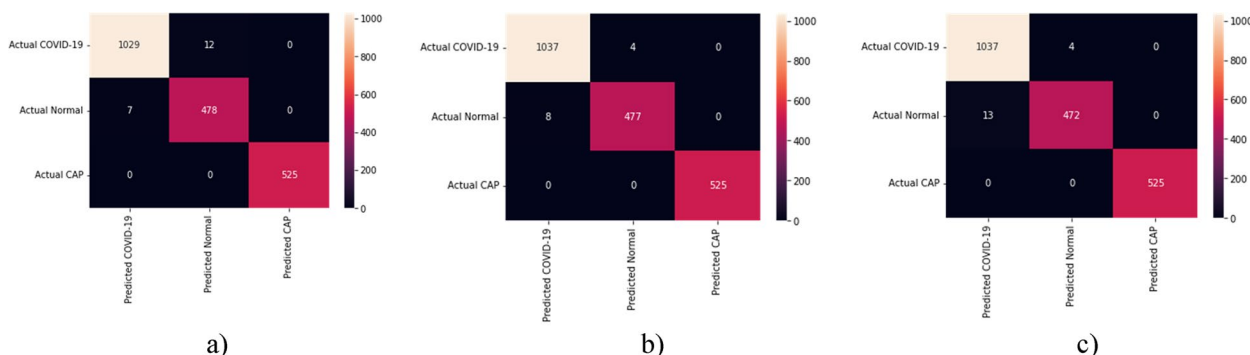


Fig. 12 CT-scan dataset confusion matrix: **a** EfficientNet-B4 network, **b** InceptionResNetV2 network, **c** InceptionV3 network

Also, considering the results obtained from the two datasets, it can be concluded that CT-Scan images have more details compared to X-ray images for diagnosis of lung diseases. Moreover, InceptionV3 has learned and has reached to early stopping point faster than other models; so, it requires less training epochs. Figure 11 shows classification results of X-ray images in form of confusion matrix. And Fig. 12 is related to CT-Scan images.

Performance evaluation

We perform a comparative analysis to demonstrate the effectiveness of the proposed model. We compare the results obtained by the proposed model with some techniques in Table 1 ([1, 2, 5–7, 9–12]). In addition, for a better judgment, the number of images of the sets used in each research is also mentioned. Table 7 compares the research of detecting COVID-19 based on CXR images with the proposed model in this research, and Table 8 compares the research of detecting COVID-19 based on CT-Scan images with the proposed model in this research. Also, we compare the data sets using the networks used in the related work during 30 steps with a training rate of 0.001 and the

results obtained from the accuracy of the proposed models and related to the set of radiographic images in Table 9. The results obtained from the set CT-Scan images can be seen in Table 10.

According to the results, it can be suggested that recommended approach accurately classifies images in all classes of COVID-19 patients, pneumonia patients, and healthy people. Therefore, this can be concluded that accurate adjustment of pre-trained CNN architecture can be use as one of the useful techniques for classification of chest X-ray images in the field of medicine.

Conclusion and discussion

Discussion

The diagnosis of COVID-19 through chest medical images is a challenging issue requiring to be overcome. In this paper, an intelligent health care model has been proposed that supports IoHT technologies for initial evaluation of COVID-19 through usage made of neural network with the help of two types of audio data, and it has been presented to screen people who suspected of COVID-19 infection as well as chest medical images. This model uses intelligent sensors to collect data.

Table 7 Comparison X-ray results with some deep learning-based methods

Reference	Network	Data type	Dataset size	Result
[1]	VGG19	CXR	4500	%86
	InceptionV3			%95
[2]	Ensemble model	CXR	15,688	%88
[9]	AlexNet	CXR	5436	%86
	VGG16			%87
	ResNet50			%91
	ResNet101			%93
[11]	ResNet152	CXR	380	%95
	VGG16			%85
	VGG19			%89
	ResNet18			%88
	ResNet50 + SVM			%94
[12]	ResNet101	CXR	10,508	%87
	EfficientNet			%96
[7]	Proposed Model	CXR	9545	%97
	ResNet50			%96
	ResNet101			%95
	GoogleNet			%96
	AlexNet			%93
	DenseNet201			%96
	The proposed model			EfficientNet-B4
	InceptionV3			%96.9
	InceptionResNetV2			%96.7

These data are stored in data repositories and used to evaluate the condition of patients. At first stage, audio data of respiratory sound have used to screen people who suspected of COVID-19 infection so that unnecessary visits to medical centers would be prevented.

Then in the next stage, medical images are sent to deep learning network so that COVID-19 would be identified.

Using technological capacities such as IoHT and artificial intelligence under critical conditions such as COVID-19 pandemic in addition to reduction of epidemic rate of the virus and protection of healthy community's health are of high importance with consideration of shortage of medical staff compared to daily references to medical centers. It is important to consider that deep convolutional

Table 8 Comparison CT-Scan results with some deep learning-based methods

Reference	Network	Data type	Dataset size	Result
[5]	ResNet50V2	CT Scan	2484	%95
	Xception			%94
	DenseNet121			%93
	Stacked model			%96
[6]	Proposed CNN	CT Scan	6354	%98
[7]	Proposed Model	CT Scan	8055	%97
	ResNet50			%96
	ResNet101			%95
	GoogleNet			%96
	AlexNet			%93
	DenseNet201			%96
[10]	Proposed Model	CT Scan	4982	%94
The proposed model	EfficientNet-B4	CT Scan	10,239	%99
	InceptionV3			%99.3
	InceptionResNetV2			%99.4

Table 9 Comparison of results obtained by the X-ray dataset with some deep learning-based methods

Reference	network	Layer #	Param #	Result
[11]	VGG16	16	Over 138 million	%96.4
[1]	VGG19	19	Over 143 million	%94.4
[7]	ResNet50	50	Over 23 million	%94.6
[9]	ResNet152	152	Over 60 million	%94.5
The proposed model	EfficientNet-B4	258	Over 19 million	%96.9
	InceptionV3	189	Over 23 million	%96.9
	InceptionResNetV2	449	Over 55 million	%96.7

neural network architectures with more steps of training show tendency towards over fitting. To prevent this, such methods like early stopping and data reinforcement have used by us; however, using dropout method in most cases is considered as an appropriate solution.

Proposed model has presented best accuracy of %94.999 for classification of audio data and %96.943 for chest X-ray images, and %99.414 for chest CT-Scan images.

Conclusion

The Internet of health things (IoHT) is an integrated platform to facilitate interactions between humans and various types of physical and virtual platforms. Considering critical condition of COVID-19 pandemic, it can play a vital role in the field of medical care, resulting in reduction of pressure on medical model. With consideration of technological progress, Nowadays the Internet of Medical Things along with artificial intelligence techniques such as machine learning and deep learning have provided new facilities covering up a wide range of functions in the field of medical care. Medical devices and sensors can collect valuable data through connection with internet so that at next stages and with the help of artificial intelligence techniques they would be processed, and their knowledge can extracted.

The available data are not considered as strong yet. However, experimental knowledge about convolutional neural networks applications is indicative of increase of

number of samples and quality of dataset having direct effect on accuracy obtained. The proposed method based on deep learning will be useful in medical diagnostic research and health care models. It is also a precise tool for medical experts to do COVID-19 screening, leading to secondary medical view.

Although these technologies have great potential to improve the treatment and diagnosis of diseases, they also face important challenges. One of these challenges is the diversity of data available in healthcare environments. Different patients with varied information will complicate modeling. A research method that can help address this issue is heterogeneous embedding learning. This method allows mapping diverse data into common vectors, leading to a better understanding of patterns present in the data. Additionally, privacy and security of medical data are crucial when utilizing these technologies. Appropriate solutions must be considered to protect this data, earn patients' trust, and enable the use of these tools in healthcare settings.

In future works, more images would be collected and deeper models for COVID-19 diagnosis would be studied. Due to current emergency status of public health, collecting big datasets is of high importance to train deep learning model. Moreover, other lung diseases would be also included in future studies. The development of a graphic interface to help radiologists in detection of COVID-19 can be aimed at in future studies. However, an IoHT-based model capable of

Table 10 Comparison of results obtained by the CT-Scan dataset with some deep learning-based methods

Reference	network	Layers #	Param #	Result
[11]	VGG16	16	Over 138 million	%91.7
[1]	VGG19	19	Over 143 million	%98.7
[7]	ResNet50	50	Over 23 million	%99
[9]	ResNet152	152	Over 60 million	%99.3
The proposed model	EfficientNet-B4	258	Over 19 million	%99
	InceptionV3	189	Over 23 million	%99.3
	InceptionResNetV2	449	Over 55 million	%99.4

producing high volume of dataset can be a great help to medical model. On the contrary, conducting a comprehensive examination and analysis of audio data features to assess the significance of various feature types for the classification of medical images constitutes a pivotal aspect within the realm of medical signal processing. The ultimate aim is to enhance the precision and efficacy of medical image classification. By meticulously scrutinizing the audio data features, this approach delves into harnessing these features optimally for medical image classification, thus representing a crucial stride towards enhancing the precision and efficacy in this domain. Also, the use of methods including deep clustering and graph representation learning is suggested as the next step in IoT-based AI studies.

Acknowledgements

Not applicable

Authors' contributions

Conceptualization, M. Mousavi; methodology, S.Hosseini.; software, M. Mousavi; validation, S.Hosseini and M. Mousavi; formal analysis, S.Hosseini and M. Mousavi; investigation, S.Hosseini; resources, M. Mousavi; data curation, S.Hosseini.; writing—original draft preparation, M. Mousavi; writing—review and editing, S.Hosseini.; visualization, M. Mousavi.; supervision, S.Hosseini.; project administration, S.Hosseini.

Funding

This research received no external funding.

Availability of data and materials

The datasets analyzed during the current study are publicly available in the, https://www.kaggle.com/datasets/himanshu007121/coughclassifier-trial?select=trial_COVID, <https://www.kaggle.com/datasets/drsurabhthorat/COVID-19-ct-scan-dataset>.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 15 April 2024 Accepted: 22 August 2024

Published online: 29 August 2024

References

- Ahmed I, Jeon G, Chehri A. An IoT-enabled smart health care system for screening of COVID-19 with multi layers features fusion and selection. *Computing*. 2023;105(4):743–60.
- Deb. S. D, Jha. R. K, K, and Tripathi. P. S, .A multi model ensemble based deep convolution neural network structure for detection of COVID19. *Biomed Sign Process Control*. 2022;71:103–126.
- Lella KK, Pja A. Automatic diagnosis of COVID-19 disease using deep convolutional neural network with multi-feature channel from respiratory sound data: cough, voice, and breath. *Alexandria Eng J*. 2022;612:1319–34.
- Chowdhury NK, Kabir MA, Rahman MM, Islam SMS. Machine learning for detecting COVID-19 from cough sounds: An ensemble-based MCDM method. *Comput Biol Med*. 2022;145:105405.
- Shorfuzzaman M. IoT-enabled stacked ensemble of deep neural networks for the diagnosis of COVID-19 using chest CT scans. *Comput*. 2021:1–22.
- Thakur S, Kumar A. X-ray and CT-scan-based automated detection and classification of COVID-19 using convolutional neural networks (CNN). *Biomed Signal Process Control*. 2021;69:102–10.
- Elpeltagy M, Sallam H. Automatic prediction of COVID– 19 from chest images using modified ResNet50. *Multimed Tools Appl*. 2021;80(17):26451–63.
- Loey M, Mirjalili S. COVID-19 cough sound symptoms classification from scalogram image representation using deep learning models. *Comput Biol Med*. 2021;139:105–16.
- Lorencin I, et al. Automatic Evaluation of the Lung Condition of COVID-19 Patients Using X-ray Images and Convolutional Neural Networks. *J Personal Med*. 2021;11(1):28–59.
- Ouyang X, et al. Dual-sampling attention network for diagnosis of COVID-19 from community acquired pneumonia. *IEEE Trans Med Imaging*. 2020;39(8):2595–605.
- Ismael AM and Şengür A. Deep learning approaches for COVID-19 detection based on chest X-ray images. *Expert Syst Appl* 2021;164:114–125.
- Marques G, Agarwal D, de la Torre Díez I. Automated medical diagnosis of COVID-19 through EfficientNet convolutional neural network. *Appl Soft Comput*. 2020;96:106–17.
- Cohen JP, Morrison P, Dao L, Roth K, Duong TQ, Ghassemi M. Covid-19 image data collection: Prospective predictions are the future. *arXiv preprint arXiv:200611988* 2020.
- Brown. C, Chauhan. J, Grammenos. A, Han. J, Hasthanasombat. A, Spathis. D, Xia. T, Cicuta. P, Mascolo. C, “Exploring automatic diagnosis of COVID-19 from crowdsourced respiratory sound data”, in: *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2020:3474–3484, <https://doi.org/10.1145/3394486.3412865>.
- Sharma. N, Krishnan. P, Kumar. R, Ramoji. S, Chetupalli. S.R, N. R, Ghosh. PK, Ganapathy. Coswara. S, — a Database of breathing, cough, and voice sounds for COVID-19 diagnosis, in: *Proc. Interspeech 2020*, 2020, pp. 4811–4815, <https://doi.org/10.21437/Interspeech.2020-2768>.
- Orlandic L, Teijeiro T, Atienza D. The COUGHVID crowdsourcing dataset, a corpus for the study of large-scale cough analysis algorithms. *Sci Data*. 2021;8(1):156.
- Cohen-McFarlane M, Goubran R, Knoefel F. Novel Coronavirus cough Database: NoCoCoDa. *IEEE Access*. 2020;8:154087–94. <https://doi.org/10.1109/ACCESS.2020.3018028>.
- Extensive COVID-19 X–Ray and CT Chest Images Dataset. [Online]. Available: <https://doi.org/10.17632/8h65ywd2jr3>.
- COVID-19 image data collection. [Online]. Available: <https://github.com/ieee8023/COVID-chestxray-dataset/tree/master/images>.
- chest-xray-pneumonia. [Online]. Available: <https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia>.
- COVID19+PNEUMONIA+NORMAL Chest X-Ray Images. [Online]. Available: <https://www.kaggle.com/sachinkumar413/COVID-pneumonia-normal-chest-xray-images>.
- Ahmed I, Ahmad A, Jeon G. An IoT-based deep learning framework for early assessment of COVID-19. *IEEE Internet Things J*. 2020;8(21):15855–62.
- Zhao B-W, He Y-Z, Su X-R, Yang Y, Li G-D, Huang Y-A, Hu P-W, You Z-H, Hu L. Motif-Aware miRNA-Disease Association Prediction Via Hierarchical Attention Network. *IEEE J Biomed Health Inform*. 2024;28:4281–94.
- Yang Y, Su X, Zhao B, Li G, Hu P, Zhang J, Hu L. Fuzzy-based deep attributed graph clustering. *IEEE Trans Fuzzy Syst*. 2023;32:1951–64.
- Zhao B-W, Wang L, Hu P-W, Wong L, Su X-R, Wang B-Q, You Z-H, Hu L. Fusing higher and lower-order biological information for drug repositioning via graph representation learning. *IEEE Trans Emerg Top Comput*. 2023;12(1):163–76.
- Zhao B-W, Su X-R, Yang Y, Li D-X, Li G-D, Hu P-W, Luo X, Hu L. A heterogeneous information network learning model with neighborhood-level structural representation for predicting lncRNA-miRNA interactions. *Comput Struct Biotechnol J*. 2024;23:2924–33.
- COVID-19 dry cough and augmented spectrograms. [Online]. Available: <https://www.kaggle.com/datasets/juanmiguellopez/COVID19-dry-cough-and-augmented-spectrograms>.

28. COVID-19 Cough Recordings [Online]. Available: https://www.kaggle.com/datasets/himanshu007121/coughclassifier-trial?select=trial_COVID.
29. McFee, B, Raffel, C, Liang, D, Ellis, D, McVicar, M, Battenberg, E, Nieto, O, "Librosa: audio and music signal analysis in Python", in: Proceedings of the 14th Python in Science Conference, SciPy, 2015, <https://doi.org/10.25080/majora-7b98e3ed003>.
30. COVID 19 CT Scan Dataset [Online]. Available: <https://www.kaggle.com/datasets/drsurabhithorat/COVID-19-ct-scan-dataset>.
31. Kashani MH, Madanipour M, Nikravan M, Asghari P, Mahdipour E. A systematic review of IoT in healthcare: Applications, techniques, and trends. *J Netw Comput Appl.* 2021;192:103–44.
32. Zhu H, et al. IoT PCR for pandemic disease detection and its spread monitoring. *Sens Actuators, B Chem.* 2020;303:127–34.
33. Yaqoob T, Abbas H, Atiquzzaman M. Security vulnerabilities, attacks, countermeasures, and regulations of networked medical devices—A review. *IEEE Commun Survays Tutorals.* 2019;21(4):3723–68.
34. Adil M and M Khan. K. Emerging iot applications in sustainable smart cities for COVID-19: Network security and data preservation challenges with future directions. *Sustainable Cities Soc* 2021;75:103–115.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.