A novel multi-scale based deep convolutional neural network for detecting COVID-19 from X-rays

Mohan Karnati [a], Ayan Seal [a,*], Geet Sahu [a], Anis Yazidi [b,c,d], Ondrej Krejcar [e,f]

[a] *Department of Computer Science and Engineering, PDPM Indian Institute of Information Technology Design & Manufacturing Jabalpur, Jabalpur, Madhya Pradesh 482005, India*
[b] *Department of Computer Science, OsloMet – Oslo Metropolitan University, Oslo, 460167, Norway*
[c] *Department of Computer Science, Norwegian University of Science and Technology, Trondheim, 460167, Norway*
[d] *Department of Plastic and Reconstructive Surgery, Oslo University Hospital, Oslo, 460167, Norway*
[e] *Center for Basic and Applied Science, Faculty of Informatics and Management, University of Hradec Kralove, Rokitanskeho 62, 500 03 Hradec Kralove, Czech Republic*
[f] *Malaysia-Japan International Institute of Technology (MJIIT), Universiti Teknologi Malaysia, Jalan Sultan Yahya Petra, 54100 Kuala Lumpur, Malaysia*

ARTICLE INFO

ABSTRACT

The COVID-19 pandemic has posed an unprecedented threat to the global public health system, primarily infecting the airway epithelial cells in the respiratory tract. Chest X-ray (CXR) is widely available, faster, and less expensive therefore it is preferred to monitor the lungs for COVID-19 diagnosis over other techniques such as molecular test, antigen test, antibody test, and chest computed tomography (CT). As the pandemic continues to reveal the limitations of our current ecosystems, researchers are coming together to share their knowledge and experience in order to develop new systems to tackle it. In this work, an end-to-end IoT infrastructure is designed and built to diagnose patients remotely in the case of a pandemic, limiting COVID-19 dissemination while also improving measurement science. The proposed framework comprises six steps. In the last step, a model is designed to interpret CXR images and intelligently measure the severity of COVID-19 lung infections using a novel deep neural network (DNN). The proposed DNN employs multi-scale sampling filters to extract reliable and noise-invariant features from a variety of image patches. Experiments are conducted on five publicly available databases, including COVIDx, COVID-19 Radiography, COVID-XRay-5K, COVID-19-CXR, and COVIDchestxray, with classification accuracies of 96.01%, 99.62%, 99.22%, 98.83%, and 100%, and testing times of 0.541, 0.692, 1.28, 0.461, and 0.202 s, respectively. The obtained results show that the proposed model surpasses fourteen baseline techniques. As a result, the newly developed model could be utilized to evaluate treatment efficacy, particularly in remote locations.

1. Introduction

COVID-19 was originally identified in China in December 2019 and has infected over a hundred million people around the world. The World Health Organization (WHO) declared a pandemic on March 11, 2020. In almost 74% of the cases, the infections are either minor (18%) or severe symptoms (56%). However, the remaining 26% vary from critical (20%) to an extreme symptoms (6%) [1]. As of today (28/05/2021), the world's cumulative number of COVID-19 infections is more than 169 million, and the death toll overpasses 3.52 million, while 151 million cases recovered completely. Moreover, the number of active instances is 14.74 million, among which 14,648,154 events are in minor condition, and 93,863 events are in serious condition [2]. Table 1 summarizes some major statistical parameters related to the pandemic COVID-19 in several countries. The novel COVID-19 disease emerges with throat inflammation, fever, and respiratory distress, then progresses to breathing difficulties. The infection could cause the severe acute respiratory syndrome, pulmonary hypertension, organ failure, and, ultimately, death of the patient [3]. Recent studies suggest that men are more likely to get affected than women. In this perspective, men represent 60% of the cases, and there were no reported substantial mortality rates among children younger than nine years [4]. Furthermore, COVID-19 infected patients must isolate themselves and adopt appropriate preventive steps to safeguard healthy individuals, thereby breaking the infection chain [4,5]. Historical data have shown that the infection rate grows exponentially rather than linearly if preventive measures are not utilized effectively, and

* Corresponding author.
 E-mail address: ayan@iiitdmj.ac.in (A. Seal).

Table 1
Statistics on COVID-19 outbreaks.

Countries	Confirmed	Deaths	Recovered
USA	33,999,680	607,726	27,701,879
India	27,555,457	318,895	24,893,410
Brazil	16,342,162	456,753	14,786,292
Russia	5,044,459	120,406	4,661,234
UK	4,473,677	127,758	4,310,572
France	5,635,629	109,165	5,284,264
Turkey	5,220,549	46,970	5,070,815
Germany	3,673,969	88,689	3,461,700
Italy	4,205,970	125,793	3,826,984

in some cases, the pandemic may reach a tipping point beyond which the infection rate becomes uncontrollable. In many circumstances, it would put a strain on the limited medical resources available for diagnosis. COVID-19 is diagnosed using at least one of the three methods listed below:

- RT-PCR: For antigen detection testing, [6,7] uses a nose blood sample and a venous blood sample. In some countries, such as India, these procedures necessitate contact between patients and physicians, which might take anything from a few hours to three days to receive results. Some studies have found that the results of numerous RT-PCR tests performed at different times for the same patient can differ, resulting in a high false negative (FN) diagnostic rate [8]. Many researchers suggested combining the RT-PCR test with additional clinical exams, such as computed tomography (CT), to improve the accuracy of the diagnosis.
- CT scan: COVID-19 patients mostly become infected with lung disease at an early stage of the disease. The most prevalent COVID-19 lung symptoms are consolidation, i.e. fluid and bone accumulations in lung blood vessels that prevent ground-glass opacity, gas exchange, and nodular shadowing. These symptoms are frequently present in the middle and lower lung regions and can be used to distinguish between people infected with non-COVID-19 and COVID-19 [9,10]. In comparison to RT-PCR, CT equipment generates images for faster COVID-19 screening [11]. CT scan-based measurement entails assessing 3D radiographic imaging of the lungs from multiple perspectives. Manual examination of COVID-19 from chest CT scans, on the other hand, is a labor-intensive, and time-consuming procedure since medical practitioners must find lesions slice-by-slice from volumetric CT images.
- Chest X-ray (CXR): In comparison to CT, CXR equipment is smaller and more portable. In hospitals, this type of resource is usually more accessible than RT-PCR and CT-scan. Furthermore, because the CXR test lasts around 15 s per subject [12], it is one of the most cost-effective pieces of evaluation equipment.

In medical treatment, a reliable computer-aided diagnostic system that analyzes CXR for precise, rapid screening and diagnosis of COVID-19 patients is required, reducing the workload on the medical staff. However, such a diagnosis is difficult to automate because CXR images of pneumonia exhibit similar types of defects in the lung territories. Therefore, relying only on classical computer vision techniques which are based on hand-crafted descriptors might be deemed to failure due to the difficulty to handle the distinctive features of pneumonia targets.

1.1. Motivation and contribution

In recent years, tremendous progress has been made in measurement science by applying deep neural networks (DNN) techniques to computer vision applications such as salient object detection [13,14], facial expression recognition [15,16], and deception detection [17], thus DNN models have become the defacto-standard nowadays. DNN has specialized in learning-rich images with high-level discriminatory semantic characteristics automatically, eliminating the need for hand-crafted descriptors. These breakthroughs have revealed that deeper models can improve results [16]. Thus, it is viable to train a DNN model to obtain promising performance in COVID-19 screening and monitoring. Moreover, technological advancement has enabled the manufacturing of low-cost portable computing devices for consumers. Cellular devices have advanced in terms of technical capabilities and processing power, and they have become a source of information, interaction, and sharing. They are now almost indispensable in our daily lives. Internet of Things (IoT) with cellular devices have permitted a far wider range of uses, not only for entertainment but also for the treatment and monitoring of health requirements, environmental surveillance, home automation, and many more [18]. Therefore, the motivation for this study is twofold. Firstly, there is a lack of resources and screening tools for identifying and monitoring COVID-19 patients, and secondly, DNN has a great potential for fetching features and accurately classifying images without any manual intervention. This work introduces a framework that includes a novel DNN enabled IoT service to intelligently measure the severity of COVID-19 lung infections by analyzing CXR images. The proposed DNN module consists of multi-scale sampling filters that allow extracting more reliable and noise-invariant features at different image patches. We have circumvent the shortcomings of the existing DNN models and achieve superior performance by carefully designing the proposed DNN model-based multi-scale sampling. All the experiments are implemented on five databases, namely COVIDx (D1) [19], COVID-19 Radiography (D2) [20], COVID-XRay-5K (D3) [21], COVID-19-CXR (D4) [22], and COVID-chestxray (D5) [23]. The proposed framework is compared with fourteen existing approaches by utilizing four well-known classification metrics viz., F1-score, recall, precision, and accuracy. Empirical evidence manifests that the proposed method outranks all the fourteen existing approaches. The integration of the proposed algorithm with an IoT framework results in an efficient and precise real-time online service for COVID-19 diagnosis. The contributions of this study can be summarized as follows:

- A detection and monitoring tool for the diagnosis of COVID-19 patients is introduced. This framework is instrumented with an IoT system that helps to oversee both potential and real cases. Thus, the newly developed equipment can be employed to observe patients efficiently, especially in a remote location.
- A novel DNN framework is designed for distinguishing non-COVID-19 from COVID-19 classes using CXR images. The use of X-ray simplifies the implementation of the proposed method in real-world scenarios. When compared to other testing procedures, X-rays are less expensive and take less time.
- The proposed DNN consists of multi-scale filters. The strength of multi-scale sampling filters to fetch robust and noise invariant facets with distinguishing power is exploited. We hypothesize that by integrating multi-scale feature extraction, we can learn more resilient convolutional filters since the scale of features varies substantially among distinct ground objects captured from several sensing devices. Moreover, our proposed DNN is simple as it has fewer layers and learning parameters.
- We give insights into the theoretical enhancement made to the DNN model and document their empowering effect through experiments. The experimental results illustrate that the computation cost is considerably lower compared with some related approaches. This confirms that our approach is more computationally efficient.

The organization of the current study is structured as follows: A concise summary of a few notable previous approaches related to COVID-19 classification is put forward in Section 2. Section 3 describes the proposed work in depth. The obtained outcomes of the proposed model along with other baseline approaches are reported in Section 4. At last, Section 5 concludes the current study.

2. Related work

Some recent developments in diagnosis the COVID-19 utilizing machine learning (ML) and deep learning (DL) techniques are thoroughly described in this section.

In the field of medical imaging, DL techniques have classically discovered a large set of applications ranging from diabetic retinopathy, histological examination, cardiac imaging, tumor detection, to mention a few. An emerging application of DL is to diagnose COVID-19 using CXR images, CT-scans [24], etc. Several researchers have published a set of pre-print papers tackling the problem of COVID-19 detection from CXR [19]. The reported results from the latter study document promising outcomes, however, they are based on small databases, which is far from the real implementation. These solutions would need to be thoroughly tested and improved before they could be put into use.

Typically, researchers rely on DL techniques to classify any specific characteristics of COVID-19 patients from CXR images. DL is known to be efficient in the detection of different lung-related diseases based on chest radiography images. A plethora of legacy studies applying ML and DL algorithms to analyze the X-ray and CT images can be found in the literature [19, 25,26]. With the upsurge of the COVID-19, many recent pieces of research have investigated the usefulness of radiological images for COVID-19 detection. In [19], Wang et al. presented a COVID-Net for COVID-19 detection. Further, Hemdan et al. [26] suggested an alternative approach named COVIDX-Net, consisting of seven DNN variants to detect COVID-19 from CXR images. However, these methods suffer from overfitting problem and are hard to implement in real-time applications since it has a larger network. To make it evident, the training and validation losses obtained by these two methods are analyzed and it is observed that the gap between training and validation losses is greater. Ohata et al. [27] employed transfer learning to further train various pre-trained DL models to fetch facets and accurately predict COVID-19. Tabik et al. [28] suggested a COVID-SDNet, which consists of several DNN networks for COVID-19 classification. However, these methods are very time-consuming. Arias et al. [29] presented an automatic detection of COVID-19 (AD-COVID19) using DNN with a segmentation approach. In [30], Wang et al. used a prior residual learning approach for classifying COVID-19 robustly. However, it has a large number of parameters, hence it is hard to implement in real-time applications, especially in health care for monitoring COVID-19 patients. Khan et al. [22] utilized pre-trained XCeption architecture and further trained it on CXR images of COVID-19 and other chest pneumonia from two separate publicly accessible databases. Furthermore, Jian et al. [25] presented a DL model that directly used pre-trained models like ResNeXt, XCeption, and Inception-V3 for COVID-19 detection from CXR images. Similarly, Apostolopoulos et al. [31] used a transfer learning approach with VGG-19 and MobileNet. These methods, have a large number of parameters and require complex computational resources to train. In [32], DarkCovid-Net is presented for classification and detection of COVID-19. In [33], five pre-trained models, namely ResNet50, ResNet101, ResNet152, InceptionV3 and Inception-ResNetV2, are employed for classifying COVID-19 effectively.

Some researchers have presented IoT-based diagnosis systems that collect relevant sensor data and process it in the cloud. With the advent of IoT, it has become a critical component of many environmental monitoring and healthcare applications.

In [34], Nguyen provided a review of the artificial intelligence (AI) methods used in COVID-19 analysis. These approaches were divided into several categories, including the use of IoT. Maghdid [35] explored that sensors on smartphones can be used to acquire health information such as temperature. Rao and Vazquez [36] investigated the utility of ML techniques on user data gathered via a web-based survey obtained from smartphones for quick COVID-19 screening. In [37], Allam and Jones suggested a method to detect potential COVID-19 patients using images from a thermal camera. Otoom et al. [38] presented an IoT-based real-time detection, observation, and inspection of COVID-19 system using eight ML algorithms, namely k-nearest neighbor (KNN), support vector machine (SVM), artificial neural network (ANN), Naive Bayes, decision stump, decision table, one rule (OneR), and zero rule (ZeroR). Zhang et al. [39] presented a residual learning diagnosis detection (RLDD) system for COVID-19 classification. A residual block was used in this method to train a DNN, which is quite large, therefore complex calculations are required. Furthermore, they presented an industrial IoT framework, but no comprehensive definition of how or where it should be implemented was provided. However, the performance of these methods falls short on small databases.

3. Proposed method

This section offers a brief overview of the proposed DL-based IoT service for evaluating CXR images and diagnosing COVID-19 effectively. Due to the reliance on classical ML approaches on human skill for feature creation, as well as DL advancements in the domain of computer vision, we propose a DL model for automatic feature engineering in this study. We will also demonstrate how our DL-based algorithm can be linked to an IoT service to create a complete diagnosis chain.

3.1. The IoT framework

Social distancing is a non-pharmaceutical method of prevention. When we are forced to stay locked up in our homes, the IoT revolution plays an important role in modern healthcare systems in terms of professional, social, and economic prospects. As a result, in the context of the current pandemic, IoT-enabled applications can be used to reduce the potential spread of COVID-19 through early and remote diagnosis. As a result, the present study introduces an end-to-end IoT framework to help virtually the patients in remote locations in the event of a pandemic. The challenges associated with each layer of the proposed framework are addressed, and design guidelines for dealing with them are discussed. This sub-section describes the developed IoT-based framework for observing and recognizing COVID-19 cases. This framework can also be used to track how well-reported patients respond to treatment and learn more about the COVID-19 disease. The proposed IoT framework is shown in Fig. 1, consisting of six steps labeled from 1–6. The doctor can upload a COVID-19 X-ray image or a group of images to an internet application from this screen. This method will extract information from images and classify them as non-COVID-19 or COVID-19. The proposed method extracts features from an image using the DNN model, followed by a softmax classifier that uses the extracted features as inputs to classify COVID-19. This method makes use of the LINDA, which is available as a web service. It consists of a processing flow that can (i) extract, (ii) train, (iii) predict, and (iv) store the statistics and results obtained for COVID-19 recognition. All computational processing for this IoT system is done in the cloud. The server is housed at the Instituto Federal de Educaço, Ciência,
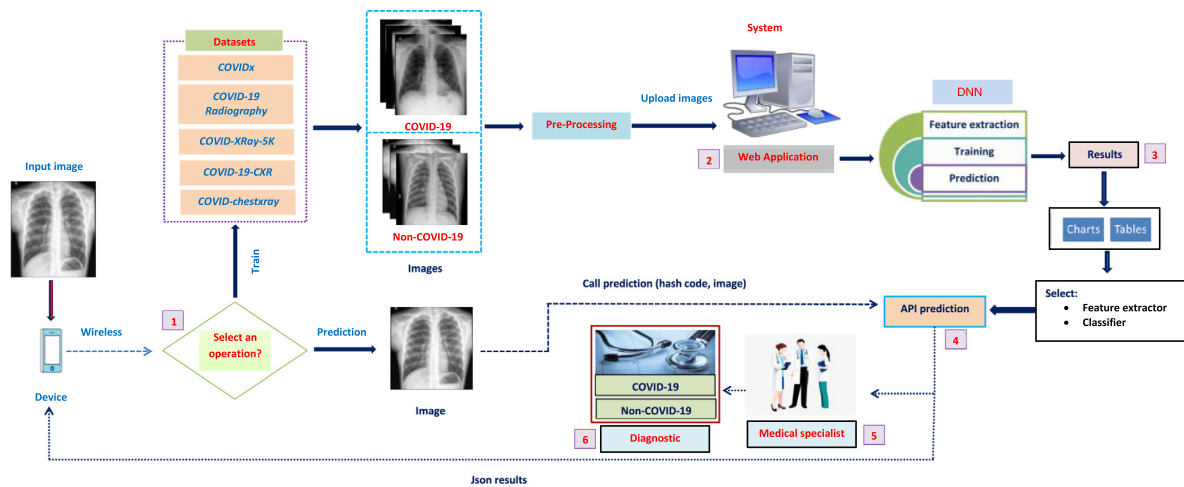
Fig. 1. Schematic illustration of the LINDA approach for IoT-based COVID-19 detection and classification in CXR images. (1) new images can be sent for training or prediction using the same software. (2) users can define system parameters using a web application; (3) training results are displayed in plots and tables; and (4) API prediction is in charge of image processing based on the parameters set in the web application, (5) medical analysis by the doctors, and, finally, (6) Diagnostic process by the proposed DNN model.
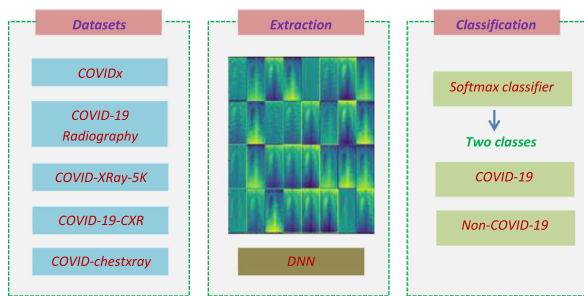

Fig. 2. Diagram showing the processing steps starting from five X-ray databases, extracting facets of the COVID-19 X-ray images based on DNN, and the classification step with softmax classifiers.

e Tecnologia do Ceará. LINDA has recently gained popularity, and it has been used to develop not only medical IoT services such as stroke classification based on cerebral vascular accident images, melanocytic lesion classification based on skin images [40,41], but also machine condition monitoring [42].

When patients exhibit COVID-19-related symptoms, smartphones, computers, and other electronic devices are permitted to transmit information and X-ray images to LINDA. A user can perform a variety of operations in LINDA, including defining the number of classes, configuring the extraction and classifier characteristics, and changing the extractors and classifiers used. LINDA also includes a graphical dashboard with metrics for evaluating the performance of the extractor and classifier. The IoT system supports the Python programming language, the PostgreSQL database, the TensorFlow, and Keras frameworks. The flow of the developed LINDA-based IoT system is depicted in Fig. 2: the first phase entails integrating the five X-ray image databases, as well as the feature extraction and classification procedures.

The data flow of the LINDA system is depicted in detail in Fig. 1. The information flow of the LINDA system will begin by sending an image from a device, as shown in Fig. 1 by the number 1. A security hash code will also be sent to the system. The system will then call the prediction API, which will select the algorithms to use based on the secure hash. The required models will be loaded into memory. If the system settings have not been completed and some changes are required, the web application (number 2) will be used to upload and categorize images. The

proposed DNN is deployed on this platform, and the algorithm was trained on five databases. Section 3.2 contains a detailed description of DNN. To learn more about LINDA, interested readers should visit [43]. The proposed method has three advantages:

- There is no need for face-to-face communication between physicians and subjects, which reduces medical staff exposure to infection.
- The proposed application diagnoses the X-ray image in less than a second, allowing faster response in case of positive cases.
- By virtue of enjoying a short development cycle, the proposed IoT-based service can be easily upgraded at a minimal cost without disrupting the service.

3.2. The proposed DNN architecture

In this study, we developed a multi-scale DNN system for extracting and recognizing COVID-19 features from CXR images. DNN automatically learns the various features from X-ray images using different scales, and these facets are learned by training the network over several iterations. In previous research, researchers discovered that convolutional sampling on fixed scales frequently limits a DNN's ability to find local invariant patterns, whereas multi-scale sampling allows a DNN to find more reliable and noise-invariant features at different image patches. To address this scope in the context of the current study, a variable filter size ($7 \times 7, 5 \times 5, 3 \times 3$) is used at various convolutional layers with strides of 1×1. Before training the network, pre-processing is performed on X-ray images as shown in Fig. 3. Fig. 4 displays the architecture of the proposed DNN. The DNN extracts robust and geometrically invariant patterns from different patches of X-ray. The input to the DNN is gray-scale images. The proposed DNN consists of 5 blocks. The first block contains 3 convolutional layers with different filter banks and $Block_2$ consisting of pooling layers, which are stacked with $Block_1$ convolutional layers as shown in Fig. 4. The features obtained from the $Block_2$ are concatenated into a single feature vector. Later, three convolution operations are applied with various filter banks (i.e. $Block_3$) on concatenated vector and combined all the features obtained from $Block_3$, then, they are stacked with a single convolutional layer consisting of 1×1 filter bank followed by max-pooling operation with filter size 4×4. Finally, we find $Block_5$ which is composed of two fully connected (FC) layers and a softmax layer of sizes 256, 512, and
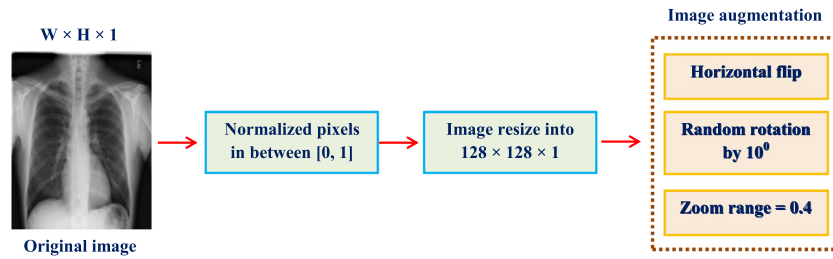
Fig. 3. Preprocessing stages for X-ray images. The preprocessing step includes the normalization of X-ray pixels between 0 and 1, image resize, and image augmentation.
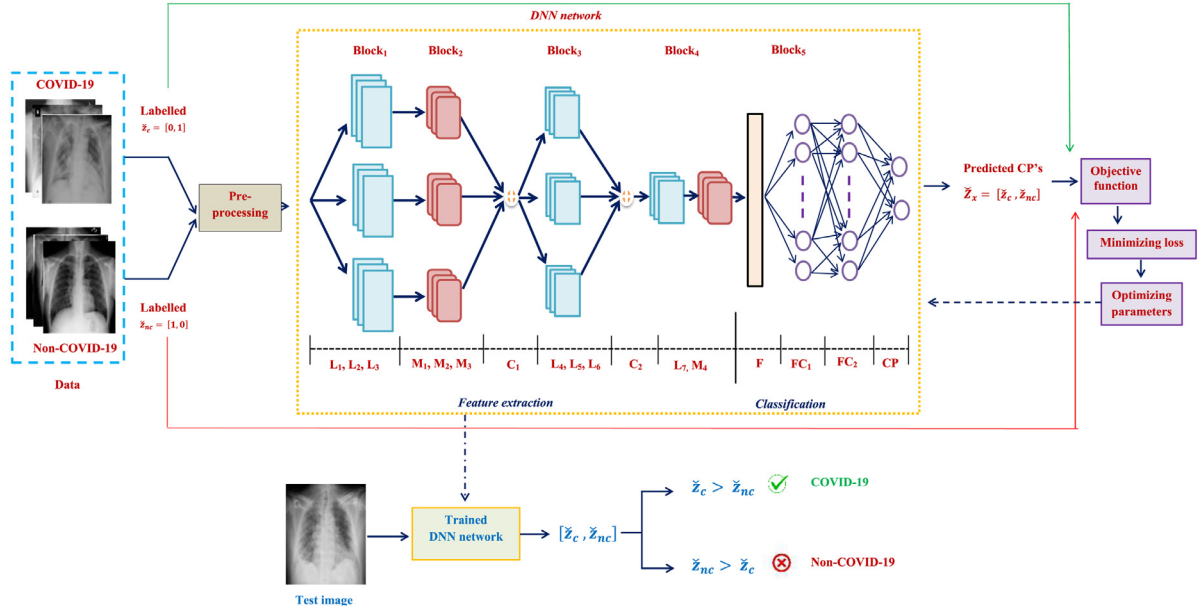

Fig. 4. Illustration of the phases of the proposed DNN framework for COVID-19 detection. First, the network is trained with labeled data. We input the images to be tested into the network in the testing phase, compute the class probability (CP) of COVID-19 and non-COVID-19, and provide the diagnostic results.

2. Technically, the first four blocks ($Block_1$, $Block_2$, $Block_3$, $Block_4$) are considered for feature extraction and the last block ($Block_5$) is employed for classifying COVID-19 using X-rays i.e. final mapping to the output. Table 2 reports the comprehensive description of each layer and its parameters.

Furthermore, filters 7×7, 5×5, and 3×3 are utilized to capture the enriched contextual information. Moreover, the 1×1 filter is used as an identity function.

$$L_1 = f_{1,p}^1 (size:7 \times 7), b_p^1 (size:1 \times 1), p=1,2,\dots,32$$
$$L_2 = f_{1,q}^1 (size:5 \times 5), b_q^1 (size:1 \times 1), q=1,2,\dots,32$$
$$L_3 = f_{1,r}^1 (size:3 \times 3), b_r^1 (size:1 \times 1), r=1,2,\dots,32$$
$$L_4 = f_{p,x}^2 (size:7 \times 7), b_x^2 (size:1 \times 1), s=1,2,\dots,64$$
$$L_5 = f_{q,y}^2 (size:5 \times 5), b_y^2 (size:1 \times 1), u=1,2,\dots,64$$
$$L_6 = f_{r,z}^2 (size:3 \times 3), b_z^2 (size:1 \times 1), v=1,2,\dots,64$$
$$L_7 = f_{c,w}^3 (size:1 \times 1), b_w^3 (size:1 \times 1), w=1,2,.128$$

(1)

Smaller window sizes (i.e., 2×2) are used for the pooling layer in the proposed method as the highest information loss occurs due to the pooling layers. A max-pooling scheme is considered in this work. The formal description of the model is defined mathematically by Eq. (1). The filter initialization values $f_{i,k}^i$ is selected at random from the distribution defined by the filter size, input, and output number of the specific layer's feature maps where a uniform distribution with upper and lower bounds of $\pm k$

is defined by $U(\pm k)$. The mathematical formulation for uniform distribution of filter initialization is shown in Eq. (2).

$$f_{1,p}^1 \sim U\left(\pm\sqrt{\frac{32}{(1+32)\times 7 \times 7}}\right), f_{1,q}^1 \sim U\left(\pm\sqrt{\frac{32}{(1+32)\times 5 \times 5}}\right)$$
$$f_{1,r}^1 \sim U\left(\pm\sqrt{\frac{32}{(1+32)\times 3 \times 3}}\right), f_{p,x}^2 \sim U\left(\pm\sqrt{\frac{64}{(1+64)\times 7 \times 7}}\right)$$
$$f_{q,y}^2 \sim U\left(\pm\sqrt{\frac{64}{(1+64)\times 5 \times 5}}\right), f_{r,z}^2 \sim U\left(\pm\sqrt{\frac{64}{(1+64)\times 3 \times 3}}\right)$$
$$f_{c,w}^2 \sim U\left(\pm\sqrt{\frac{128}{(1+128)\times 1 \times 1}}\right)$$

(2)

The total number parameters are the sum of the parameters of each layer, where the number of parameters of each convolution layer is $(f \times f \times (w_p + b)) \times f_o$, where f is a filter bank size, p is the input number of feature maps, and b is a bias i.e., 1 and f_o is the output number of feature maps. The representation of the ith convolutional layer, L_i is shown in Eq. (3). Where $\Delta(.)$ denotes an activation function of a layer. Rectified linear unit (ReLU) is an activation function adopted in this work, where $f_m \in (1, f_o)$, f_o defines the number of facets maps exist in the layer L_i, e is the number of input facets maps in previous layer i.e., L_e^{i-1}. a, b denotes the coordinates of the feature maps, and $\star$ indicates convolution operation. Eq. (3) is further elaborated as shown in

Table 2
Demonstration of the network parameters employed in the proposed DNN architecture. Here, BN → batch normalization.

R. NO	Layer	Type	Filter size	Stride	Padding	Activation	Output
1.	I	Input	–	–	–	–	$128 \times 128 \times 1$
2.	L_1	Convolution	7×7	1	Same	ReLU	$128 \times 128 \times 32$
3.	L_2	Convolution	5×5	1	Same	ReLU	$128 \times 128 \times 32$
4.	L_3	Convolution	3×3	1	Same	ReLU	$128 \times 128 \times 32$
5.	M_1	Max-pooling + BN	2×2	2	Valid	–	$64 \times 64 \times 32$
6.	M_2	Max-pooling + BN	2×2	2	Valid	–	$64 \times 64 \times 32$
7.	M_3	Max-pooling + BN	2×2	2	Valid	–	$64 \times 64 \times 32$
8.	C_1	Concatenation	–	–	–	–	$64 \times 64 \times 96$
9.	L_4	Convolution	7×7	1	Same	ReLU	$64 \times 64 \times 64$
10.	L_5	Convolution	5×5	1	Same	ReLU	$64 \times 64 \times 64$
11.	L_6	Convolution	3×3	1	Same	ReLU	$64 \times 64 \times 64$
12.	C_2	Concatenation	–	–	–	–	$64 \times 64 \times 192$
13.	L_7	Convolution	1×1	1	Valid	ReLU	$64 \times 64 \times 64$
14.	M_4	Max-pooling + BN	4×4	4	Valid	–	$16 \times 16 \times 64$
15.	F	Flatten	–	–	–	–	16384×1
16.	FC_1	Full-connection	–	–	–	ReLU	512×1
17.	FC_2	Full-connection	–	–	–	ReLU	256×1
18.	CP	Class probabilities	–	–	–	Softmax	2×1

Eq. (4).

$$L_{fm}^i = \Delta\left(L_e^{i-1} \star f_{e,fm}^i + b_{fm}^i \right), \tag{3}$$

$$L_{fm}^i(a,b) = \\ \Delta\left(\sum_{u=-m}^{m} \sum_{v=-m}^{m} L_e^{i-1}(a-u, b-v) \cdot f_{e,fm}^i(u,v) + b_{fm}^i \right) \tag{4}$$

The filter bank size of the layer L^i is $(2 \times m + 1) \times (2 \times m + 1)$. In the proposed model, the "same" padding is applied to keep the size of the feature map constant. Max-pooling operation is performed on the output of convolutional layer L_{fm}^i with filter bank of size 2×2. From each feature map, max-pooling measures the utmost importance of each patch to highlight the primary feature represented within the patch. Max-pooling also reduces the number of parameters to make the model simple, In addition, it provides feature maps that are invariant to translation, rotation, and scale. As shown in Fig. 4, the input images of 128×128 pixels are down-sampled by the max-pooling layer, resulting in filter maps of various sizes after each layer of convolution in the $Block_2$, later concatenated, with C_1 being the output obtained by the max-pooling layer. Further, $Block_3$ is stacked with C_1, containing variable sizes of filter banks for convolution operation. Followed by concatenation operation, C_2 is applied on outputs of $Block_3$, which is called $Block_4$. The convolution operation with filter bank of 1×1 is applied on output feature maps of $Block_4$ followed by a max-pooling operation, then all the feature maps are flattened to a single vector of size 16384×1, where W_i and b^i are the weight and bias of the ith FC layer. The output of the second FC layer, FC_2, is further fed into the softmax layer. The softmax layer consists of two neurons and produces a probability vector, $\hat{Z} = [\hat{z}_c, \hat{z}_{nc}]$, where, $\hat{z}_c$ is the prediction score of COVID-19 class and $\hat{z}_{nc}$ of non-COVID-19 class. The jth probability value is obtained by Eq. (5).

$$\hat{Z}_j = \frac{e^{FC_2^j}}{\sum_{j=1}^{2} e^{FC_2^j}}, j = 1, 2. \tag{5}$$

3.2.1. Network training

The proposed network trains on X-ray images and computes the probability of each class, CP. The weights of the proposed network are initialized randomly with the help of uniform distribution as shown in Eq. (2) and the adaptive momentum (Adam) optimization technique employed to tune the hyperparameters to minimize the loss between predicted class probabilities, $\hat{Z} =$

$[\hat{z}_c, \hat{z}_{nc}]$ and actual class probabilities, $Z = [z_c, z_{nc}]$ of COVID-19. The initial learning rate and weight decay are fixed as 0.00001 in Adam. As a loss function for the classifier, we use cross-entropy. Eq. (6) is employed to compute the cross-entropy.

$$\psi(Z, \hat{Z}) = -Z \log \hat{Z} \\ = -[z_c, z_{nc}] \log[\hat{z}_c, \hat{z}_{nc}] \\ = -z_c \cdot \log(\hat{z}_c) - z_{nc} \cdot \log(\hat{z}_{nc}) \tag{6}$$

with the batch size S, the loss function is given in Eq. (7).

$$\psi(Z, \hat{Z}) = -\left(\frac{1}{S} \sum_{i=1}^{S} z_c \cdot \log \hat{z}_c + (z_{nc}) \cdot \log(\hat{z}_{nc}) \right), \tag{7}$$

where $Z (= [z_1, z_2])$ is one-hot encoding vector of the actual labels. Batch size 16 is considered while training the proposed DNN since the network can occupy less memory in the proposed system.

4. Empirical evidence

This section delves into the specifics of the proposed DNN's implementation, as well as the database's details, before concluding with the empirical findings.

4.1. Experimental setting

This sub-section describes the resources used for experiments. For the training and testing of the proposed and existing models, the Keras framework and Anaconda Python 3.6 package are considered in this study. The specifications of the working system are NVIDIA Quadro P5000 graphics processor, 256-bit memory interface, 16 GB GPU RAM, Cuda core-2560, GDDR5X memory, and 288.5 GB/s bandwidth.

4.2. Experimental data

In this sub-section, we discuss about the databases and evaluation procedure of the proposed DNN model for diagnosing the COVID-19 is described. All the experiments are performed on the five publicly available databases, namely D1 [19], D2 [20], D3 [21], D4 [22], and D5 [23]. The statistical information of these databases is reported in Table 3. When dealing with a database containing a small number of images, overfitting or excessive variance of ML algorithms is common. The overfitting problem is addressed in this study by considering horizontal flip, random

Table 3
Statistical information according to the database and class.

Database	Before augmentation					
	Train		Validation		Test	
	COVID-19	non-COVID-19	COVID-19	non-COVID-19	COVID-19	non-COVID-19
D1	1299	3027	163	378	163	378
D2	960	1073	120	134	120	134
D3	148	2960	18	370	18	370
D4	256	355	32	45	32	45
D5	99	400	13	50	13	50
After augmentation						
D1	5196	12108	163	378	163	378
D2	3840	4292	120	134	120	134
D3	592	11840	18	370	18	370
D4	1024	1420	32	45	32	45
D5	396	1600	13	50	13	50

rotation by 10 degrees, and zoom range 0.4 as image augmentation strategies. Moreover, to maintain the consistency of the proposed model, all the X-ray images of five databases are resized into 128×128 and each database is divided into three groups: train, validation, and test sets. For all of the investigations, a k(=10)-fold cross-validation methodology is adopted to assess the performance of the proposed method. In other words, out of ten subsets, eight are employed for training, one is used for validation, and the remaining one is utilized for testing. The underfitting and overfitting problems may solve by considering the 10-fold cross-validation technique. Table 3 reports the number of X-ray images employed in the train, validation, and test in the ratio of 0.8, 0.1, and 0.1, respectively. The upper and lower part of Table 3 denotes the number of samples before and after image augmentation. Moreover, four well-known evaluation metrics, namely accuracy, precision, recall, and F1-score are employed for evaluation of the proposed DNN and comparative methods. The detailed description of evaluation metrics is beyond the scope of this study.

4.3. Results

In this sub-section, the empirical results of five databases are discussed. The proposed DNN is evaluated on five databases, namely D1, D2, D3, D4, and D5. Fig. 5 depicts the training procedure for five datasets. As seen in Fig. 5, accuracy improved rapidly during training until it reached an average of 10 to 20 iterations, and then progressively increased. After multiple iterations, the performance of the training and validation sets appeared to be smooth and did not grow any further. Similarly, training and validation losses decreased until it reached 10 to 20 epochs. The training loss assesses how well the model fits the training data, whereas the validation loss assesses how well the model fits new data. We discovered that the proposed DNN can achieve nearly 100% accuracy in training and the best results in validation. The training and validation losses of the proposed DNN are 0.1003, 0.1291 for D1, 0.0019, 0.0167 for D2, 0.0186, 0.01426 for D3, 0.0096, 0.0201 for D4, 0.0099, and 0.0213 for D5. As a result, it is clear that the proposed DNN structure offers considerable benefits in terms of COVID-19 identification.

To illustrate the robustness of the proposed DNN structure on five databases, metrics such as precision, recall, F1-score are measured, which is noted in Table 4. Furthermore, the accuracy of the proposed method is compared with the fourteen existing works on the five databases, and the results are noted in Table 4.

We can summarize the following:

- Table 4 clearly indicates that the proposed DNN framework obtains an average detection F1-score, recall, and precision of 96% on D1, 100% on the D2, 99% on D3, 99% on D4,

and 100% on D5 databases respectively. This indicates that the proposed model learns well on X-ray images and it is able to distinguish the features belonging to COVID-19 and non-COVID-19.

- It is observed from Table 4 that the detection accuracy of the proposed method is 96.01% on the D1, 99.61% on D2, 99.22% on D3, 98.83% on D4, and 100% on D5 databases respectively, which is far better than the accuracies obtained by the existing methods. Besides, the error rate incurred by the proposed method on the testing set is 0.1391, 0.0057, 0.0996, 0.0178, and 0.0124 on D1, D2, D3, D4, and D5 databases, which is impressive enough in comparison with the existing methods.

4.3.1. Comparative results

In this sub-section, our aim is to compare the performance of fourteen baseline approaches such as XCeption net [25], Inception_V2 [44], SVM [38], Coronet [22], COVID-SDNet [28], AD-COVID19 [29], transfer learning approach [27], Prior attention network [30], VGG-19 [31], DarkCovidNet [32], ResNet50 [33], RLDD [39] with the proposed method in the last experiment. The short description of these approaches are discussed in Section 2 however, its detailed explanation is beyond the scope of this work. In this study, all the models adopted for comparison are implemented based on specifications as stated in the original papers. Table 4 reports the average classification accuracies achieved by these methods on five publicly available databases. Also notes the values of precision, recall, and F1-score of these approaches.

Table 4 demonstrates that the proposed method is the best and it happens due to the use of the proposed DNN to fetch more reliable and noise-invariant facets at different image patches. Using this approach, we design an end-to-end IoT-enabled DL framework for fast and remote diagnosis which is our main objective. The achieved accuracy of the proposed DNN on five databases, namely D1, D2, D3, D4, and D5 are 96.01%, 99.61%, 99.22%, 98.83%, and 100%, respectively. Moreover, the measurement of the running time is a significant aspect of analyzing the proposed DNN. The training and testing time for implementing the proposed DNN, as well as comparative approaches, are detailed in Table 5. Normally, the training time of a DL model relies on the size of the input, the size of the network, number of folds, number of epochs, and other parameters. Moreover, all the experiments are conducted in the same environment to measure time efficiency. Table 5 clearly states that the proposed DNN requires an average training and testing time across all the databases. Moreover, a comparative analysis is employed to examine the false negatives predicted by the existing methods and the proposed DNN. False negatives refer to cases when a person, who has the COVID-19 disease screens negative rather
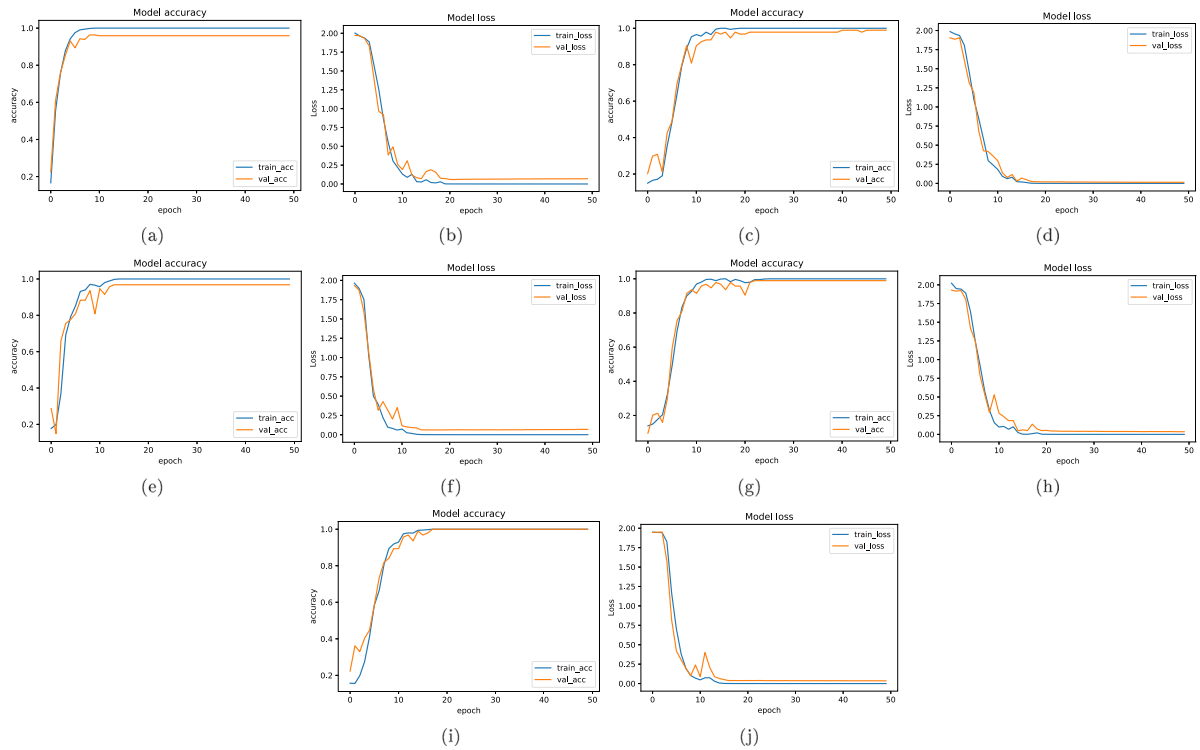
Fig. 5. Training and loss curves on five database. (a), (c), (e), (g), and (i) represents accuracy curve on D1, D2, D3, D4, and D5, (b), (d), (f), and (h) indicates loss curve on D1, D2, D3, D4, and D5.

Table 4

Analysis of state-of-the-art methods on test set of five databases using other metrics, namely accuracy, precision, recall, and F1-score in terms of%.

Ref.	Method	Year	D1				D2				D3				D4				D5			
			Accuracy	Precision	Recall	F1-score	Accuracy	Precision	Recall	F1-score	Accuracy	Precision	Recall	F1-score	Accuracy	Precision	Recall	F1-score	Accuracy	Precision	Recall	F1-score
[25]	XCeption net	2020	90.19	90	91	91	99.22	99	99	99	96.83	97	97	97	94.85	96	95	95	90.19	89	91	90
[19]	COVID-Net	2020	94.90	96	94	95	93.05	95	93	94	91.12	91	91	91	79.89	81	80	79	90.33	91	89	90
[44]	Inception_V2	2020	94.15	95	94	94	99.01	97	99	98	97.05	97	97	97	82.49	82	82	82	95.97	96	98	97
[38]	SVM	2020	91.36	90	90	90	98.03	98	98	98	98.88	99	90	94	73.19	85	73	76	91.36	91	91	91
[22]	Coronet	2020	92.03	92	92	92	99.50	99	99	99	98.01	98	99	98	85.00	85	85	85	92.03	92	92	92
[28]	COVID-SDNet	2020	94.91	96	95	95	96.01	96	96	96	95.72	97	95	96	91.33	90	92	91	94.00	94	94	94
[29]	AD-COVID19	2020	90.34	90	93	91	89.09	89	89	89	91.86	93	91	92	83.67	86	83	84	95.06	96	94	95
[27]	Transfer learning	2020	88.86	89	89	89	98.02	98	98	98	93.75	96	92	94	89.33	90	89	89	97.73	97	97	97
[30]	Prior attention	2020	90.06	92	90	91	91.99	92	92	92	96.58	98	96	97	87.63	88	86	87	93.00	93	93	93
[31]	VGG19	2020	89.96	91	90	90	96.68	97	97	97	97.61	98	98	98	88.39	88	86	87	98.00	98	98	98
[32]	DarkCovidNet	2020	88.31	88	88	88	95.37	95	95	95	98.08	98	98	98	93.18	92	94	93	95.01	95	95	95
[33]	ResNet50	2021	92.05	92	92	92	97.34	97	97	97	98.32	98	98	98	94.11	96	95	94	98.86	99	99	99
[39]	RLDD	2021	93.33	93	93	93	96.02	97	95	96	97.07	98	97	96	90.91	91	91	91	95.66	97	95	96
–	**Proposed method**	–	**96.01**	96	96	96	**99.61**	100	99	100	**99.22**	99	99	99	**98.83**	98	99	99	**100**	100	100	100

Table 5

Execution time in seconds on five databases viz. D1, D2, D3, D4, and D5 by various methods.

Ref.	Method	Training time					Testing time for all the Images				
		D1	D2	D3	D4	D5	D1	D2	D3	D4	D5
[25]	XCeption net	495.0	480.0	270.0	330.0	316.13	5.0	16.0	10.0	3.01	2.0
[19]	COVID-Net	939.0	987.0	438.33	402.01	386.99	25.0	27.05	11.76	8.33	4.09
[44]	Inception_V2	4700.0	850.0	925.0	625.0	600.01	12.0	19.0	18.02	11.23	10.0
[38]	SVM	16.75	13.2675	19.62	17.34	15.091	2.3145	1.3167	2.891	1.8472	1.2710
[22]	Coronet	2488.0	2676.0	1336.0	1720.0	1201.09	86.0	42.0	51.0	36.701	24.42
[28]	COVID-SDNet	451.00	526.66	357.92	380.06	365.66	86.0	35.0	38.33	23.64	14.96
[29]	AD-COVID19	98.01	101.03	78.72	67.00	43.94	9.04	11.35	3.78	2.00	1.33
[27]	Transfer learning	92.98	17.63	19.09	5.092	4.66	3.014	3.63	1.302	1.001	0.78
[30]	Prior attention	245.33	260.04	201.33	166.66	102.11	67.27	30.98	23.02	19.00	13.109
[31]	VGG19	2205.0	2782.06	1618.0	1920.0	1386.0	6.0	8.2	7.01	5.11	3.03
[32]	DarkCovidNet	200.0	250.0	180.0	185.0	150.0	1.03	1.769	1.25	0.996	0.841
[33]	ResNet50	934.0	960.0	700.08	658.00	500.0	17.0	19.0	15.0	14.07	10.2
[39]	RLDD	104.65	99.09	84.68	58.88	38.01	11.09	4.330	3.810	3.002	1.109
–	**Proposed method**	95.4	96	66	20.9	12.3	0.541	0.692	1.28	0.461	0.202

than positive which is included in Fig. 6. In this regard, we wish to mention that all the five databases are considered together while computing false-negative cases. It is clear from Fig. 6 that the proposed DNN has very few false negatives to existing state-of-the-art methods. In addition to this, in this study, the number

of parameters involved and memory size (in MB) required for an image are used to determine the complexity of the proposed DNN. The proposed DNN's complexity is compared to 12 current state-of-the-art DL approaches and results are reported in Table 6. Table 6 demonstrates that the number of parameters in the
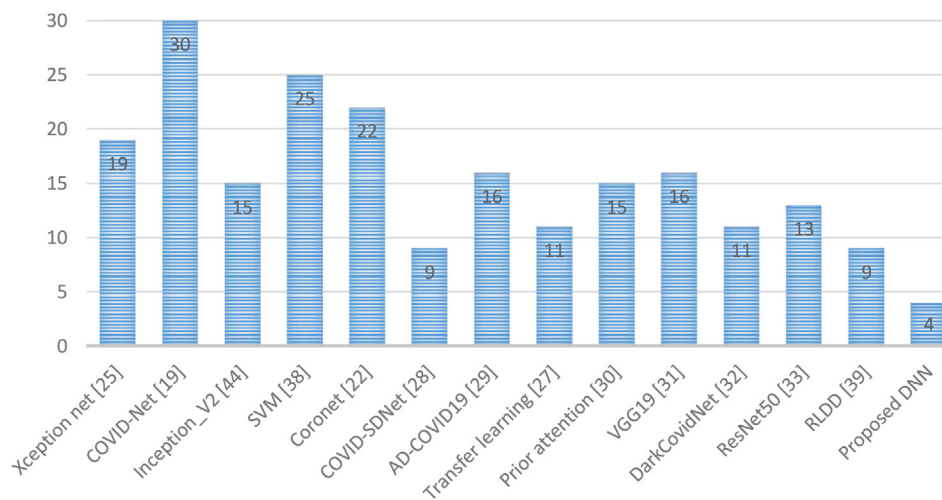
Fig. 6. A comparative analysis on number of false negatives of existing approaches as well as the proposed DNN.

Table 6
Computational efficiency analysis of the proposed DNN with other DL approaches.

Ref.	Method	No. parameters (in millions (M))	Memory size (in MB)
[25]	XCeption net	29.2	38
[19]	COVID-Net	11.75	106
[44]	Inception_V2	55.9	296
[22]	Coronet	33.9	240
[28]	COVID-SDNet	40.3	386
[29]	AD-COVID19	89.1	820
[27]	Transfer learning	11.6	120
[30]	Prior attention	15.9	110
[31]	VGG19	20.5	385
[32]	DarkCovidNet	1.12	8
[33]	ResNet50	25.6	420
[39]	RLDD	26	160
–	**Proposed method**	9.04	94

proposed DNN is smaller than the number of parameters in nineteen state-of-the-art DL techniques, demonstrating the proposed DNN's simplicity. In comparison to existing DL techniques, the memory size (in MB) required for an image is also smaller. It could satisfy the needs of many real-time COVID-19 diagnosis applications. After optimization in both time and space, the model could be equipped with real-time edge devices, such as NVIDIA TX2.

4.4. Robustness of the proposed DNN

In this sub-section, we conduct experiments to find the robustness of the proposed DNN. In basic terms, feature engineering is the process of converting chest x-ray images into desirable features using the proposed DNN in order to improve model accuracy. To compare the performance of the proposed DNN with some state-of-the-art approaches, accuracy, precision, recall, and F1-score are used, and the results are presented in Table 4. The proposed DNN outperforms all state-of-the-art techniques, as reported in Table 4. It means the proposed DNN does the feature engineering task well in comparison to other state-of-the-art methods. Nowadays, the values of these evaluation measures are no longer sufficient to demonstrate how good a DL model is. The t-SNE plots of the feature vectors obtained by the proposed DNN along with state-of-the-art approaches on five different databases are shown in Figs. 7, 8, 9, 10, and 11 in order to illustrate the effectiveness of the proposed DNN model over some of the existing methods. The features in a 2-D plot are depicted using t-distributed stochastic neighbor embedding, a dimensional reduction technique that allows us to perceive a high-dimensional database in a low-dimensional environment. Because such embedding incorporates categorization information, it visualizes the learned proposed network's most recent embeddings. It is clear from Figs. 7, 8, 9, 10, and 11 that the proposed DNN is the only method, which can extract distinguishable features for five databases separately and forms well-separated clusters when we are mapping them from higher-dimensional to a two-dimensional plane. The t-SNE plots of a few of the existing approaches are also well separated with a small margin on some of the databases. Thus, the performances of these methods are not always consistent. This experiment shows how efficient the proposed method is. Furthermore, to show the proposed method's efficient feature learning capacity, the probability vector created by the softmax layer is compared to a few existing DL techniques. Initially, the five best state-of-the-art methods are selected based on their accuracies [19,22,25,28,39]. We pick four CXR images (two are from COVID-19 patients and two are from a healthy person) randomly from the test sets. Then, the probability score of each method along with the proposed approach to these four images is estimated and displayed in Fig. 12. The label '0' and '1' in the graphs of Fig. 12 denote the probability of covid-19 infected patients and healthy persons, respectively. The probability score ranges between 0 to 1 and the ideal probability score of label '0' and label '1' for covid-19 infected patients is 1 and 0, respectively. Similarly, it is vice versa for healthy people. It can be observed from Fig. 12 that the proposed method predicts scores nearer to ideal values for both covid-19 infected patients and healthy persons. On the other hand, probability scores obtained by the existing DNN based approaches for healthy persons and infected patients are relatively far from the ideal values.
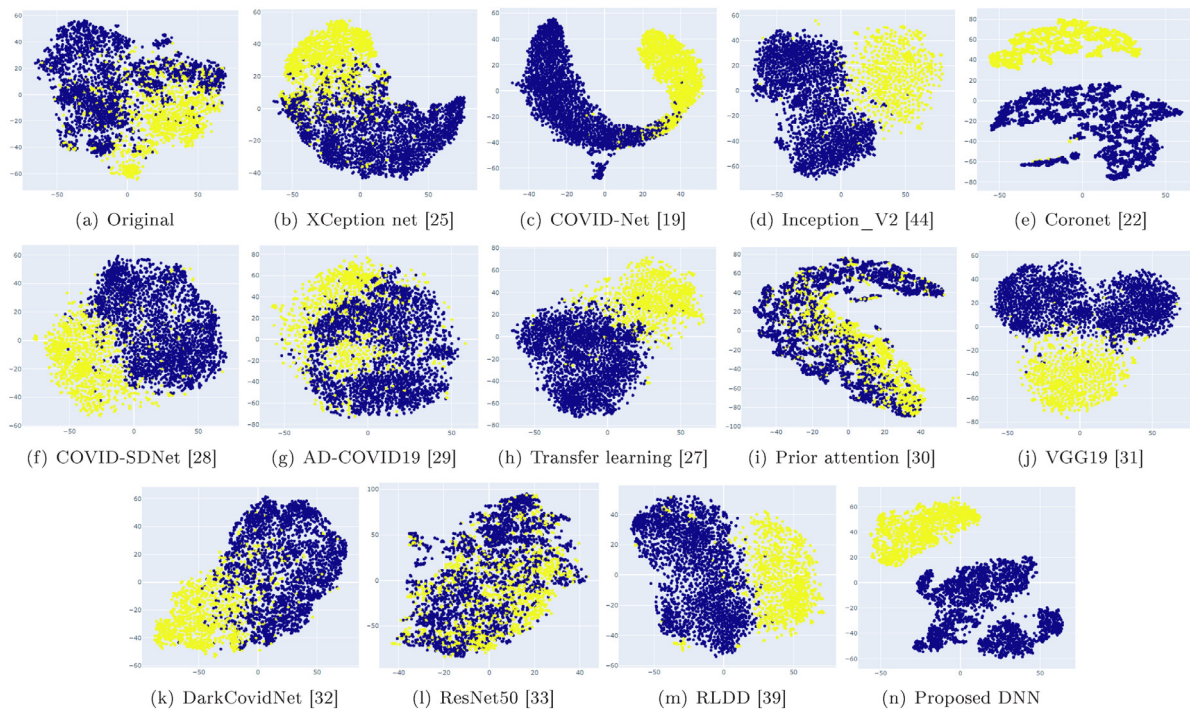
Fig. 7. Feature visualization on D1 database. Here, (a) shows the input feature map, (b)–(m) depicts the feature maps obtained from existing methods, and (n) shows the feature map obtained from the proposed DNN.
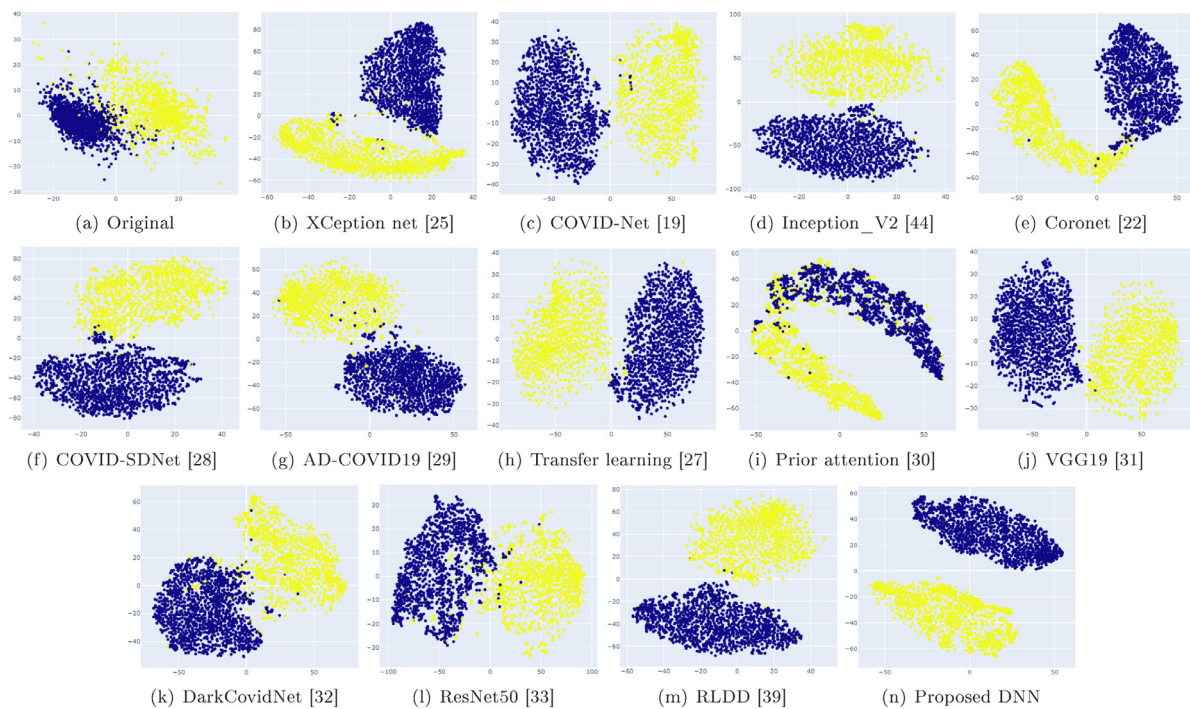

Fig. 8. Feature visualization on D2 database. Here, (a) shows the input feature map, (b)–(m) depicts the feature maps obtained from existing methods, and (n) shows the feature map obtained from the proposed DNN.

In addition to this, we observe that the proposed DNN is evaluated using a 10-fold cross-validation procedure, in which the original dataset is randomly divided into ten equal-size subsamples. Only one subsample of the ten is preserved as test data for the algorithm, while the other nine are used for training and validation. The folds are then used to repeat the cross-validation process ten times, with each of the ten subsamples serving as validation data exactly once. The performance of the proposed DNN can then be estimated by averaging (or otherwise combining) the evaluation metrics acquired from the 10 folds. However, the proposed DNN's overall accuracy differs from the individual accuracy of each fold. To identify the variance in the obtained accuracies on each database, the standard deviation of accuracies acquired in ten different folds is assessed, as illustrated in Fig. 13.
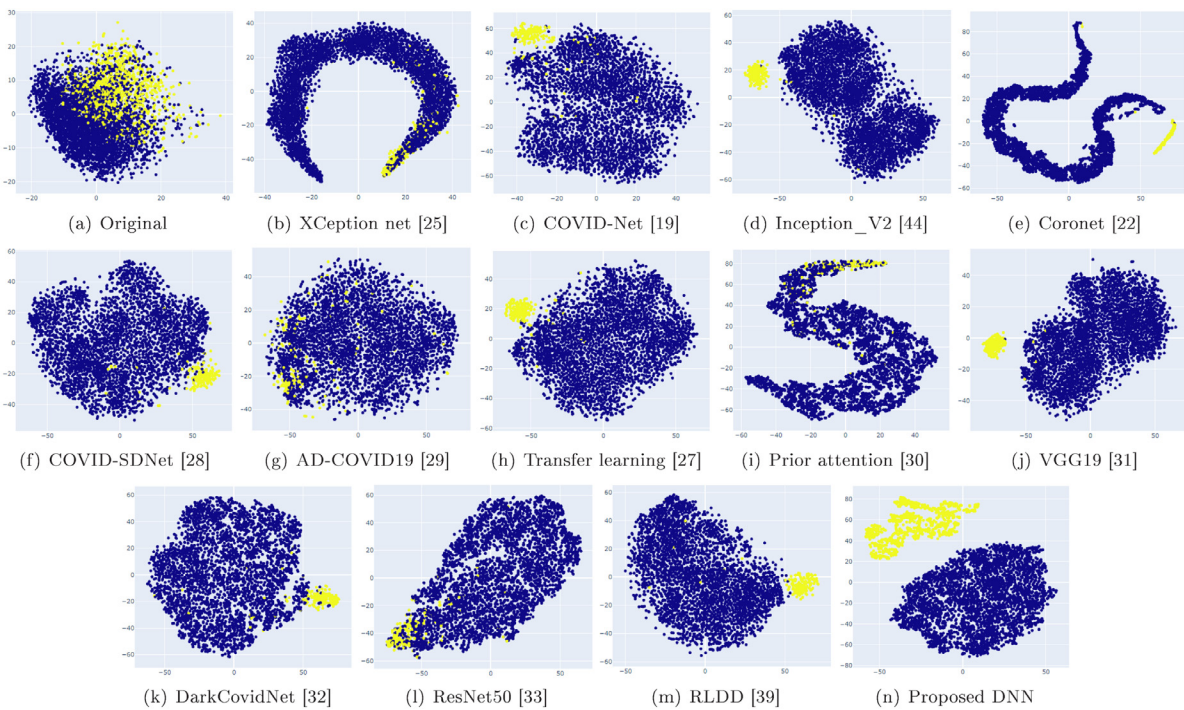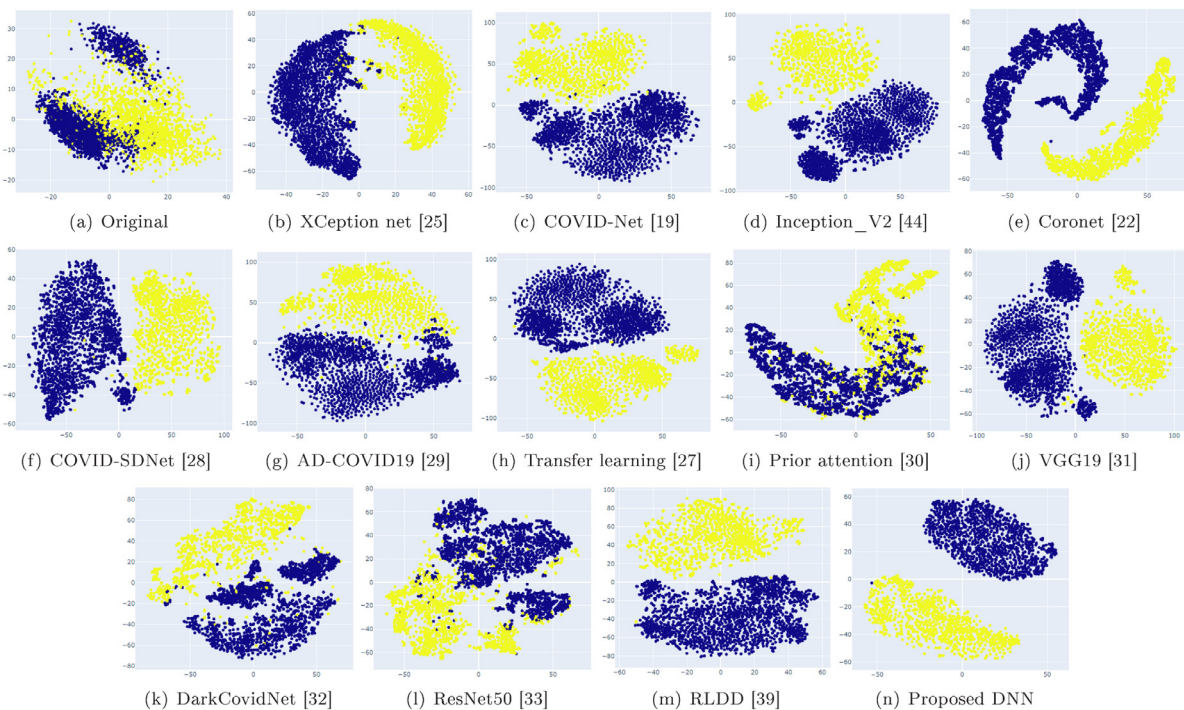
Fig. 9. Feature visualization on D3 database. Here, (a) shows the input feature map, (b)–(m) depicts the feature maps obtained from existing methods, and (n) shows the feature map obtained from the proposed DNN.

Fig. 10. Feature visualization on D4 database. Here, (a) shows the input feature map, (b)–(m) depicts the feature maps obtained from existing methods, and (n) shows the feature map obtained from the proposed DNN.

It is clear from Fig. 13 that the obtained low standard deviation on a database implies that 10-fold accuracies tend to be extremely close to the averaged accuracy of the proposed DNN.

4.5. Ablation study

To develop a DNN model from scratch for a particular problem is not an easy task especially when we are facing the data scarcity problem with a database containing few images. An ablation study is required to understand the contribution of each module of the proposed DNN. Thus, an ablation study is conducted to finalize the architecture of the proposed model in such a way that it performs well with a test set. In accordance with our experiments, the proposed DNN gives an accuracy of 96.01%, 99.61%, 99.22%, 98.83%, and 100% for the task of COVID-19 classification of the X-ray images on D1, D2, D3, D4, and
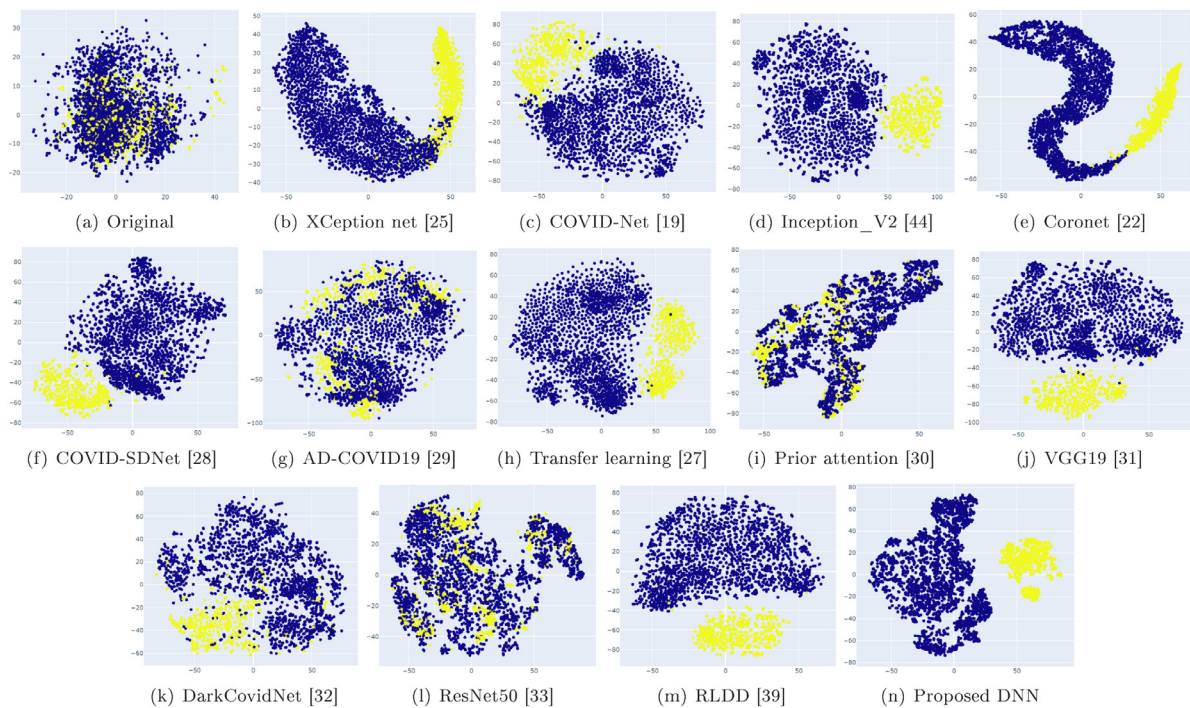
Fig. 11. Feature visualization on D5 database. Here, (a) shows the input feature map, (b)–(m) depicts the feature maps obtained from existing methods, and (n) shows the feature map obtained from the proposed DNN.

Table 7
The performance in terms of accuracy (%) on the test set of five databases with and without batch normalization layers & variation of pooling layers and without pooling layers are noted.

The seventh experiment belongs to batch normalization					
BN	D1	D2	D3	D4	D5
No	92.87	98.82	99.05	93.19	96.82
Yes	**96.01**	**99.61**	**99.22**	**98.83**	**100**
The eighth experiment belongs to pooling					
Pooling	D1	D2	D3	D4	D5
Max	**96.01**	**99.61**	**99.22**	**98.83**	**100**
Average	94.46	99.21	99.16	90.01	**100**
No	88.01	97.64	94.44	89.38	96.82

D5 databases respectively. In this sub-section, the effect of the performance of the proposed DNN is assessed by varying the model parameters. In the first experiment, the behavior of the DNN model for different activation functions is shown in Fig. 14. Fig. 14 indicates that ReLU activation gives a better performance than others. In the second experiment, the effect of the batch normalization layer on the performance of the proposed DNN is evaluated. The results of the second experiment are presented in the upper part of Table 7. In the third experiment, the effect of change in pooling layers on the performance of the proposed DNN is assessed. The results of this study are presented in the lower part of Table 7. The proposed DNN is divided into five blocks, namely $Block_1$, $Block_2$, $Block_3$, $Block_4$, and $Block_5$. We have experimented i.e., the fourth experiment, to quantify the block-wise performance of the proposed DNN. The results of the fourth experiment are shown in Table 8.

From Table 8, we can observe that although increasing layers containing blocks the proposed model gives better performance. Moreover, we can observe from the last row of Table 8 is that overfitting occurred after increasing more than 5 blocks. In addition, as seen in Fig. 15, there is a huge gap between training and validation losses when increasing more than 5 blocks. So, the model is overfitting.

Moreover, optimization plays a crucial role in the DL model for updating the weights to reduce the losses of the DL model, also called hyper-parameter tuning. There are many optimization techniques available for parameter hyper-space search. Some of the widely used optimization techniques for DL approaches are stochastic gradient descent (SGD), adaptive gradient descent (Adagrad), root mean square propagation (RMSprop), SGD with momentum, and adaptive moment estimation (Adam). In the fifth experiment, we have tested the aforementioned optimization techniques. The performance analysis of various optimization techniques is shown in Fig. 16. It can be observed from Fig. 16 that Adam optimization gave a better performance for COVID-19 classification from X-ray images. However, the proposed model obtained a satisfactory performance for other optimization techniques. Variations of different scales are examined by the sixth experiment. Table 9, shows the performance of the proposed DNN using different filter scales. To measure the importance of the multi-scale approach, conducted a seventh experiment with and without multi-scale blocks of the proposed DNN. The measurement of the proposed DNN performance with and without multi-scale is shown in Fig. 17.
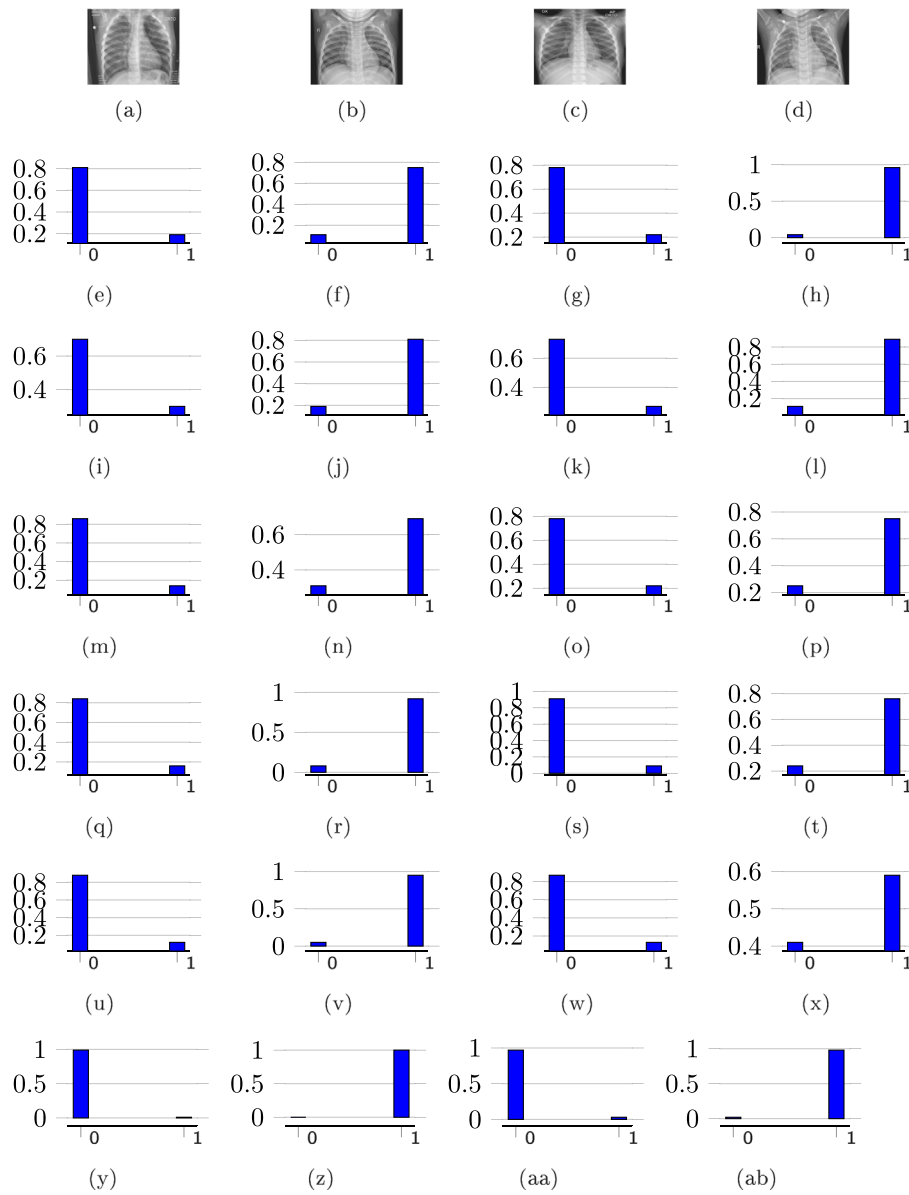
Fig. 12. The comparative probability scores of the proposed DNN along with existing methods. (a), (c) represents the COVID-19 CXR images and (b), (d) indicates the non-COVID-19 CXR images. Each row represents the probability scores of two classes. (e), (f), (g), (h) is of [25], (i), (j), (k), (l) is of [19], (m), (n), (o), (p) is for [22], (q), (r), (s), (t) is for [28], (u), (v), (w), (x) is for [39], (y), (z), (aa), (ab) is for the proposed method.
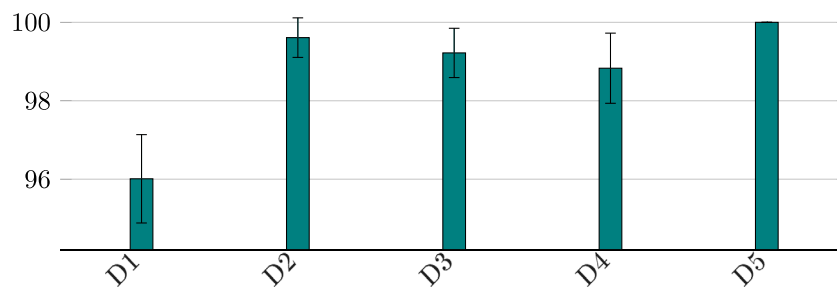

Fig. 13. The performance (averaged accuracy ±standard deviation) of the proposed DNN on five databases.

It is clear from Fig. 17 that the performance of the proposed DNN without multi-scale is not satisfactory. Thus, we can conclude that multi-scale features provide significant features to distinguish COVID-19 from non-COVID-19.

It can be observed from Table 9 that the combination of $3 \times 3, 5 \times 5, 7 \times 7$ scales gave better performance than other scales. In addition, an experiment is conducted on varying learning rates. Fig. 18, shows the performance varying while changing learning
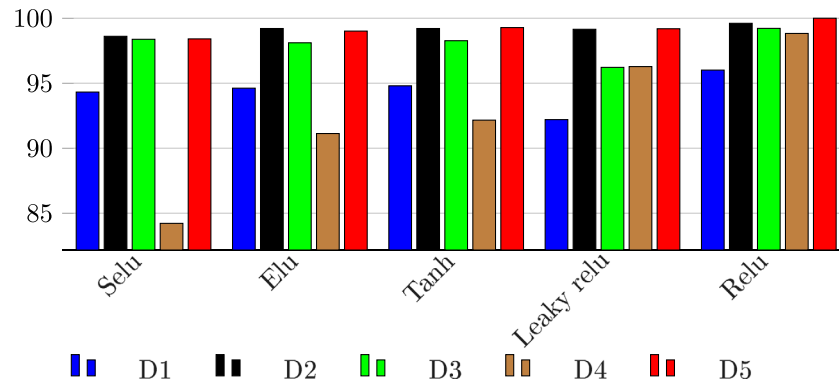
Fig. 14. Variation of COVID-19 classification accuracies (%) obtained using different activation functions, where Selu → scaled exponential linear unit, Elu → exponential linear unit, Tanh → hyperbolic tangent, ReLU → rectified linear unit.
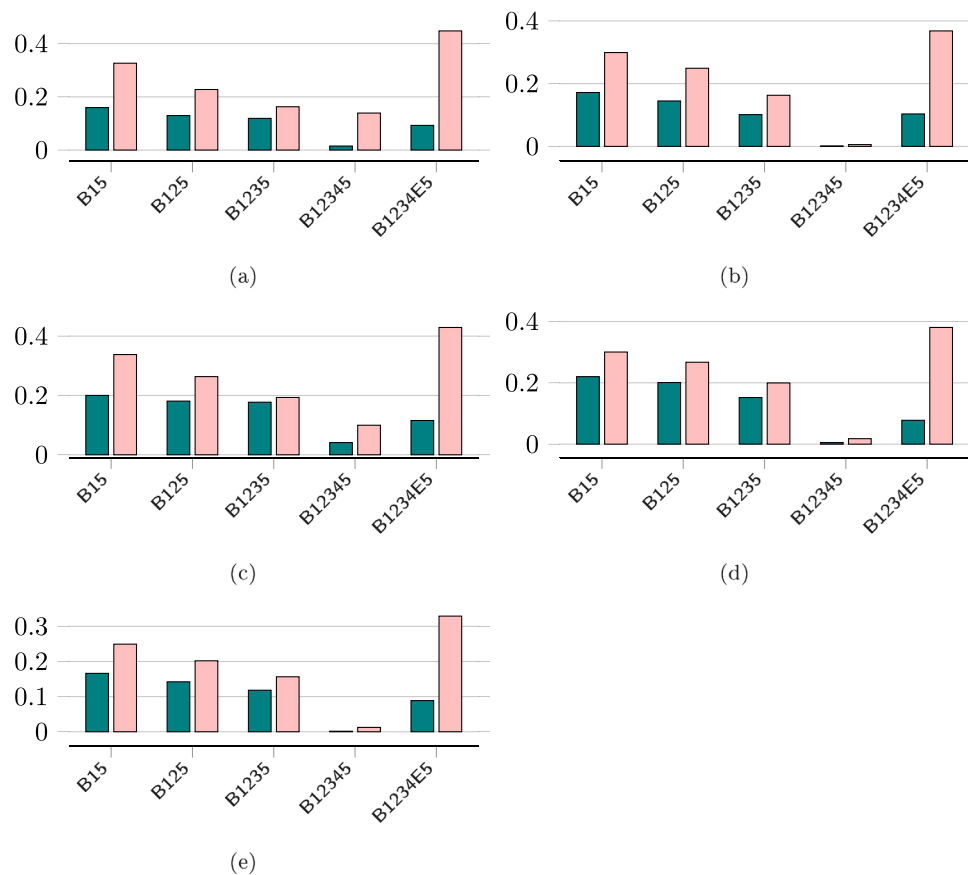

Fig. 15. The training and validation losses incurred by the proposed DNN model after adding more layers to it on five databases. (a), (b), (c), (d), and (e) indicates five databases, namely D1, D2, D3, D4, and D5. Here, B15 → $Block_1 + Block_5$, B125 → $Block_1 + Block_2 + Block_5$, B1235 → $Block_1 + Block_2 + Block_3 + Block_5$, B12345 → $Block_1 + Block_2 + Block_3 + Block_4 + Block_5$, B1234E5 → $Block_1 + Block_2 + Block_3 + Block_4 + Extra_{block} + Block_5$.

Table 8
The performance in terms of accuracy (%) on the test set of five databases with varying blocks of the proposed DNN (for example, $Block_1 + Block_5$, $Block_1$ is for feature extraction and $Block_5$ is for classification).

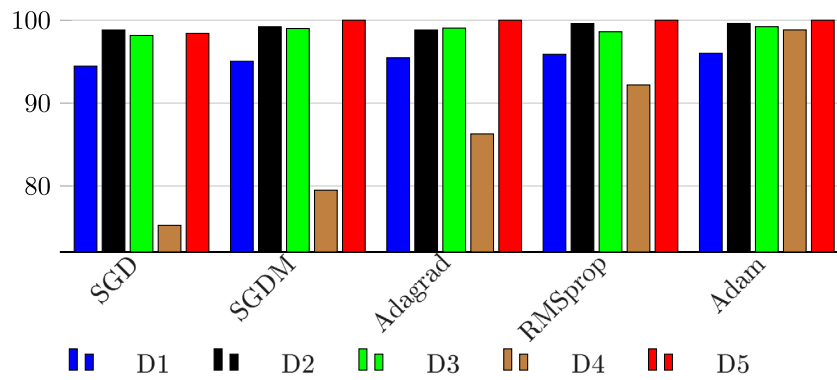| Blocks	D1	D2	D3	D4	D5
$Block_1 + Block_5$	93.88	97.64	94.45	89.38	98.41
$Block_1 + Block_2 + Block_5$	94.63	99.21	99.01	93.19	98.59
$Block_1 + Block_2 + Block_3 + Block_5$	94.91	99.21	**99.22**	94.16	96.82
$Block_1 + Block_2 + Block_3 + Block_4 + Block_5$	**96.01**	**99.61**	**99.22**	**98.83**	**100**
$Block_1 + Block_2 + Block_3 + Block_4 + Extra_{Block} + Block_5$	74.63	73.09	78.43	65.89	84.21

Fig. 16. Variation of COVID-19 classification accuracies (%) obtained on the test set using different optimization techniques.
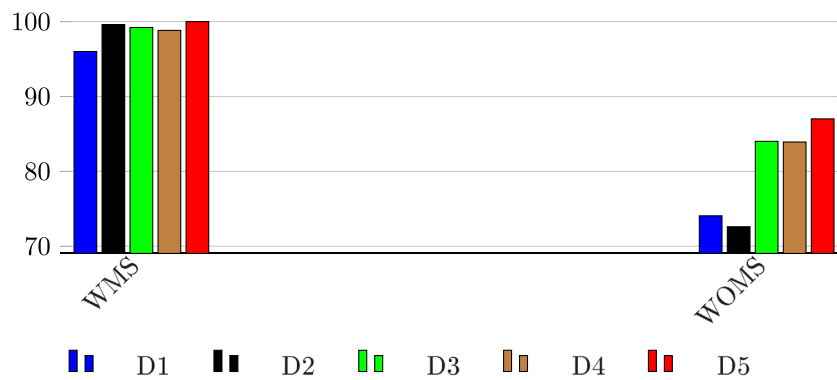

Fig. 17. The effect of the proposed DNN with and without multi-scale approach, where, WMS → With multi-scale and WOMS → Without multi-scale.

Table 9
The performance in terms of accuracy (%) on the test set of five databases with varying filter scales of the proposed DNN.

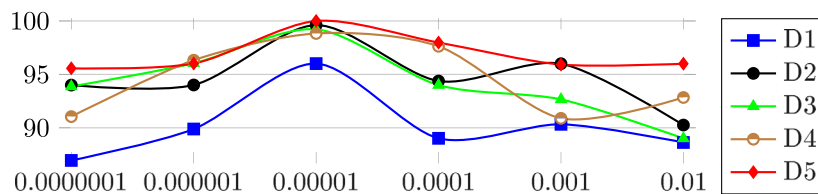| Blocks	D1	D2	D3	D4	D5
$11 \times 11, 3 \times 3, 5 \times 5$	95	98	71	86	100
$11 \times 11, 5 \times 5, 7 \times 7$	94	97	94	82	100
$11 \times 11, 3 \times 3, 7 \times 7$	82	98	95	80	100
$9 \times 9, 3 \times 3, 5 \times 5$	95	98	94	84	100
$9 \times 9, 5 \times 5, 7 \times 7$	85	96	94	86	95
$9 \times 9, 3 \times 3, 7 \times 7$	92	97	89	83	100
$1 \times 1, 3 \times 3, 5 \times 5$	92	98	98	90	98
$1 \times 1, 5 \times 5, 7 \times 7$	91	98	98	84	98
$1 \times 1, 3 \times 3, 7 \times 7$	91	98	97	85	76
$3 \times 3, 5 \times 5, 7 \times 7$	**96.01**	**99.61**	**99.22**	**98.83**	**100**

Fig. 18. Evaluation of different learning rates on five databases.

rates. From Fig. 18, it is observed that the learning rate at 0.00001 obtained higher classification accuracy.

5. Conclusion

In this study, a DNN enabled IoT framework is introduced for fast and accurate detection of COVID-19. Five databases viz., D1, D2, D3, D4, and D5 are considered in this study to manifest the

efficiency of the proposed method over existing approaches. One of the key benefits of integrating IoT into healthcare is reducing the exposition to contagion and automating the diagnosis, thus making the medical staff more concentrated on patients. Connected to this, the DNN framework is employed to fetch more reliable and noise-invariant facets at various image patches. The proposed method acquires an average recognition accuracy

of 96.01%, 99.61%, 99.22%, 98.83%, and 100% respectively. Experimental outcomes also manifest that the proposed method outranks fourteen contemporary approaches by adopting the average time i.e., training and testing time. Compared to the existing methods, the proposed model predicts very few FP's and FN's, which is shown in Fig. 6. Furthermore, it is worth investigating to deploy the proposed model in some real-life settings. The results obtained in this study are very promising and this work can be extended by considering multiple factors in the future. For future work, we intend to enhance the diversity of the database by adding new X-ray images of patients with COVID-19, as soon as these images are available, and by including X-ray exams of other lung-related diseases. Further, more efforts will be given to exploring how to identify COVID-19 in the early stages and how the prior attention mechanism can be employed in other medical image analysis problems.

CRediT authorship contribution statement

Mohan Karnati: Conception and design of study, Acquisition of data, Analysis and/or interpretation of data, Drafting the manuscript, Revising the manuscript critically for important intellectual content, . **Ayan Seal:** Conception and design of study, Acquisition of data, Analysis and/or interpretation of data, Drafting the manuscript, Revising the manuscript critically for important intellectual content. **Geet Sahu:** Conception and design of study, Acquisition of data, Analysis and/or interpretation of data, Drafting the manuscript, Revising the manuscript critically for important intellectual content. **Anis Yazidi:** Conception and design of study, Acquisition of data, Analysis and/or interpretation of data, Drafting the manuscript, Revising the manuscript critically for important intellectual content. **Ondrej Krejcar:** Conception and design of study, Acquisition of data, Analysis and/or interpretation of data, Drafting the manuscript, Revising the manuscript critically for important intellectual content.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

References

[1] Y.-H. Xu, J.-H. Dong, W.-M. An, X.-Y. Lv, X.-P. Yin, J.-Z. Zhang, L. Dong, X. Ma, H.-J. Zhang, B.-L. Gao, Clinical and computed tomographic imaging features of novel coronavirus pneumonia caused by SARS-CoV-2, J. Infection 80 (4) (2020) 394–400.

[2] worldometere, Covid-19 coronavirus pandemic, 2020, URL https://www.worldometers.info/coronavirus/?fbclid=IwAR2UgycDn8i64zB71xUGm5svanZxQEI_U6IEEzgiNRtMnVLtBQtyKqPW_e8.

[3] E. Mahase, Coronavirus: covid-19 has killed more people than SARS and mers combined, despite lower case fatality rate, 2020.

[4] L. Yan, H.-T. Zhang, Y. Xiao, M. Wang, Y. Guo, C. Sun, X. Tang, L. Jing, S. Li, M. Zhang, et al., Prediction of criticality in patients with severe Covid-19 infection using three clinical features: a machine learning-based prognostic model with clinical data in wuhan, 2020, medRxiv.

[5] A. Singh, S.K. Chandra, M.K. Bajpai, Study of non-pharmacological interventions on COVID-19 spread, 2020, medRxiv.

[6] W. Wang, Y. Xu, R. Gao, R. Lu, K. Han, G. Wu, W. Tan, Detection of SARS-CoV-2 in different types of clinical specimens, JAMA 323 (18) (2020) 1843–1844.

[7] C. Vermeiren, X. Marchand-Senécal, E. Sheldrake, D. Bulir, M. Smieja, S. Chong, J.D. Forbes, K. Katz, Comparison of copan eswab and floqswab for COVID-19 PCR diagnosis: working around a supply shortage, J. Clin. Microbiol. (2020).

[8] E. Prompetchara, C. Ketloy, T. Palaga, Immune responses in COVID-19 and potential vaccines: Lessons learned from SARS and MERS epidemic, Asian. Pac. J. Allergy Immunol. 38 (1) (2020) 1–9.

[9] Y. Fang, H. Zhang, J. Xie, et al., Sensitivity of chest ct for covid-19: comparison to rt-pcr, Radiology 200432 (2020).

[10] Z. Ye, Y. Zhang, Y. Wang, Z. Huang, B. Song, Chest CT manifestations of new coronavirus disease 2019 (COVID-19): a pictorial review, Eur. Radiol. (2020) 1–9.

[11] B. Wang, S. Jin, Q. Yan, H. Xu, C. Luo, L. Wei, W. Zhao, X. Hou, W. Ma, Z. Xu, et al., AI-assisted CT imaging analysis for COVID-19 screening: Building and deploying a medical AI system, Appl. Soft Comput. 98 (2021) 106897.

[12] H.Y.F. Wong, H.Y.S. Lam, A.H.-T. Fong, S.T. Leung, T.W.-Y. Chin, C.S.Y. Lo, M.M.-S. Lui, J.C.Y. Lee, K.W.-H. Chiu, T.W.-H. Chung, et al., Frequency and distribution of chest radiographic findings in patients positive for COVID-19, Radiology 296 (2) (2020) E72–E78.

[13] A.K. Gupta, A. Seal, P. Khanna, A. Yazidi, O. Krejcar, Gated contextual features for salient object detection, IEEE Trans. Instrum. Measur., 1–13.

[14] A.K. Gupta, A. Seal, P. Khanna, E. Herrera-Viedma, O. Krejcar, Almnet: Adjacent layer driven multiscale features for salient object detection, IEEE Trans. Instrum. Meas. 70 (2021) 1–14, http://dx.doi.org/10.1109/TIM.2021.3108503.

[15] K. Mohan, A. Seal, O. Krejcar, A. Yazidi, FER-net: facial expression recognition using deep neural net, Neural Comput. Appl., 1–12.

[16] K. Mohan, A. Seal, O. Krejcar, A. Yazidi, Facial expression recognition using local gravitational force descriptor-based deep convolution neural networks, IEEE Trans. Instrum. Meas. 70 (2020) 1–12.

[17] M. Karnati, A. Seal, A. Yazidi, O. Krejcar, Lienet: A deep convolution neural networks framework for detecting deception, IEEE Trans. Cogn. Dev. Syst. (2021).

[18] I. Yaqoob, E. Ahmed, I.A.T. Hashem, A.I.A. Ahmed, A. Gani, M. Imran, M. Guizani, Internet of things architecture: Recent advances, taxonomy, requirements, and open challenges, IEEE Wirel. Commun. 24 (3) (2017) 10–16.

[19] L. Wang, Z.Q. Lin, A. Wong, Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images, Sci. Rep. 10 (1) (2020) 1–12.

[20] T. Rahman, COVID-19 radiography database, 2020, URL https://www.kaggle.com/tawsifurrahman/covid19-radiography-database.

[21] S. Minaee, R. Kafieh, M. Sonka, S. Yazdani, G.J. Soufi, Deep-covid: Predicting covid-19 from chest x-ray images using deep transfer learning, Med. Image Anal. 65 (2020) 101794.

[22] A.I. Khan, J.L. Shah, M.M. Bhat, Coronet: A deep neural network for detection and diagnosis of COVID-19 from chest x-ray images, Comput. Methods Programs Biomed. (2020) 105581.

[23] C. Li, M. Wang, G. Wu, K. Rana, N. Charoenkitkarn, J. Chan, COVID19 chest X-ray classification with simple convolutional neural network, in: CSBio'20: Proceedings of the Eleventh International Conference on Computational Systems-Biology and Bioinformatics, 2020, pp. 97–100.

[24] Q. Yan, B. Wang, D. Gong, C. Luo, W. Zhao, J. Shen, J. Ai, Q. Shi, Y. Zhang, S. Jin, et al., Covid-19 chest CT image segmentation network by multi-scale fusion and enhancement operations, IEEE Trans. Big Data 7 (1) (2021) 13–24.

[25] R. Jain, M. Gupta, S. Taneja, D.J. Hemanth, Deep learning based detection and analysis of COVID-19 on chest X-ray images, Appl. Intell. (2020) 1–11.

[26] E.E.-D. Hemdan, M.A. Shouman, M.E. Karar, Covidx-net: A framework of deep learning classifiers to diagnose covid-19 in x-ray images, 2020, arXiv preprint arXiv:2003.11055.

[27] E.F. Ohata, G.M. Bezerra, J.V.S. das Chagas, A.V.L. Neto, A.B. Albuquerque, V.H.C. de Albuquerque, P.P. Reboucas Filho, Automatic detection of COVID-19 infection using chest X-ray images through transfer learning, IEEE/CAA J. Autom. Sin. 8 (1) (2020) 239–248.

[28] S. Tabik, A. Gómez-Ríos, J.L. Martín-Rodríguez, I. Sevillano-García, M. Rey-Area, D. Charte, E. Guirado, J.-L. Suárez, J. Luengo, M. Valero-González, et al., Covidgr dataset and COVID-sdnet methodology for predicting COVID-19 based on chest X-ray images, IEEE J. Biomed. Health Inf. 24 (12) (2020) 3595–3605.

[29] J.D. Arias-Londoño, J.A. Gomez-Garcia, L. Moro-Velázquez, J.I. Godino-Llorente, Artificial intelligence applied to chest X-ray images for the automatic detection of COVID-19. a thoughtful evaluation approach, IEEE Access 8 (2020) 226811–226827.

[30] J. Wang, Y. Bao, Y. Wen, H. Lu, H. Luo, Y. Xiang, X. Li, C. Liu, D. Qian, Prior-attention residual learning for more discriminative COVID-19 screening in CT images, IEEE Trans. Med. Imaging 39 (8) (2020) 2572–2583.

[31] I.D. Apostolopoulos, T.A. Mpesiana, Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks, Phys. Eng. Sci. Med. (2020) 1.

[32] T. Ozturk, M. Talo, E.A. Yildirim, U.B. Baloglu, O. Yildirim, U.R. Acharya, Automated detection of COVID-19 cases using deep neural networks with X-ray images, Comput. Biol. Med. 121 (2020) 103792.

[33] A. Narin, C. Kaya, Z. Pamuk, Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks, Pattern Anal. Appl. 24 (3) (2021) 1207–1220.

[34] T.T. Nguyen, Artificial intelligence in the battle against coronavirus (COVID-19): a survey and future research directions, 10, 2020, Preprint.

[35] H.S. Maghdid, K.Z. Ghafoor, A.S. Sadiq, K. Curran, K. Rabie, A novel ai-enabled framework to diagnose coronavirus covid 19 using smartphone embedded sensors: Design study, 2020, arXiv preprint arXiv:2003.07434.

[36] A.S.S. Rao, J.A. Vazquez, Identification of COVID-19 can be quicker through artificial intelligence framework using a mobile phone–based survey when cities and towns are under quarantine, Infect. Control Hosp. Epidemiol. 41 (7) (2020) 826–830.

[37] Z. Allam, D.S. Jones, On the coronavirus (COVID-19) outbreak and the smart city network: universal data sharing standards coupled with artificial intelligence (AI) to benefit urban health monitoring and management, in: Healthcare, Vol. 8, (1) Multidisciplinary Digital Publishing Institute, 2020, p. 46.

[38] M. Otoom, N. Otoum, M.A. Alzubaidi, Y. Etoom, R. Banihani, An IoT-based framework for early identification and monitoring of COVID-19 cases, Biomed. Signal Process. Control 62 (2020) 102149.

[39] M. Zhang, R. Chu, C. Dong, J. Wei, W. Lu, N. Xiong, Rldd: An advanced residual learning diagnosis detection system for COVID-19 in iIoT, IEEE Trans. Ind. Inf. (2021).

[40] C.M. Dourado, S.P.P. Da Silva, R.V.M. Da Nóbrega, P.P. Rebouças Filho, K. Muhammad, V.H.C. De Albuquerque, An open IoHT-based deep learning framework for online medical image recognition, IEEE J. Sel. Areas Commun. 39 (2) (2020) 541–548.

[41] D.D.A. Rodrigues, R.F. Ivo, S.C. Satapathy, S. Wang, J. Hemanth, P.P. Reboucas Filho, A new approach for classification skin lesion based on transfer learning, deep learning, and IoT system, Pattern Recognit. Lett. 136 (2020) 8–15.

[42] Q. Hu, E.F. Ohata, F.H. Silva, G.L. Ramalho, T. Han, P.P. Reboucas Filho, A new online approach for classification of pumps vibration patterns based on intelligent IoT system, Measurement 151 (2020) 107138.

[43] C.M. Dourado Jr., S.P.P. da Silva, R.V.M. da Nobrega, A.C.d.S. Barros, P.P. Reboucas Filho, V.H.C. de Albuquerque, Deep learning IoT system for online stroke detection in skull computed tomography images, Comput. Netw. 152 (2019) 25–39.

[44] D. Dansana, R. Kumar, A. Bhattacharjee, D.J. Hemanth, D. Gupta, A. Khanna, O. Castillo, Early diagnosis of COVID-19-affected patients based on X-ray and computed tomography images using deep learning algorithm, Soft Comput. (2020) 1–9.