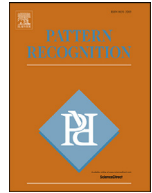




Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



# MetaCOVID: A Siamese neural network framework with contrastive loss for $n$ -shot diagnosis of COVID-19 patients

Mohammad Shorfuzzaman<sup>a</sup>, M. Shamim Hossain<sup>b,c,\*</sup>

<sup>a</sup> Department of Computer Science, College of Computers and Information Technology (CCIT), Taif University, Taif, Saudi Arabia

<sup>b</sup> Chair of Pervasive and Mobile Computing, King Saud University, Riyadh 11543, Saudi Arabia

<sup>c</sup> Department of Software Engineering, College of Computer and Information Sciences, King Saud University, Riyadh 11543, Saudi Arabia

## ARTICLE INFO

### Article history:

Received 3 August 2020

Revised 23 September 2020

Accepted 13 October 2020

Available online 17 October 2020

### Keywords:

COVID-19 diagnosis

Multi-shot learning

Contrastive loss

CXR images

Siamese network

## ABSTRACT

Various AI functionalities such as pattern recognition and prediction can effectively be used to diagnose (recognize) and predict coronavirus disease 2019 (COVID-19) infections and propose timely response (remedial action) to minimize the spread and impact of the virus. Motivated by this, an AI system based on deep meta learning has been proposed in this research to accelerate analysis of chest X-ray (CXR) images in automatic detection of COVID-19 cases. We present a synergistic approach to integrate contrastive learning with a fine-tuned pre-trained ConvNet encoder to capture unbiased feature representations and leverage a Siamese network for final classification of COVID-19 cases. We validate the effectiveness of our proposed model using two publicly available datasets comprising images from normal, COVID-19 and other pneumonia infected categories. Our model achieves 95.6% accuracy and AUC of 0.97 in diagnosing COVID-19 from CXR images even with a limited number of training samples.

© 2020 Elsevier Ltd. All rights reserved.

## 1. Introduction

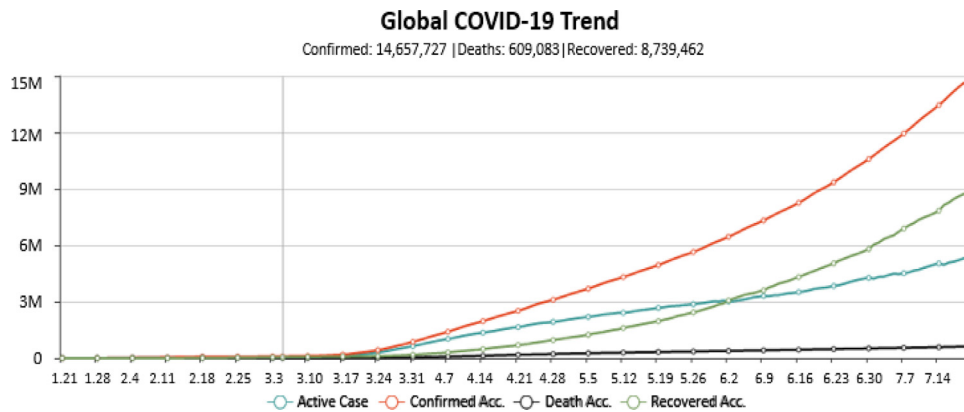
The coronavirus disease 2019 (COVID-19) initially identified in December 2019 in the city of Wuhan in China has rapidly spread throughout the world within a very short period of time resulting in an ongoing pandemic. Since the outbreak it has affected over two hundred countries and territories across the globe with more than 14.5 million cases reported. Fig. 1 depicts the global trend of COVID-19 as of July 20, 2020 including total number of confirmed, active, death, and recovered cases [1]. The outbreak was declared as Public Health Emergency of International Concern (PHEIC) by the World Health Organization (WHO) [2] on January 30, 2020. The virus is extremely contagious and is mostly transmitted between individuals through close contact. To stop its rapid transmission, it is crucial to gain a good understanding of the genetic characteristics of the virus. The genome size of the virus varies from approximately 26 to 32 kgbases being one of the largest among single stranded RNA (ribonucleic acid) viruses. The average diameter of the particle of the virus is around 120 nm [3]. Various common symptoms are found in the infected patients such as cough, fever,

shortness of breath, fatigue, loss of smell, and pneumonia. Health complications include pneumonia, acute respiratory distress syndrome, and other infections. Precise and on time diagnosis are being hampered due to undiscovered treatment, scarcity of resources, and harsh conditions of laboratory environment. This has heightened the challenges to reduce the spread of the virus. Accurate and speedy identification of suspected patients at the early phase may possibly play a critical role in timely quarantine and progressive cure. Thus, swift identification of potential infection by coronavirus is incredibly crucial for timely control of epidemic and public health welfare.

Identification of coronavirus infection is primarily done by nucleic acid test also called a PCR (polymerase chain reaction) test which examines for the existence of antibodies for an infection. However, results from recent studies show that this type of pathogenic laboratory testing though being a diagnostic gold standard suffers from limitations since it is time-consuming and produces high false negative cases [5]. Furthermore, deploying COVID-19 tests at a large scale is very expensive and is not affordable by many developing and underdeveloped countries. Hence, development of Artificial Intelligence (AI) based diagnosis and testing methods will be very beneficial. Variety of such AI based medical applications include cancer detection [15], Epilepsy Seizure Detection [16], pathology detection [26] multi-modal skin lesion

\* Corresponding author.

E-mail addresses: [m.shorf@tu.edu.sa](mailto:m.shorf@tu.edu.sa) (M. Shorfuzzaman), [mshossain@ksu.edu.sa](mailto:mshossain@ksu.edu.sa) (M.S. Hossain).



**Fig. 1.** COVID-19 trend in global scale. Graph shows total number of confirmed, active, death, and recovered cases. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

classification [17], polyp detection [23] and so on. In favor of this, researchers are taking global initiatives to use AI as a potentially powerful tool to come up with cost-effective and fast diagnostic procedures to control the ongoing epidemic [32]. The key research goals include COVID-19 transmission, its early diagnosis, development of effective treatment, and understanding its socio-economic impact [4].

Computer aided diagnosis (CAD) systems capable of processing chest X-ray (CXR) images and computed tomography (CT) scans along with state-of-the-art deep learning techniques could be very beneficial for the health professionals in diagnosing COVID-19 cases. Some studies in the literature have already demonstrated the effectiveness of using various deep learning techniques [27] to identify positive COVID-19 cases from chest X-ray (CXR) images and computed tomography (CT) scans [24] and to monitor the disease progress over time. Since deep learning algorithms generally require huge amount of training data to produce effective prediction results, the existing methods trained on limited training samples (due to the lack of large COVID-19 public dataset availability) are likely to suffer from model generalizability to new data. To alleviate this problem of data scarcity, researchers adopted various techniques such as data augmentation and generative adversarial network (GAN) [6,7]. Nevertheless, these techniques are highly dependent on the appropriate selection of parameters. Various hand-tuned data augmentation technique suffers from over-fitting problem [6] whereas techniques related to generating images through GAN face challenges in emulating real patient data which leads to unanticipated bias during model testing [7]. Furthermore, some studies have adopted transfer learning technique by using various pre-trained CNN models. Since these CNN models are pre-trained on a large non-medical dataset (i.e., ImageNet), substantial amount of fine-tuning which generally requires longer training period is necessary to produce promising diagnostic results.

Lately,  $n$ -shot (specially, one-shot and few-shot) learning has gained immense popularity in research community for analyzing medical images with a limited sample size. In general terms, for example, one-shot learning refers to the task of classifying an image to a particular class given a single (or few) training samples of each class. Specifically, one or more samples from each image class are used to prepare (train) the model which in turn can classify unseen images in future. One of the meta learning models that recently gained success in implementing few-shot learning (especially one-shot) in various domains is Siamese network. In a Siamese network architecture, identical deep convolutional neural networks (CNNs) are trained to extract feature vectors discriminating between samples of each image class which are then contrasted to verify the similarity of the input images. This paper

presents a trainable  $n$ -shot deep meta learning framework to classify COVID-19 cases with limited training CXR images. We use a fine-tuned CNN model called VGG16 [8] as backbone encoder network to generate feature embeddings from the input images and leverage pairwise contrastive loss function to adjust the network weights. More specifically, we have used CXR images from two public datasets to pre-train the embedding CNN network to generate feature representations that are used by the Siamese network which learns a metric space for  $n$ -shot classification of unseen images without any retraining. In summary, following are the contributions of our work: (a) A meta learning framework called Meta-COVID based on Siamese neural network is presented for diagnosis of COVID-19 patients from chest X-ray images, (b) The proposed work focuses on the benefit of using contrastive loss and  $n$ -shot learning in framework design, (c) A fine-tuned pre-trained VGG encoder is used to capture unbiased feature representations to improve feature embeddings from the input images, (d) The COVID-19 diagnosis problem is formulated as a  $k$ -way,  $n$ -shot classification problem where  $k$  and  $n$  represent the number of class labels and data samples used for model training, (e) Performance evaluation is presented to demonstrate the efficacy of the proposed framework with a limited dataset.

The remainder of the paper is organized as follows. Related work is presented in Section 2. Section 3 and 4 present method, dataset, and experiments with performance results. Lastly, Section 5 concludes the paper with future work.

## 2. Related work

Relevant to the proposed research, our literature study will largely contain existing research effort in the area of COVID-19 diagnosis using AI techniques. Deep learning which is a specialized form of machine learning in the domain of AI has shown great potential in medical image analysis during the last decade [9]. Substantial research has been conducted using deep learning in various medical fields such as disease prediction, diagnosis of pulmonary nodules, and classification of benign and malignant tumors and so on. According a recent study from the researchers at UN Global Pulse [10], it is shown that AI applications can be as accurate as humans in detecting COVID-19 and offer faster and cheaper solutions in diagnosing the virus than standard test kits thus saving radiologists' valuable time. As part of this, researchers are primarily concentrating on techniques based on statistical learning for the detection of potential coronavirus infection from CXR images and computed tomography (CT scans). Some research initiatives in progress are provided below.

A relatively earlier (in the beginning of the outbreak) effort done by a group of researchers in Renmin University of Wuhan, China [11] proposed an AI model for the diagnosis of COVID-19 cases using CT scans. The model uses UNet++ [12] architecture for coronavirus detection using CT scan features and makes use of more than 40,000 images from 106 patients for model training. Experimental results demonstrate that the radiologists' efforts in terms of time can substantially be decreased by using this model.

Xu et al. [13] presented a deep learning model to screen coronavirus disease from viral pneumonia (of type Influenza-A) and normal cases with pulmonary CT scans. They have first identified candidate infection regions using segmentation and the separated images are then classified using a classification model based on location attention. The model was trained and evaluated using 618 CT samples consisting of 219 COVID-19, 224 Influenza-A, and 175 healthy cases. The samples were collected from three hospitals in China that are designated for COVID-19 treatment. The model showed only a moderate level (86.7%) of accuracy score using the curated dataset. In a subsequent effort, Wang et al. [5] proposed a robust diagnostic model based on deep neural network which works based on graphical features generated from CT scans and saves crucial disease control time. They used a fine-tuned Inception [14] pre-trained model in their architecture and utilized a dataset containing CT scans of COVID-19 patients and patients with other non-COVID viral pneumonia. Mei et al. [46] proposed a framework that leverages deep learning techniques to diagnose COVID-19 from CT scans using deep neural networks from scratch.

Besides CT scans, several studies have used CXR images for the detection of COVID-19. Since it is relatively easier to find CXR images than CT scans especially in rural areas, they can be a viable alternative to CT images. Wang and Wong [18] proposed an AI system called COVID-Net to diagnose COVID-19 from chest X-ray images containing samples from healthy, COVID-19 and other pneumonia infected patients. The limitation of this study is that the authors trained and tested their model using an imbalanced dataset which contains very few (less than 100) COVID-19 images as opposed to about 16,000 images from healthy and other non-COVID pneumonia patients. Additionally, authors in [19] presented a similar study that uses CXR images to detect coronavirus infection through transfer learning mechanism. Chakraborty [20] also developed a CXR based model using deep neural network that can achieve significant performance improvement even when the size of the dataset is limited. Nevertheless, the model lacks generalizability and needs fine-tuning to produce more stable results.

Another laudable effort made by researchers from Delft Imaging project [21] which developed an AI model for diagnosing COVID-19 from CXR images. Their model is called CAD4COVID which is built upon an existing AI model previously developed for diagnosing tuberculosis. It triages COVID-19 suspected patients. Hossain et al. [38] presented a healthcare framework based on 5 G network that makes use of CXR and CT images for COVID-19 in interpreting their predictions by extracting critical features related to COVID-19 to gain deeper understanding. Abbas et al. [22] have also leveraged deep learning techniques for the diagnosis of COVID-19 from CXR images using CNN and other pre-trained models such as ResNet [25].

Although a recent report [28] has shown the success of a number of Chinese hospitals in deploying AI-assisted radiology technologies in combating COVID-19, radiologists have shown their concern [29] that the shortage of available data to train the AI diagnostic models is a major challenge. This is substantiated by the fact that large body of the AI models in the literature have used datasets containing limited COVID-19 samples.

To address the challenge of limited training samples, He et al. [30] proposed a framework based on deep neural network that is able to attain significant improvement in accuracy in COVID-19 de-

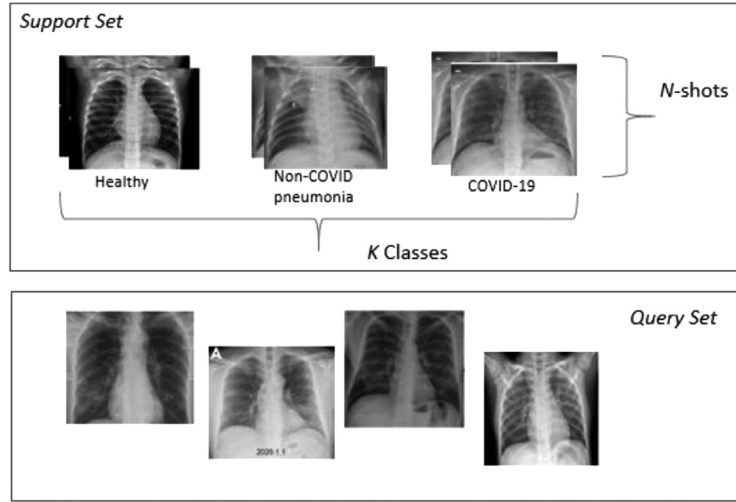
tection with a limited dataset consisting of CT scans. They have developed a synergistic approach to combine transfer learning with self-supervised contrastive learning to extract unbiased features to avoid overfitting problem. The experimental results demonstrate the superiority of their approach in comparison with several state-of-the-art models. In a subsequent effort, Chen et al. [31] develop a meta learning model with prototypical network to predict coronavirus infections from chest CT images that requires a small dataset for training. They have used momentum based contrastive learning to extract feature vectors from the input images that are used by the prototypical network to make predictions on unseen CT images for potential COVID-19 encounters. Validation results with two publicly available CT scans datasets suggest the effectiveness of their model compared to several other relevant methods.

Existing techniques in the literature that use CXR images for COVID-19 diagnosis mostly use custom CNN architecture or pre-trained transfer learning models which require large training data to produce rich feature encoding. In contrast, we have proposed an end-to-end trainable  $n$ -shot deep meta learning framework based on Siamese neural network to classify COVID-19 cases with limited training CXR images. Our proposed model is computationally efficient that can achieve better or the same level of performance as the pre-trained and other custom CNN models that require longer training time. Also, techniques in the literature mostly use de-facto categorical or binary cross entropy loss function. In contrast, we have used contrastive loss function which results in faster model convergence with fewer experiments and hyperparameter updates. Furthermore, most of the existing models use image augmentation to improve model generalizability even with prolonged training time. Alternatively, the proposed meta learning framework shows faster model convergence and greater generalizability by using contrastive loss function and appropriate hyper-parameter optimization such as learning schedule and regularization through dropout technique used in various layers in the model.

### 3. Problem formulation

The shortage of COVID-19 CXR images brings substantial challenges in developing tools for effectively diagnosing COVID-19 cases using deep learning based techniques. To tackle this challenge, we leverage deep meta learning and devise the diagnosis problem of COVID-19 from CXR images as a  $n$ -shot classification problem which can be considered as an instantiation of meta learning in the paradigm of supervised learning. Meta learning has recently emerged as a trending research area in the field of Artificial Intelligence (AI) and is believed to be a steppingstone for attaining Artificial General Intelligence (AGI). It is often referred to as "learning to learn" and is capable of learning new skills and generalizing to new tasks quickly by means of limited training samples. The aim of  $n$ -shot learning is to make classification of unseen data given only a limited set of training examples (of size  $n$ ). At one end, if there is only a single example of each class it is referred to as one-shot learning. On the other hand, few-shot learning requires that each class should have a few training examples, typically, up to five.

The problem of  $n$ -shot image classification can be defined as  $K$ -way  $N$ -shot episodic task where  $K$  denotes the number of target class labels in the dataset and  $N$  denotes the number of available images (samples) for each of the classes. Given, we have a dataset,  $D$ , we sample  $N$  data points (images) from each of the  $K$  classes present in our data set and we call it as support set. Similarly, we sample  $Q$  different images from each of the classes and call it as query set. The goal is to classify the images of the query set based on  $K$  classes and  $KN$  total images available in the support set. Fig. 2 demonstrates an  $n$ -shot classification scenario in visual form. In our case, we have three different types CXR images



**Fig. 2.** An example of a  $K$ -way  $N$ -shot learning problem where  $K = 3$  and  $N = 2$  in the support set. Query set images need to be classified from 3 available classes {normal, COVID-19 positive, and non-COVID pneumonia}.

in our dataset, namely, normal, COVID-19 positive, and non-COVID pneumonia cases. Hence, we view the diagnosis problem both as 3-class (images from all three categories) and 2-class (normal and COVID-19 positive images) classification problem. In this context, our COVID-19 diagnosis problem can be identified as three-way  $N$ -shot and two-way  $N$ -shot learning problem.

#### 4. Methodology

This section presents the generic high-level architecture of our deep Siamese network (as shown in Fig. 3) for  $n$ -shot learning to diagnose COVID-19 cases which is widely known as metric learning-based approach to meta learning. We will describe all major components of the architecture including base CNN encoder for feature embedding, contrastive learning, and training strategies.

##### 4.1. Metric-based learning

In the metric-based meta learning setting, the goal is to learn the appropriate metric space. The fundamental concept is mostly related to nearest neighbors techniques and kernel density estimation where the predicted probability is calculated across a set of given output labels,  $y$ , as a weighted sum of labels of the examples from the support set. A kernel function,  $f_{\theta}$ , is used to create the weight which essentially estimates the similarity between two different input data points. The predicted probability over the samples of a support set,  $S$ , can be formulated as:

$$P_{\theta}(y|x, S) = \sum_{(x_i, y_i) \in S} y_i \cdot f_{\theta}(x, x_i) \quad (1)$$

The performance of a metric-based meta learning model largely depends on learning an appropriate kernel. A useful metric would represent the relationship among the inputs such as similarity in the latent space to facilitate highly accurate predictions. In our case, we want to learn the similarity between two images. For this purpose, a convolutional neural network is used to extract the features from two images and finds the similarity by computing the distance between features of these two images. This approach is widely used in metric-based learning algorithms such as Siamese networks that we have used in our meta learning framework besides other metric-based networks such as matching, relation and prototypical networks.

##### 4.2. Deep Siamese network model

As shown in the preceding diagram, a Siamese network contains two identical parallel networks both sharing the same weights and architecture where each of the networks accepts a different input image and the output from them are combined to make the final prediction. More specifically, the goal is to have two identical base neural networks that take an actual image and another candidate image as input and can learn a function to produce the similarity output between these two images. The concern is that how we can essentially train such a neural network encoder that can learn this similarity function. Ideally, a convolutional neural network could be used without any constraints. Apparently, it would be desirable to use a custom CNN model which is smaller and computationally efficient and can achieve the same level of performance in feature encoding as the pre-trained models. However, such a CNN model requires large training data to produce rich feature encoding. Since we have a limited dataset, we leverage the power of CNN models pre-trained on large ImageNet [39] data which in recent times have shown promising results in solving computer vision problems such as medical imaging. Hence, we have used a fine-tuned pre-trained VGG-16 [8] as base encoder to obtain feature embeddings from the input images to ultimately compute similarity among them.

Let's consider that we have two input images,  $x_1$  and  $x_2$ . After passing the image first image,  $x_1$ , through the top encoder, we receive a feature embedding of  $x_1$  denoted as  $z_1(x_1) = \text{VGG16}(x_1)$  where  $z_1(x_1)$  is the output generated from the average pooling layer. Similarly, the second image,  $x_2$  is fed to the identical bottom encoder sharing the same weights,  $w$ , to get a different feature embedding of  $x_2$  denoted as  $z_2(x_2)$ . Then, in the latent space of feature embeddings, we feed these two embeddings to an energy function,  $E$ , which will give us the similarity between the two inputs. We use  $L_1$  component wise distance as our energy function which can be expressed as follows:

$$E_w(x_1, x_2) = d_w(x_1, x_2) = \|z_1(x_1, x_2) - z_2(x_1, x_2)\| \quad (2)$$

The value of  $E$  will be smaller if the input images ( $x_1$  and  $x_2$ ) are similar and vice versa. In reality, if it is less than a supplied threshold value the images are similar and, if not, they are different. Finally, this distance value can be incorporated in loss function (described in the next subsection) to tune the base encoder through back propagation for improved feature embeddings. A feedforward linear layer and sigmoid activation function are used

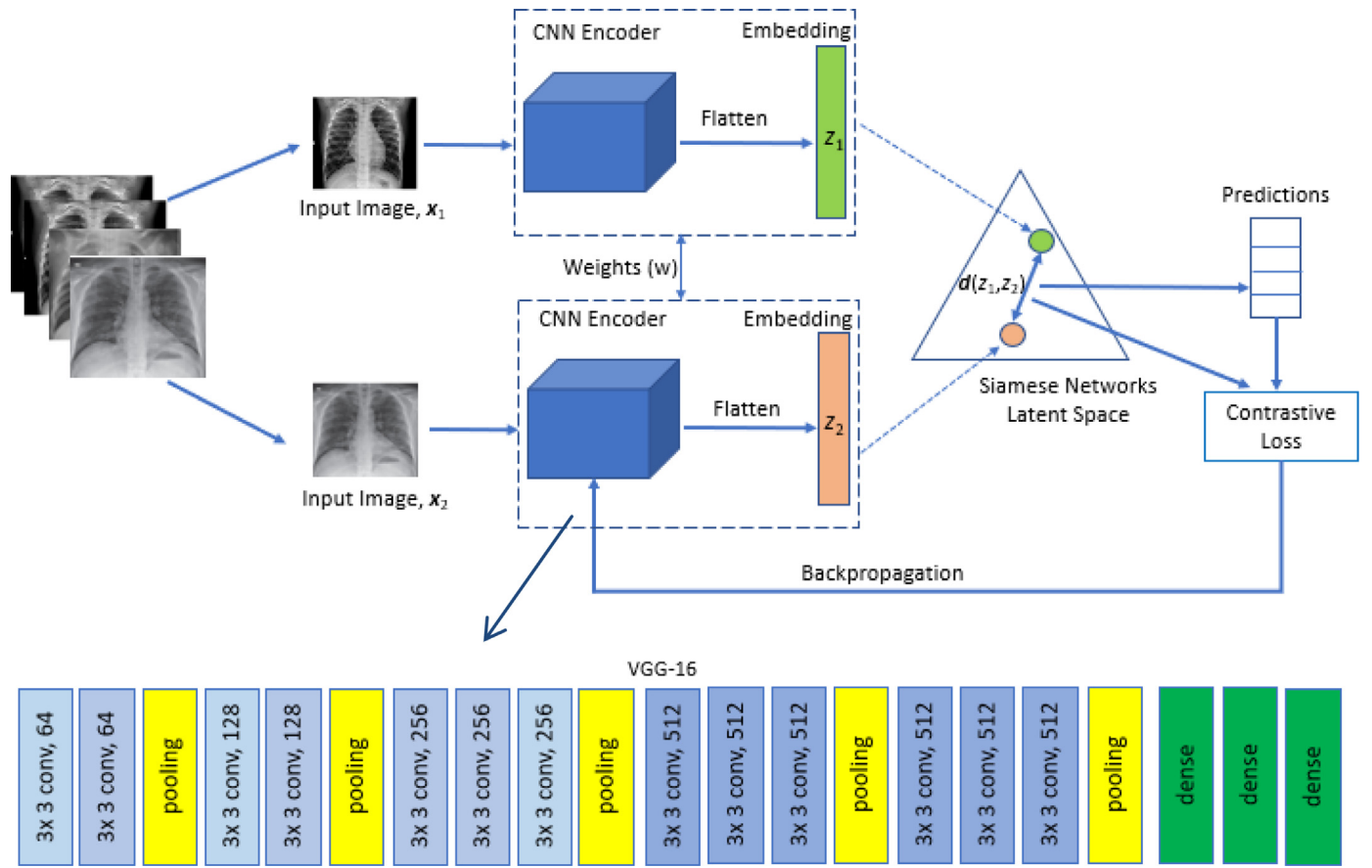


Fig. 3. High-level architecture of deep Siamese neural network for  $n$ -shot COVID-19 classification. (Zooming may be required for superior view). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

to convert the distance to a probability,  $p$ , which indicates whether the input images belong to the same target class or not. Given the input images from a support set,  $S$ , and a test image,  $t$ , the predicted target class is computed as follows where  $o(t)$  denotes the true class of the test image and  $\hat{o}(\cdot)$  is the predicted class.

$$\hat{o}_S(t) = o(\operatorname{argmax} p(t, x_i)), x_i \in S \quad (3)$$

#### 4.3. Loss function

Through the use of an appropriate loss function, the base encoder network can learn parameters to obtain a better encoding of the input image. Since Siamese networks make binary classification by classifying if the input images are similar or not, using binary cross-entropy loss function would be a natural choice. However, we have also considered using contrastive loss function due to the nature of classification strategy adopted by Siamese networks based on the similarity of pairs of input images. Hence, we have used both the loss functions in this study for performance evaluation.

##### 4.3.1. Contrastive loss

Originally proposed by Hadsell et al. [33], the contrastive loss function requires pairs of input samples as opposed to individual samples. The idea is that the base encoder network is penalized differently by the loss function in accordance with the classes of the input images. In particular, the loss function makes the model produce more similar feature embeddings if the target classes are the same and less similar feature embeddings if the classes are different. Mathematically, the contrastive loss is formulated as follows:

$$\text{Loss} = (1 - y) * \frac{1}{2} (d_w)^2 + (y) * \frac{1}{2} \{\max(0, m - d_w)\}^2 \quad (4)$$

In the preceding equation, the value of  $y$  is the true label, which will be 0 when the two input images are similar and 1 if they are dissimilar, and  $d_w$  is the distance measure between feature embeddings of the input images. Now,  $y$  equals to 0 implies that the amount of loss contributed by similar pairs would be simplified to the first term only and  $d_w$  is minimized. On the contrary if  $y = 1$  then the loss will be simplified to the second term and  $d_w$  is maximized to  $m$ , a hyperparameter called margin. Thus, when input pairs are dissimilar, and if their distance is greater than the margin, they do not incur a loss (as shown in Fig. 4).

##### 4.3.2. Binary cross entropy

Binary cross entropy loss also called log loss is used to estimate the performance of a classifier with an output probability ranging from 0 to 1. The loss value will increase if the predicted probability deviates from the true label. This can be formulated as follows where  $y$  and  $p$  represent the class label and prediction probability, respectively:

$$L = -y \log p + (1 - y) \log (1 - p) \quad (5)$$

This loss function is used to train the network so that it can differentiate between similar and dissimilar images if we provide one training example from positive and negative categories and aggregate both the losses as below:

$$L = L_{pos} + L_{neg} \quad (6)$$

#### 4.4. Training strategy

Formally, we have in our hand a  $k$ -way,  $n$ -shot classification problem for classifying CXR images. We have our dataset,  $D$ , with

**Algorithm 1**Training algorithm for  $k$ -way  $n$ -shot learning.

---

**Input:** Batch size  $N$ , Number of epochs  $numEpochs$ , Dataset  $D$ , fine-tuned VGG-16 encoder model  $M$  with parameter  $\theta$ , Loss function  $L$ , margin  $m$

**Initialize**  $posPairs$ ,  $negPairs$ ,  $posDist$ ,  $negDist$  for training and validation

$D_{train}$ ,  $D_{test}$  = split dataset,  $D$

$\theta_0 = w_0$

**for**  $i$  **do**  $numEpochs$

**for**  $b$  **do**  $getBatches()$

$X_b, Y_b$  = random batch from  $D_{train}$

$posPairs = getPositivePairs(X_b, Y_b)$

$negPairs = getNegativePairs(X_b, Y_b)$

$posDist_b = L1\_distance(M(posPairs, \theta_b))$  using Eq. (2)

$negDist_b = L1\_distance(M(negPairs, \theta_b))$  using Eq. (2)

$dist_b = \text{concat}(posDist_b, negDist_b)$

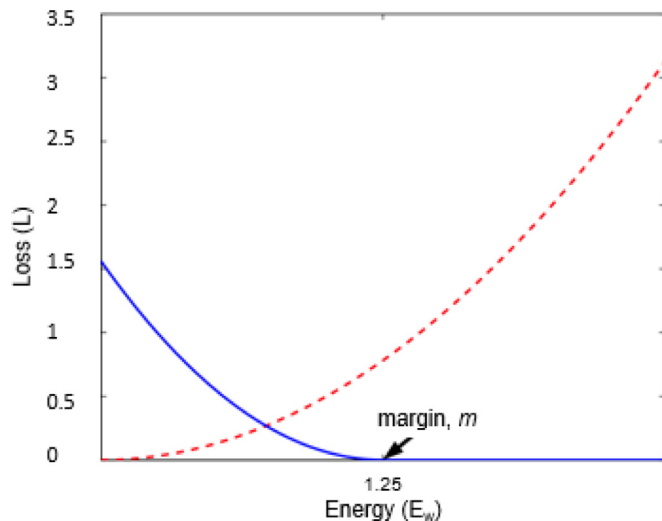
$L_b = Loss(dist_b, m, posPairs, negPairs)$  using Eq. (4)

    Update parameter  $\theta_b$  with new weight,  $w$

**end for**

**end for**

---



**Fig. 4.** Contrastive loss showing the margin  $m$ . The blue solid line signifies the loss function for the dissimilar pairs and the dotted red line refers to the same for similar pairs [33]. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

a training ( $D_{train}$ ) and a test ( $D_{test}$ ) split. Now, the training set will contain  $n$  samples from each of  $k$  classes thus totaling  $k.n$  examples in  $D_{train}$  and  $D_{test}$  consists of several samples for evaluation. We train our model in an episodic fashion which dictates that in each episode, we sample a few data points from our meta-training dataset,  $D_{train}$ , prepare our support set and query set, and train on the support set and test on the query set. So, over series of episodes, our model will learn how to learn from a smaller dataset.

Now, our model will be trained using a batch of training tasks  $\{Task_b\}$  to ultimately solve a test task,  $Task_{test}$ . We can consider that an episode consists of a classification tasks,  $Task_b$ , from the batch of tasks for training which is identical to the classification task,  $Task_{test}$  and will use  $k.n$  number of image samples for training and  $Q$  query images for evaluation. At the end of each episode the model parameters will be updated with the goal to maximize the classification performance of our model on  $Q$  query images through backpropagation of the calculated loss. In this way, our model will learn to solve an unseen task by gaining experience through a series of training tasks. An example  $n$ -shot classification training scenario is given in Fig. 5 where we create episodes consisting of tasks each of which is defined by a support set and a query set containing sample images from meta-training dataset. Moreover, each of these tasks is similar to the test  $n$ -shot classifica-

tion task containing images for support and query sets. Finally, the detailed training strategy of our model is shown in Algorithm 1.

## 5. Experiments

To demonstrate the effectiveness of our  $n$ -shot meta learning approach to COVID-19 diagnosis and inspect the effects of using different  $k$ -way,  $n$ -shot variants and loss functions, we extensively evaluate our proposed Siamese network model with two publicly available CXR datasets. Doctors frequently use CXR and CT scans for the diagnosis of various common diseases such as pneumonia, cancers, lung inflammation, and internal organ injuries. Given the fact that CXR imaging machines are available in nearly all hospitals we decide to use CXR images instead of using CT scans or other types of image data. In the subsequent sections, we will present the datasets used with preprocessing, experimental settings, and results with discussion.

### 5.1. Dataset description and preprocessing

In this study, we use COVID-19 patients' CXR images that are acquired from an open source database created by Dr. Joseph Cohen [34] in his GitHub repository. Dr. Joseph Cohen is continuously uploading CXR pictures of COVID-19 patients having acute respiratory distress syndrome (ARDS), COVID-19, pneumonia, Middle East respiratory syndrome (MERS), and severe acute respiratory syndrome (SARS). At present, there are about 230 CXR images of COVID-19 patients. In general, the CXR images of confirmed COVID-19 cases show various shapes of "pure ground glass" also referred to as hazy lung opacity during the disease development. To create a classifier from CXR images it is also necessary to have related CXR images of patients who do not have COVID-19. Fortunately, Kaggle [35] has a repository of CXR images of pneumonia and healthy patients. This dataset contains a total of 5856 images from both pneumonia and healthy categories. We have considered both image sources as a dataset to build, train, and test the proposed model. Since the second data source contains samples of healthy and other non-COVID pneumonia CXR images which are far more than COVID-19 positive image samples. To avoid creating an imbalanced dataset which may lead to ambiguous accuracies, we have only gathered 226 CXR images from each of all three categories such as normal, non-COVID pneumonia, and COVID-19 positive in the curated dataset. Thus, we build a final balanced dataset containing a total of 678 CXR. We divide the dataset for generating Support set and Query set based on various values of  $n$  in  $n$ -shot classification task. To pre-train the base VGG-16 encoder to be used in generating feature embeddings on the input images to the Siamese network, we use the same dataset consisting of CXR

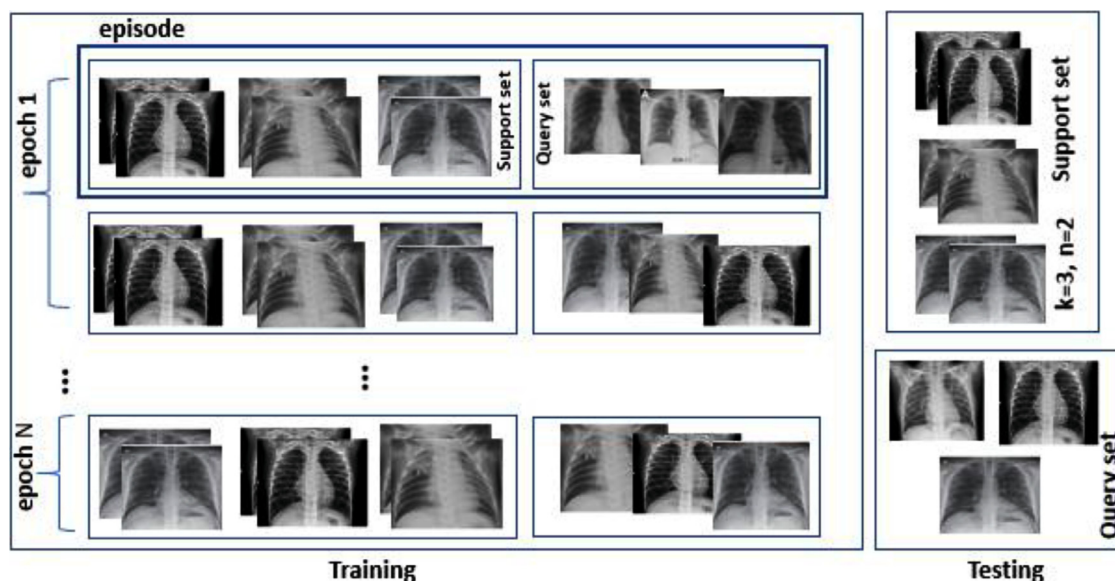


Fig. 5. An example training strategy for 2-shot, 3-class image classification task.

**Table 1**  
Dataset split statistics.

Class	Pre-training of VGG-16 encoder network		Siamese network ( $n$ -shot learning)	
	Training	Testing	Training	Testing
Normal	160	66	10	216
Non-COVID pneumonia	160	66	10	216
COVID-19	160	66	10	216
Total	480	198	30	648

images as mentioned above and split the dataset into training and validation sets with a ratio of 0.7:0.3. Table 1 shows the statistics of the dataset split.

### 5.1.1. Pre-processing

Due to the fact that the images in the dataset were collected from different locations with various clinical settings, the intensity and quality of images vary considerably. Nevertheless, we avoid extensive pre-processing of our CXR images in the dataset to gain improved generalization ability of our proposed Siamese network model. This in turn makes our model further robust to artifacts and noises present in the images while extracting feature embeddings from the input images. Thus, we only used few standard pre-processing tasks including image resizing, normalization, and histogram-equalization to optimize the model training method. The size of the CXR images in the dataset varies from  $365 \times 465$  to  $1125 \times 859$  pixels. Hence, we re-scale all images to a size of  $100 \times 100$  pixels to get a consistent image dimension for the entire dataset. Additionally, we perform intensity normalization also called scaling which is an important pre-processing task to expedite model convergence by eliminating feature biases and attaining a uniform distribution for the dataset. We convert the image pixel values from  $[0, 255]$  to  $[0, 1]$  to obtain a standard normal distribution by using min-max normalization technique. Finally, we apply histogram equalization on the input images in all three RGB channels to improve image contrast. This is usually done by effectively stretching out the most often used intensity values which allows the areas with poorer local contrast to achieve a better contrast.

### 5.2. Experimental settings

The pre-trained base encoder network and the proposed Siamese network models are implemented using TensorFlow. We use Google Colab notebook environment for model training and testing which provides free GPU access. It currently provides NVIDIA Tesla P100 GPU with 16GB RAM and comes with pre-installed Python 3.x packages and Keras API with backend TensorFlow.

Towards the end of the pretrained model we add a flatten layer which is followed by a dense layer with 5120 neurons, sigmoid activation function, and L2 kernel regularizer which is reported to have achieved significant performance improvement with a large number of kernels. Encodings (feature vectors) of the two input images are generated using this preceding dense layer. Then, we add a customized layer to the model to compute L1 distance by taking the absolute difference between the encodings. Finally, we add a dense layer with a sigmoid unit to generate the similarity score. We have used both contrastive and binary cross-entropy loss functions for model learning. In addition, Adam optimizer is used for model training and optimizing with an initial learning rate of 0.0001. Subsequently, the learning rate is updated using *ReduceLROnPlateau* callback provided by Keras which monitors the performance metric and if no improvement is observed for a 'patience' number of epochs the learning rate is reduced. Furthermore, we have used another callback from Keras called *EarlyStopping* to stop the training process when triggered upon monitoring some performance measure.

We evaluate our model using the following six metrics: accuracy, recall, precision, specificity, F1-score, and AUC (Area Under Curve). Here, accuracy calculates the proportion of predictions that



**Table 2**Performance results for various  $n$ -shot settings with contrastive loss. 3-way represents 3-class labels.

Model	Accuracy	Precision	Recall	Specificity	F1-score	AUC
MetaCOVID (3-way, 7-shot)	0.925	0.945	0.936	0.953	0.940	0.955
MetaCOVID (3-way, 8-shot)	0.936	0.951	0.945	0.965	0.938	0.962
MetaCOVID (3-way, 9-shot)	0.948	0.966	0.955	0.975	0.947	0.974
MetaCOVID 3-way, 10-shot)	0.956	0.970	0.960	0.980	0.965	0.975

**Table 3**Performance results for various 3-way,  $n$ -shot settings with cross entropy loss.

Model	Accuracy	Precision	Recall	Specificity	F1-score	AUC
MetaCOVID (3-way, 7-shot)	0.890	0.927	0.915	0.935	0.916	0.933
MetaCOVID (3-way, 8-shot)	0.915	0.935	0.919	0.940	0.922	0.948
MetaCOVID (3-way, 9-shot)	0.923	0.938	0.939	0.948	0.938	0.954
MetaCOVID 3-way, 10-shot)	0.938	0.949	0.953	0.964	0.950	0.957

**Table 4**

Performance comparison between the proposed Siamese network model (with 3-way, 10-shot learning) and other pre-trained CNN models.

Model	Acc.	Precision	Recall	Specificity	F1-score	AUC
InceptionV3	0.875	0.826	0.950	0.800	0.883	0.900
Xception	0.955	0.977	0.956	0.988	0.966	0.980
InceptionResNetV2	0.900	0.833	1.00	0.800	0.908	0.900
VGG-16	0.933	0.956	0.956	0.976	0.956	0.954
MetaCOVID (3-way, 10-shot)	0.956	0.970	0.960	0.980	0.965	0.975

precisely matches with the real values. Precision also called positive predictive value (PPV) is the fraction of true positive cases over all positive predictions. Recall or sensitivity also called true positive rate (TPR) which is very critical in medical applications refers to the fraction of all COVID-19 positive incidents that are accurately categorized as positive. Specificity also called true negative rate (TNR) refers to the proportion of all negative cases that are accurately categorized as negative. For example, the percentage of healthy people who are correctly diagnosed as not carrying the virus. F1-score denotes the harmonic mean which is calculated from precision and recall by taking their weighted average. Finally, AUC refers to the area under the receive operating characteristic (ROC) curve that demonstrates how TPR increases with the decrease in false positive rate (FPR).

### 5.3. Evaluation results and discussion

To evaluate the efficacy of our proposed meta-learning model in diagnosing COVID-19 cases we took the following approach. We start by evaluating our model for various 3-way,  $n$ -shot learning settings with both contrastive and cross-entropy losses. A comparison of performance results is also done with different pre-trained CNN models such as Inception [36], Xception [37], Inception ResNet v2 [36], VGG-16 [8]. Finally, we present the results obtained from 2-class (normal, COVID-19) variant of our classification problem with different  $n$ -shot learning settings.

Since our method is based on multi-shot learning, we are interested in investigating how does model performance change with the number of shots. As such, we carry a number of experiments to observe the relationship between the number of shots and the performance.

Table 2 shows the performance results of our Siamese network model for 3-class detection problem including normal, non-COVID pneumonia, and COVID-19 cases using  $n$ -shot meta learning strategy with contrastive loss where  $n$  varies from 7 to 10. It is apparent from the results that the accuracy of the model and other performance metrics gradually increase with the increased number of shots. This is due to the fact that with increased shots the model

exploits the benefit of more available pairs of images where it has to distinguish a similar image from different ones.

We also produce performance results (as shown in Table 3) with similar learning settings but with binary cross-entropy loss which is usually deemed to be a natural choice for classification problem. Generally, the performance results obtained with contrastive loss function seem to be better than the results obtained with cross-entropy loss function. This is due to the fact that our Siamese network model works based on similarity of pairs of images and contrastive loss function is reported in the literature to be more effective than cross-entropy loss. Moreover, as shown in Fig. 6, the model training and validation with contrastive loss function appears to be more stable and further shows better convergence even though with longer training epochs (approximately 175 epochs in contrast with approximately 100 epochs with cross-entropy loss). To make sure that our model does not overfit to training data, we have used dropout regularization technique in the fine-tuned base encoder. In addition, early stopping coupled with a learning schedule (ReduceLROnPlateau callback from Keras) is used to stop the training process at the right moment to reduce overfitting.

We also compare (as shown in Table 4) the performance of our Siamese network model with other pre-trained CNN models. It is noticed that our model (MetaCOVID) with 10-shot learning setting shows comparable or in some cases better performance than the best performing pre-trained Xception model in performance metrics including specificity, sensitivity, and accuracy. The proposed model produces impressive values of sensitivity (96.0%) and specificity (98.0%) which are deemed to be very critical performance estimates for applications in medical settings. This is promising owing to the fact that our Siamese meta learning model is trained only with a limited sample (10) of training examples from each category of CXR images. In addition, the proposed meta learning model is relatively smaller in size and contains comparatively fewer trainable parameters. It clearly shows the benefit of using meta learning approach to classifying COVID-19 patients over other contemporary methods.

Finally, to evaluate the effectiveness of our model in identifying COVID-19 cases from normal CXR images only (2-class prob-

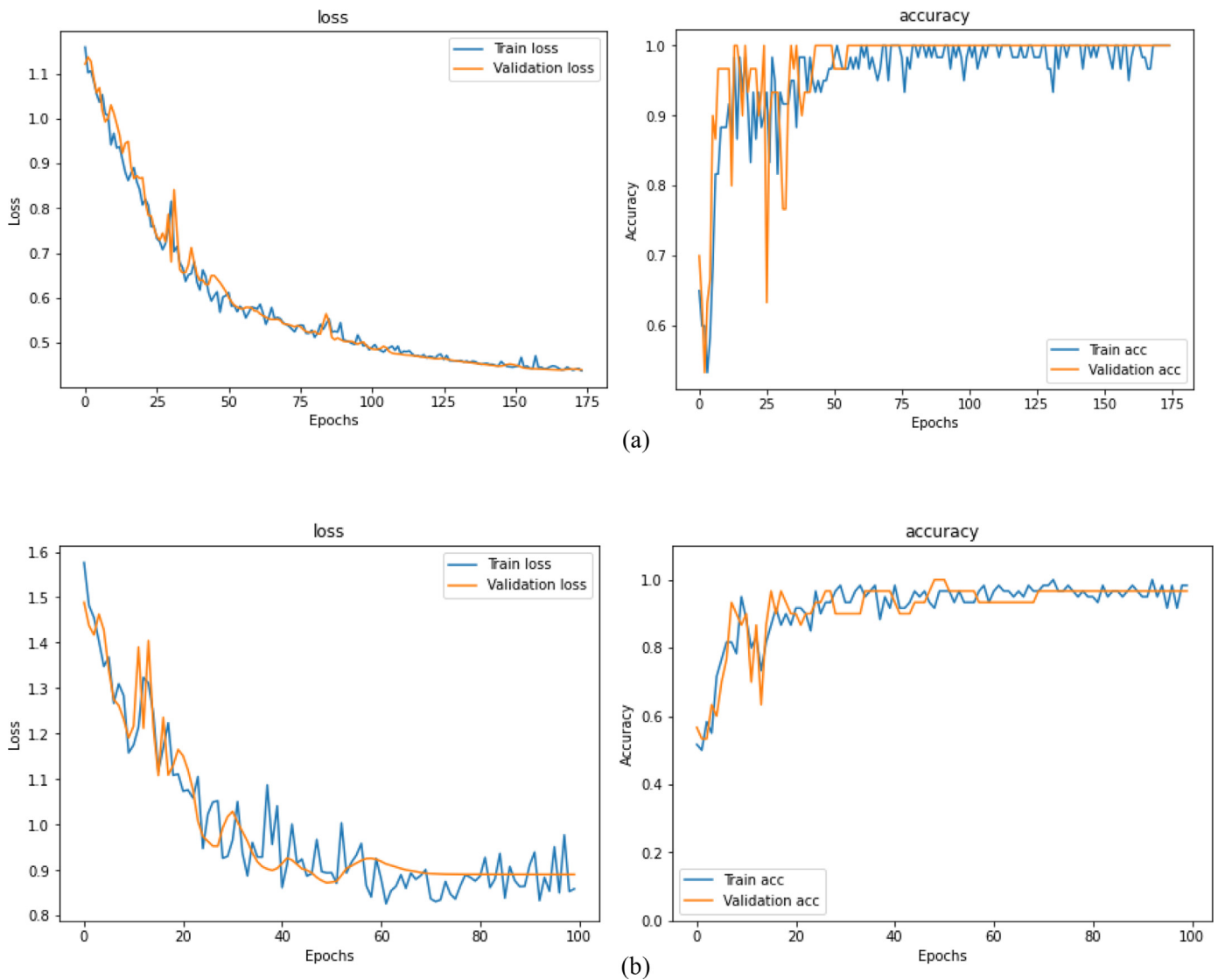


Fig. 6. Training and validation accuracy and loss for 3-way, 10-shot learning settings with (a) contrastive loss (b) cross-entropy loss.

Table 5

Performance results of our model with contrastive loss for various 2-way,  $n$ -shot settings for 2-class (normal, COVID-19) classification.

Model	Accuracy	Precision	Recall	Specificity	F1-score	AUC
MetaCOVID (2-way, 7-shot)	0.940	0.955	0.945	0.958	0.949	0.965
MetaCOVID (2-way, 8-shot)	0.948	0.963	0.955	0.975	0.958	0.975
MetaCOVID (2-way, 9-shot)	0.950	0.975	0.965	0.980	0.969	0.982
MetaCOVID 2-way, 10-shot)	0.965	0.980	0.970	0.984	0.974	0.989

lem) we perform experiments with similar settings in 3-class problem. Table 5 shows the performance results with contrastive loss and various  $n$ -shot settings in classifying healthy and COVID-19 patients. As expected, the model shows better performance in all metrics for diagnosing COVID-19 patients in 2-class scenario.

### 6. Conclusions and future work

This study is one step towards better understanding of the dynamics of COVID-19 pandemic and proposing a state-of-the-art AI based solution for efficient and fast diagnosis system for COVID-19 infections which is the need of the time. The proposed research aims to achieve this through the integration of a meta learning

network model with contrastive loss and pre-trained CNN encoder. Specifically, we use a fine-tuned pre-trained VGG-16 network encoder to capture unbiased feature representations that are robust to overfitting and leverage a Siamese network for final classification of COVID-19 cases. We show that our proposed model with contrastive loss and various  $n$ -shot learning settings offer a highly accurate yet practical solution for automatically diagnosing COVID-19 cases to accelerate line of treatment for patients. Our best model with 10-shot learning setting achieves an accuracy of 95.6% in diagnosing COVID-19 cases with impressive values of sensitivity (96.0%) and specificity (98.0%) which are deemed to be very critical performance estimates for applications in medical settings. Furthermore, our proposed model exhibits comparable or in some

cases better performance than the studied fine-tuned pre-trained CNN models. This is promising due to the fact that our meta learning model is trained only with a limited sample (10) of training examples from each category of CXR images. Simultaneously, it is essential to pinpoint some of the shortcomings of this work which can possibly be tackled in future research. The major drawback is the inadequate interpretability of our model since effective diagnosis requires that results obtained from such interpretability study should be clinically verified by an expert radiologist. As an immediate future work, we plan to extend our work by producing qualitative results with the aid of a model interpretation tool to gain deeper understanding of what our model is learning from the input data during training and validation. We also plan to better tackle COVID-19 diagnosis problem as a multi-modal data fusion problem where various types of clinical data such as patient vitals, location, and population density will be used in addition to image data.

### 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.

### Acknowledgment

The authors are grateful to the [Deanship of Scientific Research at King Saud University](#), Riyadh, Saudi Arabia, for funding this work through the Vice Deanship of Scientific Research Chairs: Chair of Pervasive and Mobile Computing.

### References

- [1] Global COVID-19 Trend, COVID-19 Dashboard, CoronaBoard, URL: <https://coronaboard.com/>. Retrieved July 20, 2020.
- [2] Statement on the second meeting of the International Health Regulations Emergency Committee Regarding the Outbreak of Novel Coronavirus (2019-nCoV), World Health Organization, 2005 30 January 2020. Archived from the original on 31 January 2020. Retrieved July 20, 2020.
- [3] M.A. Shereen, S. Khan, A. Kazmi, N. Bashir, R. Siddique, COVID-19 infection: origin, transmission, and characteristics of human coronaviruses, *J. Adv. Res.* 24 (2020) 91–98 <https://doi.org/10.1016/j.jare.2020.03.005>.
- [4] M.A. Rahman, M.S. Hossain, N.A. Alrajeh, N. Guizani, B5G and explainable deep learning assisted healthcare vertical at the edge: COVID-19 perspective, *IEEE Netw.* 34 (4) (July/August 2020) 98–105.
- [5] S. Wang, B. Kang, J. Ma, et al., A deep learning algorithm using CT images to screen for corona virus disease (COVID-19), *Eur. Radiol.* (2020 Jul 2) 1–11.
- [6] A. Zhao, G. Balakrishnan, F. Durand, et al., Data augmentation using learned transformations for one-shot medical image segmentation, in: Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, 2019, pp. 8543–8553.
- [7] S. Pereira, A. Pinto, V. Alves, C.A. Silva, Brain tumor segmentation using convolutional neural networks in MRI images, *IEEE Trans. Med. Imag.* 35 (5) (2016) 1240–1251.
- [8] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, in: Proc. of the 3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7–9, 2015.
- [9] Y. Liu, G. Liu, Q. Zhang, Deep learning and medical diagnosis, *Lancet* 394 (10210) (2019) 1709–1710 382.
- [10] J. Bullock, A. Luccioni, K.H. Pham, C.C. Lam, M. Luengo-Oroz, Mapping the landscape of Artificial Intelligence applications against COVID-19, *arXiv:2003.11336v2 [cs.CY]* 23 Apr 2020.
- [11] J. Chen, L. Wu, J. Zhang, L. Zhang, D. Gong, Y. Zhao, S. Hu, Y. Wang, X. Hu, B. Zheng, K. Zhang, H. Wu, Z. Dong, Y. Xu, Y. Zhu, X. Chen, L. Yu, H. Yu, Deep learning-based model for detecting 2019 novel coronavirus pneumonia on high-resolution computed tomography: a prospective study, *medRxiv* (2020).
- [12] Z. Zhou, M.M.R. Siddiquee, N. Tajbakhsh, J. Liang, Unet++: a nested U-net architecture for medical image segmentation, in: *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, Springer, 2018, pp. 3–11.
- [13] X. Xu, X. Jiang, C. Ma, P. Du, Deep learning system to screen coronavirus disease 2019 pneumonia, *Engineering* (2020), doi:10.1016/j.eng.2020.04.010.
- [14] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, et al., Going deeper with convolutions, in: Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 1–9.
- [15] B. Gecer, S. Aksoy, E. Mercan, L. Shapiro, D. Weaver, J. Elmore, Detection and classification of cancer in whole slide breast histopathology images using deep convolutional networks, *Pattern Recognit.* 84 (2018).
- [16] M.S. Hossain, G. Muhammad, S.U. Amin, M. Al Sulaiman, Applying Deep Learning for Epilepsy Seizure Detection and Brain Mapping Visualization, *ACM Trans. Multimedia Comput. Commun. Appl.* 15 (1s) (2019) 1–17.
- [17] L. Bi, D.D.F. Feng, M. Fulham, J. Kim, Multi-label classification of multi-modality skin lesion via hyper-connected convolutional neural network, *Pattern Recognit.* 107 (2020) 107502.
- [18] L. Wang and A. Wong, COVID-Net: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest radiography images, Mar 2020.
- [19] 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.* 43 (2020) 635–640, doi:10.1007/s13246-020-00865-4.
- [20] S. Chakraborty, An attempt- Detection of COVID-19 presence from chest X-ray scans using CNN & class activation maps. <https://towardsdatascience.com/detection-of-covid-19-presence-from-chest-x-ray-scans-using-cnn-class-activation-maps-c1ab0d7c294b>, Retrieved July 20, 2020.
- [21] Delft Imaging, CAD4COVID: triage for COVID-19 using artificial intelligence on chest X-rays, <https://www.delft.care/cad4covid/>, Retrieved July 20, 2020.
- [22] A. Abbas, M.M. Abdelsamea, M.M. Gaber, Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network, *Appl. Intell.* (2020), doi:10.1007/s10489-020-01829-7.
- [23] R. Zhang, Y. Zheng, C.C. Poon, D. Shen, J.Y. Lau, Polyp detection during colonoscopy using a regression-based convolutional neural network with a tracker, *Pattern Recognit.* 83 (2018) 209–219.
- [24] H. Bai, R. Wang, Z. Xiong, B. Hsieh, et al., AI augmentation of radiologist performance in distinguishing covid-19 from pneumonia of other etiology on chest CT, *Radiology* (2020) 201491.
- [25] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778.
- [26] G. Muhammad, M.S. Hossain, N. Kumar, EEG-Based Pathology Detection for Home Health Monitoring, *IEEE J. Sel. Areas Commun.* (2020), doi:10.1109/JSAC.2020.3020654.
- [27] X. Yang, et al., Deep relative attributes, *IEEE Trans. Multimedia* 18 (6) (2016) 1832–1842.
- [28] J. Liu, Artificial Intelligence Assisted Radiology Technologies Aid COVID-19 Fight in China, Mar 27, 2020, <https://www.itnonline.com/article/artificial-intelligence-assisted-radiology-technologies-aid-covid-19-fight-china>, Retrieved July 23, 2020.
- [29] C. Ross and R. Robbins, Debate flares over using AI to detect COVID-19 in lung scans, Mar 30 2020, <https://www.statnews.com/2020/03/30/debate-over-artificial-intelligence-to-detect-covid-19-in-lung-scans/>, Retrieved July 23, 2020.
- [30] X. He, X. Yang, S. Zhang, J. Zhao, Y. Zhang, E. Xing, P. Xie, Sample-efficient deep learning for COVID-19 diagnosis based on CT scans, *IEEE Trans. Med. Imag.* (2020).
- [31] X. Chena, L. Yao, T. Zhou, J. Dong, Y. Zhang, Momentum contrastive learning for few-shot COVID-19 diagnosis from chest CT images, *Med. Image Anal.* (2020).
- [32] A. Yassine, M.S. Hossain, COVID-19 networking demand: an auction-based mechanism for automated selection of edge computing services, *IEEE Trans. Netw. Sci. Eng.* (2020), doi:10.1109/TNSE.2020.3026637.
- [33] R. Hadsell, S. Chopra, Y. LeCun, Dimensionality reduction by learning an invariant mapping, in: Proc. of 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), New York, NY, USA, 2006, pp. 1735–1742.
- [34] J. Cohen. COVID-19 image data collection, 2020. <https://github.com/ieee8023/covid-chestxray-dataset>.
- [35] Chest X-Ray Images (Pneumonia), <https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia>.
- [36] C. Szegedy, S. Ioffe, V. Vanhoucke, A. Alemi, Inception-v4, Inception-ResNet and the impact of residual connections on learning, in: Proc. of the Thirty-First AAAI Conference on Artificial Intelligence, 2017, pp. 4278–4284.
- [37] F. Chollet, Xception: deep learning with depthwise separable convolutions, in: Proc. of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 1800–1807.
- [38] M.S. Hossain, G. Muhammad, N. Guizani, Explainable AI and mass surveillance system-based healthcare framework to combat COVID-19 like pandemics, *IEEE Netw.* 34 (4) (2020) 126–132.
- [39] J. Deng, W. Dong, R. Socher, L.J. Li, K. Li, et al., ImageNet: a large-scale hierarchical image database, in: Proc. of 2009 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2009, pp. 248–255.

**Mohammad Shorfuzzaman** is an Associate Professor in the Department of Computer Science, College of Computers and Information Technology (CCIT) at Taif University, Taif, Saudi Arabia. He is a member of BDAAG (Big Data Analytics and Applications) research group in CCIT. His primary research interests include applied artificial intelligence in the areas of computer vision and natural language processing, big data, and cloud computing.

**M. Shamim Hossain** is a Professor at the Department of Software Engineering, College of Computer and Information Sciences, King Saud University, Riyadh, Saudi Arabia. He is also an adjunct professor at the School of Electrical Engineering and Com-

puter Science, University of Ottawa, Canada. He received his Ph.D. in Electrical and Computer Engineering from the University of Ottawa, Canada. His-research interests include cloud networking, smart environment (smart city, smart health), AI, deep learning, edge computing, Internet of Things (IoT), multimedia for health care, and multimedia big data. He has authored and coauthored more than 300 publications including refereed journals, conference papers, books, and book chapters. Recently, he co-edited a book on "Connected Health in Smart Cities", published by Springer. He has served as cochair, general chair, workshop chair, publication chair, and TPC for over 20 IEEE and ACM conferences and workshops. He is the chair of IEEE Special Interest Group on Artificial Intelligence (AI) for Health with IEEE ComSoc eHealth Technical Committee. Currently, he is the Co-Chair of the special session "AI- Enabled technologies for smart health monitoring", to be held with IEEE I2MTC 2021. He was the co-chair of the 3rd IEEE ICME Workshop on Multimedia Services and Tools for smart-health (MUST-SH 2020). He is a recipient of a number of awards, including the Best Conference Paper Award and the 2016 ACM Transactions on Multimedia Computing, Communications and Applications (TOMM) Nicolas D. Georganas Best Paper Award, and the 2019 King Saud University Scien-

tific Excellence Award (Research Quality). He is on the editorial board of the IEEE Transactions on Multimedia, IEEE Multimedia, IEEE Network, IEEE Wireless Communications, IEEE Access, Journal of Network and Computer Applications (Elsevier), International Journal of Multimedia Tools and Applications (Springer), Human-centric Computing and Information Sciences (Springer), Games for Health Journal, and International Journal of Information Technology, Communications and Convergence (Inderscience). He also presently serves as a lead guest editor of IEEE Network, ACM Transactions on Internet Technology, ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM) and Multimedia systems Journal. Previously, he served as a guest editor of IEEE Communications Magazine, IEEE Network, IEEE Transactions on Information Technology in Biomedicine (currently JBHI), IEEE Transactions on Cloud Computing, International Journal of Multimedia Tools and Applications (Springer), Cluster Computing (Springer), Future Generation Computer Systems (Elsevier), Computers and Electrical Engineering (Elsevier), Sensors (MDPI), and International Journal of Distributed Sensor Networks. He is a senior member of both the IEEE, and ACM.