Transfer learning based novel ensemble classifier for COVID-19 detection from chest CT-scans

Nagur Shareef Shaik [*], Teja Krishna Cherukuri

TATA Consultancy Services Ltd., Hyderabad, Telangana, India

ARTICLE INFO

ABSTRACT

Coronavirus Disease 2019 (COVID-19) is a deadly infection that affects the respiratory organs in humans as well as animals. By 2020, this disease turned out to be a pandemic affecting millions of individuals across the globe. Conducting rapid tests for a large number of suspects preventing the spread of the virus has become a challenge. In the recent past, several deep learning based approaches have been developed for automating the process of detecting COVID-19 infection from Lung Computerized Tomography (CT) scan images. However, most of them rely on a single model prediction for the final decision which may or may not be accurate. In this paper, we propose a novel ensemble approach that aggregates the strength of multiple deep neural network architectures before arriving at the final decision. We use various pre-trained models such as VGG16, VGG19, InceptionV3, ResNet50, ResNet50V2, InceptionResNetV2, Xception, and MobileNet and fine-tune them using Lung CT Scan images. All these trained models are further used to create a strong ensemble classifier that makes the final prediction. Our experiments exhibit that the proposed ensemble approach is superior to existing ensemble approaches and set state-of-the-art results for detecting COVID-19 infection from lung CT scan images.

1. Introduction

The novel coronavirus (COVID-19) is the deadliest infectious virus, unleashing an outbreak that shook the whole world and resulted in millions of deaths. The first case of coronavirus was reported in late December 2019 in Wuhan province (China) with symptoms similar to pneumonia [62]. This virus swiftly spread around the globe, initially through person-to-person and later by community-to-community transmissions, becoming a worldwide health hazard [42]. Coronavirus strains are made up of a positive-oriented single-stranded RNA-type, and their ability to mutate quickly makes it impossible to prescribe a standard medication. The International Committee on Taxonomy of Viruses (ICTV) designated the coronavirus as Severe Acute Respiratory Syndrome CoronaVirus 2 (SARS-CoV-2), and the World Health Organization (WHO) proclaimed it a pandemic in March 2020. Due to the rapid spread of the virus and lack of standard medication, this virus resulted in 3.7 million fatalities across the globe by May 2021 [46]. Hence, identifying coronavirus-infected individuals is critical for halting the transmission of the virus and combating the pandemic.

Reverse Transcriptase Polymerase Chain Reaction (RT-PCR), a type of molecular test, is the most extensively used test for detecting the presence of Covid virus in humans. If the RT-PCR test is conducted three weeks after witnessing symptoms, it is 90% accurate in discovering positive cases [15]. However, RT-PCR kits are not widely available, and the test takes at least 6 h to get the final result. Another type of test that may be performed for COVID-19 identification is the lateral flow test (LFT), however it is not as reliable as RT-PCR. Analyzing individual Lung Computerized Tomography (CT) scan images is another strategy that may be used to detect the presence of coronavirus when molecular tests are not available and in time-critical circumstances to obtain immediate results. Besides confirming the existence of the virus, CT-Scan analysis also indicates the percent of infection as well, which aids in offering appropriate treatment. Medical professionals and radiologists acknowledged that CT-Scan analysis detects COVID-19 with 90% accuracy, which is equivalent to RT-PCR [14]. However, large scale manual examination of CT scan images is impractical, and radiologists must rely on Computer Aided Diagnosis (CAD) tools to analyze CT-scan images for COVID-19 detection.

The research community has been rigorously working on developing effective CAD tools for diagnosing COVID infection from medical images such as X-rays and CT scans. Preliminary research focused on the use of Machine Learning based models to identify pathogenic lungs infected

---

with coronavirus [34]. Mesut et al. developed a machine learning based system to classify X-ray images into one of three categories: normal, pneumonia and coronavirus. Fuzzy color-based preprocessing was applied on original X-ray images before they were passed to pre-trained models such as MobileNetV2, SqueezeNet to extract features. Resultant representations were stacked and subjected to Social Mimic optimization to derive discriminant features, which were then used to train a Support Vector Machine (SVM) model for classification [57]. Jawad et al. introduced another machine learning framework in which the original X-ray image dataset was augmented using Generative Adversarial Network (GAN) and then sent to a Convolutional Neural Networks (CNN) for representation learning. The high dimensional representations that resulted were reduced by applying Principal Component Analysis (PCA) and trained using Logistic Regression (LR) model for classification. The authors claimed that using GAN for augmentation and PCA for dimension reduction improved the performance of the classification model [45]. Instead of using different models for feature extraction and classification, a single model was employed for both feature extraction and classification. This is another improvement in the field [25]. Various pre-trained CNN models such as VGG16, NASNet, EfficientNet, DenseNet121, and Xception were adapted and fine-tuned using the X-ray image datasets with the aim of detecting coronavirus [38].

Most of the above discussed works mainly focused on X-ray image analysis for Covid prediction. Significant research has been done for identification of Covid-19 infection based on chest CT scan images. Matteo et al. suggested a light CNN model based on SqueezeNet for COVID-19 detection from chest CT scan images [43]. Another variant of Convolutional Neural Network architecture was designed and implemented to handle CT scan images for COVID-19 identification [16]. Another unique model, weighted filter based CNN architecture, was proposed with a minimum number of layers. Weighted filters enabled the CNN model to prioritize a certain set of features, resulting in increased accuracy [48]. Similarly, numerous CNN architectures have been proposed and developed in the literature for COVID detection from medical images [1]. Most of these works are limited in terms of feature optimization as they used CNN models for representation learning. Addressing this, a bi-stage feature selection approach was proposed for selecting optimum features acquired from CNNs trained on chest CT scan images [51]. However, CNN needs large-scale labeled medical images to develop a robust CAD system. Addressing this issue, various transfer learning and data augmentation based methodologies have been proposed in the literature [30]. However, most of them suffer from uni-lateral decisions, which means that all of those approaches were confined to decisions from a single model, without any mechanism to cross-check the correctness of those models.

In this article, we present a novel ensemble strategy for making final predictions that combines the predictive capabilities of multiple state-of-the-art transfer learning-based pre-trained models. Initially, Lung CT scan images are processed and passed to one of the deep pre-trained models such as VGG16, VGG19, InceptionV3, ResNet50, ResNet50V2, InceptionResNetV2, Xception, and MobileNet. The convolution base of the model is frozen and the features obtained from them are fine-tuned on a three layered deep neural network with an induced dropout rate of 0.3 after each layer. The Softmax probabilities obtained from these models serve as input to the proposed ensemble classifier, which is a composite neural network architecture that is jointly trained using multi-modal prediction probabilities. Our extensive experimental studies demonstrate that the proposed novel ensemble classifier is robust and outperforms existing ensemble approaches in the literature for the COVID-19 detection task. Summary of our major contributions in this work include:

● Preprocessing of Lung CT scan images before passing them to pre-trained models for feature extraction.

● Fine-tuning eight different pre-trained models such as VGG16, VGG19, ResNet50, ResNet50V2, InceptionV3, InceptionResNetV2, Xception, and MobileNet with frozen convolution base and trainable deep neural network head for classification.
● Designing a Composite Neural Network architecture which ensembles multi-modal Softmax probabilities.
● Conducting experiments on SARS-CoV-2 and COVID-CT chest CT scan image datasets and presenting a comparative study with literature.

This article is organized as follows: Section 2 presents a thorough description of the literature related to the theme of the work under consideration. Section 3 outlines the work flow of the proposed model along with a comprehensive description about each module of the proposed model developed for COVID-19 detection from CT scan images. The datasets used for experimental studies, as well as the experimental setup and performance measurements, are described in detail under section 4. This section also includes experimental results, discussions about the outcomes of the experiments, and a comparative study. Section 5 summarizes this study, as well as future insights.

2. Background study

Recent advances in the Computer Vision Field (CVF) have grabbed the attention of the research community [10,11,58] and they started exploring ways to apply these models for medical image analysis applications such as tumor classification [8,9,53], retinopathy identification [6,7,52] and even for COVID-19 diagnosis [34]. This section of the manuscript summarizes recent works proposed in the literature for COVID-19 detection from X-ray and CT scan images.

2.1. Convolutional neural network based methods

Convolutional Neural Networks (CNNs) are sparse Neural Network architectures that have been in wide usage for solving a variety of complex visual recognition tasks. Unlike traditional machine learning approaches, these deep models do not rely on hand-crafted features because the architecture itself is designed with feature extraction capabilities [35]. Convolutions, followed by Pooling operations, are used for feature extraction and sub-sampling, while the final dense layers present in the classification head [21] predict the class label. These CNN models are well-known for generating cutting-edge outcomes in a wide range of hard situations, including object identification and semantic segmentation [19].

Following the success of Convolutional Neural Network models, the research community focused on developing numerous CNN architectures for detecting COVID-19 from chest X-ray and chest CT-scan images. A Deep Convolutional Neural Network architecture is designed and trained using chest X-ray images for automating the process of COVID-19 detection [37]. Another novel CNN architecture, ReCoNet, is designed with residual connections and multi-level processing for the same purpose and has been tested on the same dataset [2]. Similarly, another CNN-based model is designed and trained for reliable COVID-19 identification utilizing multiple chest imaging such as CT scans, X-rays, and ultrasound scan images [20]. Ardakani et al. conducted a thorough examination of the performance of 10 different convolutional neural network architectures for covid prediction using CT scan images [4]. Similarly, different CNN architectures have been designed by modifying several architectural components, such as the number of layers, type of kernels, dimensions of kernels, and so on. However, the amount of labeled radiology images used for training limits the performance and generalization of those models [59]. In order to address this issue, several transfer learning-based approaches for COVID-19 identification have been developed in the literature [29,41].

2.2. Transfer learning based methods

Transfer Learning is a prominent Deep Learning approach in which the knowledge gained while training that model using large scale datasets can be transferred to another model specifically designed to solve a similar or related problem [17,59]. Transfer learning strategies often use pre-trained models, which are deep neural network models trained on huge labeled data and whose weights are readily available for usage [47]. VGGNet, InceptionV3, ResNet, Xception, MobileNet are popular state-of-the-art pre-trained models trained on the ImageNet dataset. In this section, we will explore significant transfer learning methodologies established in the literature for Covid-19 prediction. Eduardo Soares et al. introduced a CT scan dataset for COVID detection and employed various pre-trained models like ResNet, GoogLeNet, VGG16, and AlexNet for COVID infection detection [3]. A hybrid model combining CNN and Bi-LSTM was designed for COVID infection detection from CT scans [5]. Transfer learning based CAD system is built using chest radio graph images [36]. In contrast to traditional transfer learning, a deep transfer learning based approach was proposed to identify corona infection from chest X-ray images [31]. In similar lines, various transfer learning approaches have been proposed in the literature. Rahaman et al. developed the most competitive and different transfer learning approach for the COVID-19 detection task [44].

Most of the strategies discussed above relied on pre-trained models for feature extraction or fine-tuning. Few of them offered comparative studies among popular models fine-tuned on chest radio graphs, but they did not focus on fusing their predictive power, instead relayed on predictions from a single model. Few of them spotlighted the use of ensemble approaches to build a strong classifier from weak candidate models. Ebenezer et al. designed stacking-based ensemble by binding numerous pre-trained model predictions with a single Softmax neuron [49]. Both of these works employed final prediction labels of candidate models to train the final single neuron ensemble classifier. In contrast to these models, we suggest a unique ensemble strategy that uses prediction probabilities, that represent the confidence scores with respect to the classes, and passes them to a composite deep neural network that aggregates candidate models.

3. Proposed approach

The main aim of this work is to introduce a new ensemble approach that can learn from candidate CNN model predictions to produce more accurate predictions of Covid-19 based on chest CT scan images. The proposed approach consists of two modules, design of candidate model architecture and novel ensemble classifier. The technical features of these modules are covered in the sub-sections below.

3.1. Architecture of the candidate models

A candidate model is a neural network model comprising a pre-processing module, a pre-trained convolutional base, and a classification head with regularization. Initially, chest CT scan images are passed through a preprocessing module for feature-wise normalization. These Normalized images are further reshaped according to the inputs accepted by the pre-trained models. These processed images are then fed to a frozen Convolutional base with ImageNet weights, this stage extracts semantic feature maps. Instead of directly passing the high dimensional spatial features to the classification head, these feature maps are compressed using the Global Average Pooling (GAP) layer and are eventually passed through the proposed classification head. The classification head is designed with a combination of dense and dropout layers to prevent the model from overfitting. Final dense layer of the model is connected to a Softmax layer with the number of neurons corresponding to the number of classes, in our case two neurons are used to represent Covid positive and negative cases. This Softmax layer generates class-specific probabilities, which are further fed to the proposed ensemble classifier. Fig. 1 depicts the architecture of the candidate model proposed for Coronavirus prediction. Each block of the candidate model that is used as part of the proposed model is explored in detail in the sub-sections that follow.

3.1.1. Preprocessing module

The main task of this module is to perform feature-wise normalization on the input images. The images will be coerced into a zero-centered mean and unit standard deviation distribution using this module. This is achieved by computing the mean and variance of each image pixel value. The normalized images are further reshaped to the size accepted by a pre-trained convolutional base, which extracts feature representations. For instance, VGGNet, ResNet50, ResNet50V2, MobileNet accepts ($224 \times 224 \times 3$) shaped images and InceptionV3, Xception, InceptionResNetV2 accepts ($299 \times 299 \times 3$) shaped images as input. Hence, Normalized images are reshaped accordingly before passing to the ConvBase module.

3.1.2. Pre-trained convolution base

This module is the heart of the candidate model. The main objective of this module is to extract semantic spatial representations from chest CT scans. We use eight different state-of-the-art pre-trained models like VGGNet with 16 and 19 layers [54], ResNet50, ResNet50V2 [22], InceptionV3 [56], Xception [13], MobileNet [24] and InceptionResNetV2 [55] to form Convolutional Base with frozen ImageNet weights. Table 1 provides architecture details of pre-trained models such as expected input dimension, number of layers, number of parameters, number of frozen layers in the convolution base, and the number of spatial feature maps obtained from frozen convolution base. The 2D spatial representations obtained from ConvBase module are high dimensional and are squeezed by applying the global average pooling operation. Final squeezed representations serve as input to the classification head.

3.1.3. Regularized classification head

The proposed classification head module consists of three ReLU activated dense layers, with 512, 512, and 256 neurons respectively. Each dense layer is supported by a dropout layer at a rate of 0.3 to avoid overfitting. In addition, weights of fully connected layers are L_1 regularized and biases are L_1L_2 regularized to further avoid overfitting of the model during the training process. All three modules are stacked and trained in an end-to-end fashion to generate predictions. Induced dropout and parameter regularization prevent the model from overfitting and allow it to produce meaningful predictions with reasonable accuracy. However, the performance of candidate models varies depending on the convolutional base used for feature extraction. In order to create a strong COVID detection system, we design a composite neural network which ensembles all candidate models by leveraging candidate model predictions.
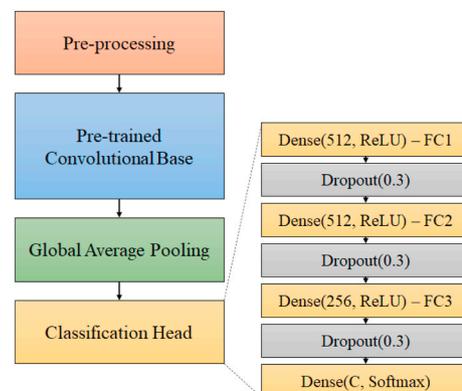
Fig. 1. Architecture of candidate model.

Table 1
Detailed Characteristics of Pre-trained models used as Convolution Base for Candidate Models.

Model	Default Input Size	# Layers	# Parameters	# Frozen Layers	# Features
VGG16	(224 × 224 × 3)	16	138,357,544	13	4096
VGG19	(224 × 224 × 3)	19	143,667,240	16	4096
ResNet50	(224 × 224 × 3)	50	25,636,712	48	2048
ResNet50V2	(224 × 224 × 3)	50	25,613,800	48	2048
InceptionV3	(299 × 299 × 3)	48	23,851,784	46	2048
Xception	(299 × 299 × 3)	36	22,910,480	34	2048
MobileNet	(224 × 224 × 3)	28	4,253,684	26	1024
InceptionResNetV2	(299 × 299 × 3)	164	55,873,736	162	1536

3.2. Novel ensemble classifier

Ensemble is the process of creating a strong model from multiple weak candidate models that have been trained to perform the same task. Numerous such ensemble techniques have been proposed in the literature for classification and regression tasks [18]. Max Voting [12], Stacking [27], Bagging [60] and Boosting are popular ensemble techniques that have demonstrated effective performance for a variety of tasks, including COVID-19 infection detection [28]. The downside of these models is that they rely solely on the final prediction label and ignore the confidence factor of the label. To avoid such problems, we propose a novel ensemble approach that focuses on confidence labels predicted by multiple candidate models.

Initially, prediction probabilities from eight candidate models are passed to a composite neural network with a fusion layer, which is then followed by two ReLU activated fully connected layers with 256 neurons each. A dropout rate of 0.2 is applied after each fully connected layer, and all of the parameters of this composite network are regularized to prevent the model from overfitting. The Model loss is computed by taking into account the cross-entropy loss between the predictions and the ground truths, and the model parameters are optimized using ADAM optimizer. Fig. 2 depicts the overall picture of the proposed novel ensemble approach for COVID-19 infection detection using chest CT scan images.

Let X be an input chest CT Scan image of dimension ($H \times W \times D$) passed to candidate models, and (P_0, P_1, .., P_c) be the probabilities obtained from Softmax output layer of candidate models and represents how likely the given input X belongs to a class C. For eight candidate models, we get eight probability vectors as output and these output probabilities serve as input to a composite neural network. Instead of concatenating them to form a single vector, we treat each probability vector individually and pass them to a composite model which has eight dedicated input layers. This approach enables the model to learn candidate-specific probability while avoiding overlap with other candidate model probabilities. The fusion layer, which is placed after the input layer, allows the model to select the optimal feature vector from the input. The fully connected layers placed after the fusion layer learn the non-linear discriminant information required to accurately detect COVID-19 infection.

In general, ensemble models use the most likely class label as input. This may mislead the final ensemble model deviating the focus on the maximum win label, resulting in greater misclassifications. The significance of the proposed model is that it attempts to derive optimal decision by leveraging the class probabilities obtained from several state-of-the-art candidate CNN models. Instead of just relying on majority votes given by candidate models, it reduces misclassifications by leveraging both maximum votes and label confidence with respect to the actual label during training. End-to-end training of multiple inputs targeting a single goal helps the model to achieve superior performance over other ensemble strategies.

4. Experimental discussions

This section summarizes the experiments conducted to validate the proposed novel ensemble approach for detecting COVID-19 infection from chest CT scan images. A comprehensive description of the datasets utilized for the experimental studies is given at the beginning of this section, followed by the experimental setup and evaluation measures used to assess the efficiency of the proposed ensemble approach. Finally,
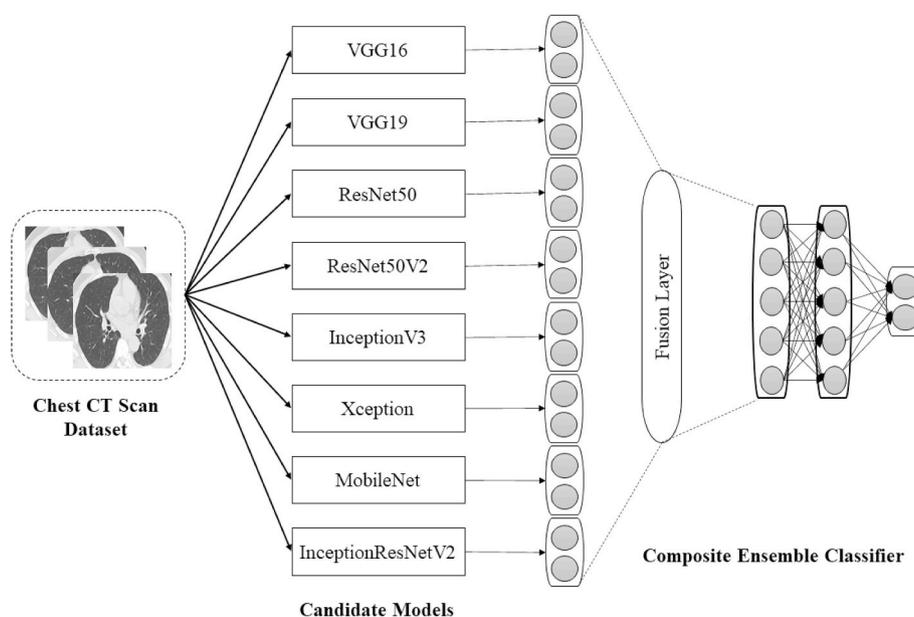
Fig. 2. Architecture of proposed Novel Ensemble Classifier; Each fully connected layer post fusion layer consists of 256 neurons with ReLu activation and parameter regularization; Dropout with rate 0.2 is applied after every fully connected layer.

we present a comprehensive analysis of the experimental results on two benchmark datasets, along with comparative studies.

4.1. Summary of the CT-Scan datasets

The effectiveness of the proposed ensemble classifier is validated using two benchmark CT scan datasets, SARS-CoV-2 [3], and COVID-CT [61]. The subsections that follow provide details about each dataset, such as the source of the data, and number of images available for each category (COVID-Positive and COVID-Negative).

4.1.1. SARS-CoV-2

The SARS-CoV-2 dataset consists of 2482 chest CT scan images, 1252 of which are from patients infected with SARS-CoV-2, and 1230 from patients who are not infected with SARS-CoV-2 but have other pulmonary diseases. These CT scans are available in PNG format and have spatial dimensions ranging from 104×119 to 416×512. The samples of this dataset were collected from Sao Paulo hospitals, Brazil, and it includes 120 individuals, both male and female.

4.1.2. COVID-CT

The COVID-CT dataset is a small-scale chest CT scan image dataset. This dataset consists of 746 images, 349 of which belong to COVID-Positive and 397 of which belong to COVID-Negative category. All COVID-Positive CT Scan images were collected from medRxiv and bioRxiv. These CT scans have spatial dimensions ranging from 124×153 to 1485×1853 and are available in PNG format. Fig. 3 represents chest CT scan images with and without SARS-CoV-2 infection.

4.2. Setup for experiments

All the experiments are conducted in the Google Co-laboratory with the Keras and TensorFlow libraries. Since no standard test splits are provided for both datasets, we adopted hold-out set validation strategy as per [3,23,26,32,40,50] and used 80% of the data for training and 20% for validation. As most of the recent works followed an image-wise split on these datasets for validation, we too followed the same strategy to make a standard validation split in order to establish a fair comparison with them.

We empirically determined the values of various hyper-parameters by observing the model performance by setting parameters within certain ranges. For the proposed model, we investigated learning rate ranging from 0.005 to 0.05, dropout rates ranging from 0.1 to 0.5, and weight & bias regularization rates ranging from 0.001 to 0.1. We have determined the parameters after conducting extensive experiments with

these values, and fixed the hyper-parameters that resulted in highest performance. The details of each hyper-parameter is reported in Table 2. In all the experiments, the ensemble model is allowed to run for 100 epochs to achieve consistent performance. Despite the fact that 100 epochs is a fairly short number of epochs for training deep networks, we observed that the model quickly converged, and there was no substantial improvement in loss after 100 epochs. We have set epochs to 500 for fine-tuning candidate models. Other hyper-parameters, such as loss and optimizer, are selected based on the task at hand and type of training data available. Table 2 provides the details of fixed hyper-parameters used while performing experiments on our proposed ensemble approach and relevant studies. We used the check pointing strategy throughout the training phase to save the best model with the least amount of loss. The results reported for all the experiments refer to the measures obtained for the test data alone.

4.3. Model evaluation measures

Different classification metrics such as Accuracy, Precision, Recall, F_1 Score and AUC are used to determine the efficacy of the proposed approach for the Coronavirus detection task. Accuracy is the ratio of the number of correctly estimated CT scan images to the actual number of CT scan images being examined. Precision reveals what percentage of the diseased images are actually diseased, while recall reveals what percentage of the actual diseased are predicted as diseased. F_1 Score is the harmonic mean of the Precision and recall. Mathematical

Table 2
Hyper-parameter values set for Candidate and Ensembling models.

Hyper Parameter	Value
Candidate Model - Batch Size	32
Ensembling Model - Batch Size	32
Candidate Model - Epochs	500
Ensembling Model - Epochs	100
Candidate Model - Optimizer	ADAM
Ensembling Model - Optimizer	ADAM
Candidate Model - Learning Rate	0.003
Ensembling Model - Learning Rate	0.001
Dropout Rate for Candidate Model	0.3
Dropout Rate for Ensembling Model	0.2
Weight Regularization	L1(0.01)
Bias Regularization	L1_L2(0.01)
Hidden layer activation	ReLU
Output layer activation	Softmax
Candidate Model - Loss	Cross-entropy
Ensembling Model - Loss	Cross-entropy

Normal Chest CT Scan image samples (COVID-19 Negative)

Chest CT Scan image samples with SATS-CoV-2 Infection (COVID-19 Positive)
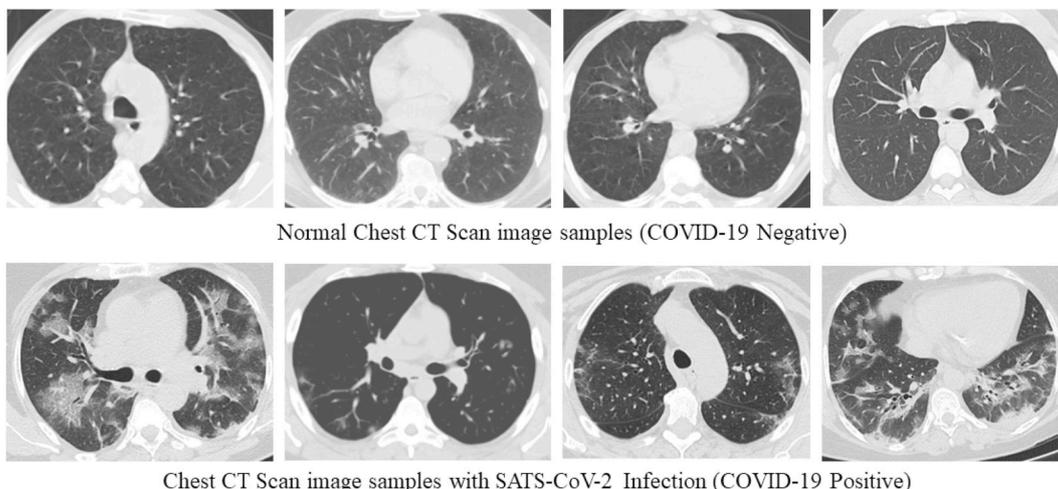
Fig. 3. Chest CT Scan Sample images with and without COVID-19 Infection.

expressions for computing each of these metrics are as follows:

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} \tag{1}$$

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

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

$$F_1Score = \frac{2 * Precision * Recall}{Precision + Recall} \tag{4}$$

In equations (1)–(4), TP and FP correspond to the number of true positives and number of false positives respectively. While TN and FN correspond to the number of true negatives and false negatives.

4.4. Result analysis on SARS-CoV-2 dataset

The SARS-CoV-2 dataset has been considered for our preliminary experiments. Deep pre-trained networks including VGG16, VGG19, ResNet50, ResNet50V2, InceptionV3, Xception, MobileNet, Inception-ResNetV2 are used as Convolutional Base for candidate models to learn feature representations from the CT scan images. Regularized classification head provides initial predictions which are later used as inputs to the proposed ensemble model. Table 3 represents the results of several baseline candidate models. Modified ResNet with 50 layers, MobileNet, and VGG with 16 layers have demonstrated their architectural stability for the COVID detection task by producing accurate predictions when compared to other models.

InceptionResNetV2 (IRV2) achieved the lowest classification accuracy of 91.35%, while ResNet50V2 achieved the highest classification accuracy of 97.79%. The performance of the InceptionV3 and MobileNet models is similar to that of the IRV2 and ResNet50V2 models. ResNet50V2 outperforms other models not only in terms of accuracy, but also in terms of Precision, Recall F_1-Score, and AUC.

We compare the proposed ensemble approach with the top five models (5-clf) and all models (8-clf) to check the negative impact of the weak models on the performance of strong models in ensemble approaches. This comparison is included in Table 4, along with comparisons of existing popular ensemble approaches such as Max Voting, Random Forest bagging classifier, Gradient Boosting classifier. We can observe that the proposed ensemble approach with 5-clf and 8-clf outperforms the existing ensemble approaches for the task of COVID detection on the SARS-CoV-2 dataset with an accuracy of 98.99%. This is the best accuracy so far on this dataset, to the best of our knowledge. The accuracy produced by both 5-clf and 8-clf is the same, but they differ in terms of other performance indicators considered here. The precision of the 5-clf model is 99.02%, which is the highest among all models, while the recall of the 8-clf model is 99%, which is also the highest among all models.

Table 4
Comparing the performance of various ensembling models with proposed approach on SARS-CoV-2 dataset images.

Ensembling Approach	Accuracy	Precision	Recall	F1-Score	AUC
Max Voting	98.39	98.37	98.40	98.39	98.40
Bagging (Random Forest)	96.18	96.26	96.32	96.18	98.44
Boosting (Gradient Boosting)	98.19	98.17	98.25	98.19	98.25
Proposed Ensembling (5-clf)	98.99	99.02	98.97	98.99	98.97
Proposed Ensembling (8-clf)	98.99	98.98	99.00	98.99	99.00

4.5. Result analysis on COVID-CT dataset

We extend our experimental studies with the COVID-CT dataset to check the effectiveness of the proposed ensemble approach with small scale datasets. Table 5 provides the results of various baseline candidate models trained on COVID-CT dataset. With this dataset, VGG with 16 layers achieved the highest accuracy of 92%, while all other models were confined to below 90% accuracy. In Comparison, VGG19, MobileNet and ResNet50V2 models result in better accuracy than InceptionV3, ResNet, Xception, and IRV2 models. From these results, we may conclude that the VGGNet architecture is better suitable for a small-scale dataset with lower mis-classification rates for the COVID-19 detection task.

Table 6 compares the proposed ensemble approach with existing ensemble techniques. As discussed above, to validate the impact of weak models in making a final decision, we perform ensemble of the top five candidate models, VGG16, VGG19, ResNet50V2, MobileNet, and Xception networks (5-clf) and all models (8-clf). It can be observed that both 5-clf and 8-clf based ensemble models are comparable in terms of accuracy, however they differ in other scores such as precision, recall, and F_1-Score. The 5-clf model achieves a precision of 93.60%, while the 8-clf model achieves a recall of 93.54%, which are highest among all the models.

4.6. Comparison with state-of-the-art methods

In this section, we present a comparative study of the proposed ensemble approach with several state-of-the-art models developed for detecting COVID infection from chest CT scan images. We compare the proposed novel ensemble classification model with a set of models trained on deep feature representations of chest CT scan images from both SARS-CoV-2 and COVID-CT datasets.

Table 7 compares the proposed model performance with the existing models in terms of accuracy, precision, recall and F1 Score. Deep architectures such as xDNN [3], DenseNet201 [26] and Modified VGG-19 [40] achieve reasonably good performance in terms of various measures on SARS-CoV-2 dataset. In terms of accuracy, xDNN surpasses the other two models with a 1.10%–2.30% improvement. However, our proposed ensemble approach exhibits superior performance and outperforms

Table 3
Comparing the performance of various pre-trained models fine-tuned on SARS-CoV-2 dataset images.

Model	Accuracy	Precision	Recall	F1-Score	AUC
VGG16	96.18	96.16	96.2	96.17	96.2
VGG19	94.77	94.95	94.66	94.75	94.66
ResNet50	92.15	92.13	92.18	92.14	92.18
ResNet50V2	97.79	97.77	97.84	97.78	97.84
InceptionV3	91.55	91.62	91.67	91.55	91.67
Xception	94.77	94.75	94.83	94.77	94.83
MobileNet	97.38	97.41	97.35	97.38	97.35
InceptionResNetV2	91.35	91.33	91.39	91.34	91.39

Table 5
Comparing the performance of various pre-trained models fine-tuned on COVID-CT dataset images.

Model	Accuracy	Precision	Recall	F1-Score	AUC
VGG16	92.00	91.91	91.91	91.91	91.91
VGG19	88.67	88.58	88.46	88.52	88.46
ResNet50	84.67	84.52	84.42	84.47	84.42
ResNet50V2	88.00	87.96	87.72	87.82	87.72
InceptionV3	83.33	83.27	83.65	83.27	83.65
Xception	85.33	86.51	84.30	84.80	84.30
MobileNet	88.67	88.5	88.61	88.55	88.61
InceptionResNetV2	82.00	81.90	81.58	81.71	81.58

Table 6

Comparing the performance of various ensembling models with proposed approach on COVID-CT dataset images.

Ensembling Approach	Accuracy	Precision	Recall	F1-Score	AUC
Max Voting	91.33	91.29	91.16	91.22	91.16
Bagging (Random Forest)	87.33	87.84	86.83	87.08	86.83
Boosting (Gradient Boosting)	84.67	88.34	82.98	83.67	82.98
Proposed Ensembling (5-clf)	93.33	93.60	92.97	93.21	92.97
Proposed Ensembling (8-clf)	93.33	93.17	93.54	93.29	93.54

Table 7

Comparing the performance of various recently published deep learning based models with proposed ensembling approach on SARS-CoV-2, COVID-CT datasets.

Dataset	Model	Accuracy	Precision	Recall	F1-Score
SARS-CoV-2	xDNN [3]	97.30	99.10	95.50	97.30
	DenseNet201 [26]	96.20	–	–	–
	Modified VGG-19 [40]	95.00	95.30	94.00	94.30
	Convolutional SVM [39]	96.00	–	–	–
	Proposed (5-clf)	98.99	99.02	98.97	98.99
	Proposed (8-clf)	98.99	98.98	99.00	98.99
COVID-CT	Decision Function [32]	88.30	–	–	86.70
	ResNet101 [50]	80.30	78.20	85.70	81.80
	DenseNet169 (Self-trans Approach) [23]	86.00	–	–	85.00
	Capsule Network [33]	–	84.00	–	–
	SqueezNet based CNN [43]	85.00	85.00	87.00	86.00
	Transfer Learning Ensemble [49]	86.00	–	89.00	85.00
	Proposed (5-clf)	93.33	93.60	92.97	93.21
	Proposed (8-clf)	93.33	93.17	93.54	93.29

xDNN, hybrid Convolutional SVM [39] and the rest of the baseline models on SARS-CoV-2 dataset. To the best of our knowledge, 98.99% is the state-of-the-art accuracy so far on the SARS-CoV-2 dataset.

Similar trends have been observed in the case of COVID-CT dataset as well. DesnseNet169 [23], ResNet101 [50] and SqueezNet based CNN [43] achieve acceptable results on COVID-CT dataset, whereas ensemble of such deeper architectures (Decision Function [32] and Transfer Learning Ensemble [49]) outperforms individual models. On the COVID-CT dataset, our proposed ensemble approach outperforms existing ensembles models in the literature and sets state-of-the-art scores by achieving an accuracy of 93.33%.

From this comparison study, we claim that the proposed novel ensemble approach minimizes misclassifications and provides more generalized predictions for COVID-19 detection on providing processed chest CT scan images as input candidate models whose predictions are in turn used by the proposed ensemble model.

Using processed chest CT scan images as input to the candidate models, we claim that the proposed new ensemble technique minimizes mis-classifications and gives more generic predictions for COVID-19 detection. These predictions are then utilized by the suggested ensemble model and further improves accuracy.

5. Conclusion and future work

The main objective of this work is to introduce an effective ensemble approach for detecting SARS-CoV-2 infection from chest CT scan images. The proposed approach is unique in terms of candidate model design used for initial predictions and composite ensemble classifier design for ensemble candidate models. Deep pre-trained convolutional bases of VGGNet, ResNet, MobileNet, InceptionV3, Xception, and IRV2 allow

candidate models to learn discriminant features from CT scan images. Regularized classification head prevents candidate models from over-learning and makes generalized predictions. The Proposed composite ensemble classifier receives candidate prediction probabilities as an input and creates a strong discriminator for final prediction. Our experimental studies on two benchmark datasets, SARS-CoV-2 and COVID-CT, validates the effectiveness of the proposed ensemble approach for COVID-19 detection. The results from both datasets indicate that, in comparison to existing models and conventional ensemble techniques, the suggested ensemble approach is substantially efficient in terms of various performance measures.

This work might be extended in the future to allow for the detection of a variety of lung infections from CT scan images. Furthermore, we may also try to design an attention-based ensemble approach that prioritizes certain candidate models based on their individual performance for the task considered.

CRediT authorship contribution statement

Nagur Shareef Shaik: Conceptualization, Methodology, Software, Formal analysis, Writing - Original Draft. **Teja Krishna Cherukuri:** Software, Investigation, Validation, Data Curation, Visualization.

References

[1] A. Abbas, M.M. Abdelsamea, M.M. Gaber, Classification of covid-19 in chest x-ray images using detract deep convolutional neural network, Appl. Intell. 51 (2021) 854–864.

[2] S. Ahmed, M.H. Yap, M. Tan, M.K. Hasan, Reconet: Multi-Level Preprocessing of Chest X-Rays for Covid-19 Detection Using Convolutional Neural Networks, medRxiv, 2020.

[3] P. Angelov, E. Almeida Soares, Sars-cov-2 Ct-Scan Dataset: A Large Dataset of Real Patients Ct Scans for Sars-Cov-2 Identification, medRxiv, 2020.

[4] A.A. Ardakani, A.R. Kanafi, U.R. Acharya, N. Khadem, A. Mohammadi, Application of deep learning technique to manage covid-19 in routine clinical practice using ct images: results of 10 convolutional neural networks, Comput. Biol. Med. 121 (2020), 103795.

[5] M.F. Aslan, M.F. Unlersen, K. Sabanci, A. Durdu, Cnn-based transfer learning–bilstm network: a novel approach for covid-19 infection detection, Appl. Soft Comput. 98 (2021), 106912.

[6] J.D. Bodapati, N.S. Shaik, V. Naralasetti, Composite deep neural network with gated-attention mechanism for diabetic retinopathy severity classification, J. Ambient Intell. Humanized Comput. (2021) 1–15.

[7] J.D. Bodapati, N.S. Shaik, V. Naralasetti, Deep convolution feature aggregation: an application to diabetic retinopathy severity level prediction, Signal, Image Video Process. 1–8 (2021).

[8] J.D. Bodapati, N.S. Shaik, V. Naralasetti, N.B. Mundukur, Joint training of two-channel deep neural network for brain tumor classification, Signal, Image Video Process. 15 (2021) 753–760.

[9] J.D. Bodapati, S.N. Shareef, V. Naralasetti, N.B. Mundukur, Msenet: multi-modal squeeze-and-excitation network for brain tumor severity prediction, Int. J. Pattern Recogn. Artif. Intell. (2021), 2157005.

[10] J.D. Bodapati, N. Veeranjaneyulu, Abnormal network traffic detection using support vector data description, in: Proceedings of the 5th International Conference on Frontiers in Intelligent Computing: Theory and Applications, Springer, 2017, pp. 497–506.

[11] J.D. Bodapati, N. Veeranjaneyulu, S. Shaik, Sentiment analysis from movie reviews using lstms, Ingénierie Des. Systèmes Inf. 24 (2019).

[12] T.B. Chandra, K. Verma, B.K. Singh, D. Jain, S.S. Netam, Coronavirus disease (covid-19) detection in chest x-ray images using majority voting based classifier ensemble, Expert Syst. Appl. 165 (2021) 113909.

[13] F. Chollet, Xception: deep learning with depth wise separable convolutions, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 1251–1258.

[14] A. Dangis, C. Gieraerts, Y.D. Bruecker, L. Janssen, H. Valgaeren, D. Obbels, M. Gillis, M.V. Ranst, J. Frans, A. Demeyere, et al., Accuracy and reproducibility of low-dose submillisievert chest ct for the diagnosis of covid-19, Radiology: Cardiothoracic Imaging 2 (2020), e200196.

[15] J. Dinnes, J.J. Deeks, S. Berhane, M. Taylor, A. Adriano, C. Davenport, S. Dittrich, D. Emperador, Y. Takwoingi, J. Cunningham, et al., Rapid, point-of-care antigen and molecular-based tests for diagnosis of sars-cov-2 infection, Cochrane Database Syst. Rev. 3 (2021).

[16] C. Do, L. Vu, An approach for recognizing covid-19 cases using convolutional neural networks applied to ct scan images, in: Applications of Digital Image Processing XLIII, International Society for Optics and Photonics, 2020, p. 1151034.

[17] V. Dondeti, J.D. Bodapati, S.N. Shareef, N. Veeranjaneyulu, Deep convolution features in non-linear embedding space for fundus image classification, Rev. d'Intelligence Artif. 34 (2020) 307–313.

[18] X. Dong, Z. Yu, W. Cao, Y. Shi, Q. Ma, A survey on ensemble learning, Front. Comput. Sci. 14 (2020) 241–258.

[19] S. Gidaris, N. Komodakis, Object detection via a multi-region and semantic segmentation-aware cnn model, in: Proceedings of the IEEE International Conference on Computer Vision, 2015, pp. 1134–1142.

[20] G. Gilanie, U.I. Bajwa, M.M. Waraich, M. Asghar, R. Kousar, A. Kashif, R.S. Aslam, M.M. Qasim, H. Rafique, Coronavirus (covid-19) detection from chest radiology images using convolutional neural networks, Biomed. Signal Process Control 66 (2021) 102490.

[21] J. Gu, Z. Wang, J. Kuen, L. Ma, A. Shahroudy, B. Shuai, T. Liu, X. Wang, G. Wang, J. Cai, et al., Recent advances in convolutional neural networks, Pattern Recogn. 77 (2018) 354–377.

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

[23] 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, medrxiv, 2020.

[24] A.G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, H. Adam, Mobilenets: Efficient Convolutional Neural Networks for Mobile Vision Applications, 2017 arXiv preprint arXiv:1704.04861.

[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. 51 (2021) 1690–1700.

[26] A. Jaiswal, N. Gianchandani, D. Singh, V. Kumar, M. Kaur, Classification of the covid-19 infected patients using densenet201 based deep transfer learning, J. Biomol. Struct. Dyn. (2020) 1–8.

[27] E. Jangam, A.A.D. Barreto, C.S.R. Annavarapu, Automatic detection of Covid-19 from chest Ct Scan and Chest X-Rays images using deep learning, transfer learning and stacking, Appl. Intell. 1–17 (2021).

[28] R. Kundu, P.K. Singh, M. Ferrara, A. Ahmadian, R. Sarkar, Et-net: an ensemble of transfer learning models for prediction of covid-19 infection through chest ct-scan images, Multimed. Tool. Appl. (2021) 1–20.

[29] C. Li, Y. Yang, H. Liang, B. Wu, Transfer learning for establishment of recognition of covid-19 on ct imaging using small-sized training datasets, Knowl. Base Syst. 218 (2021) 106849.

[30] M. Loey, G. Manogaran, N.E.M. Khalifa, A deep transfer learning model with classical data augmentation and cgan to detect covid-19 from chest ct radiography digital images, Neural Comput. Appl. (2020) 1–13.

[31] 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.

[32] A.K. Mishra, S.K. Das, P. Roy, S. Bandyopadhyay, Identifying covid 19 from chest ct images: a deep convolutional neural networks based approach, J. Healthcare Eng. 2020 (2020).

[33] A. Mobiny, P.A. Cicalese, S. Zare, P. Yuan, M. Abavisani, C.C. Wu, J. Ahuja, P.M. de Groot, H. Van Nguyen, Radiologist-level Covid-19 Detection Using Ct Scans with Detail-Oriented Capsule Networks, 2020 arXiv preprint arXiv:2004.07407.

[34] Y. Mohamadou, A. Halidou, P.T. Kapen, A review of mathematical modeling, artificial intelligence and datasets used in the study, prediction and management of covid-19, Appl. Intell. 50 (2020) 3913–3925.

[35] L. Nanni, S. Ghidoni, S. Brahnam, Handcrafted vs. non-handcrafted features for computer vision classification, Pattern Recogn. 71 (2017) 158–172.

[36] B.N. Narayanan, R.C. Hardie, V. Krishnaraja, C. Karam, V.S.P. Davuluru, Transfer-to-transfer learning approach for computer aided detection of covid-19 in chest radiographs, A&I 1 (2020) 539–557.

[37] 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. (2021) 1–14.

[38] B. Nigam, A. Nigam, R. Jain, S. Dodia, N. Arora, B. Annappa, Covid-19: automatic detection from x-ray images by utilizing deep learning methods, Expert Syst. Appl. 176 (2021) 114883.

[39] U. Özkaya, Ş. Öztürk, S. Budak, F. Melgani, K. Polat, Classification of Covid-19 in Chest Ct Images Using Convolutional Support Vector Machines, 2020 arXiv preprint arXiv:2011.05746.

[40] H. Panwar, P. Gupta, M.K. Siddiqui, R. Morales-Menendez, P. Bhardwaj, V. Singh, A deep learning and grad-cam based color visualization approach for fast detection of covid-19 cases using chest x-ray and ct-scan images, Chaos, Solit. Fractals 140 (2020) 110190.

[41] S. Pathan, P. Siddalingaswamy, P. Kumar, M.P. Mm, T. Ali, U.R. Acharya, Novel ensemble of optimized cnn and dynamic selection techniques for accurate covid-19 screening using chest ct images, Comput. Biol. Med. (2021), 104835.

[42] L.T. Phan, T.V. Nguyen, Q.C. Luong, T.V. Nguyen, H.T. Nguyen, H.Q. Le, T. T. Nguyen, T.M. Cao, Q.D. Pham, Importation and human-to-human transmission of a novel coronavirus in vietnam, N. Engl. J. Med. 382 (2020) 872–874.

[43] M. Polsinelli, L. Cinque, G. Placidi, A light cnn for detecting covid-19 from ct scans of the chest, Pattern Recogn. Lett. 140 (2020) 95–100.

[44] M.M. Rahaman, C. Li, Y. Yao, F. Kulwa, M.A. Rahman, Q. Wang, S. Qi, F. Kong, X. Zhu, X. Zhao, Identification of covid-19 samples from chest x-ray images using deep learning: a comparison of transfer learning approaches, J. X Ray Sci. Technol. (2020) 1–19.

[45] J. Rasheed, A.A. Hameed, C. Djeddi, A. Jamil, F. Al-Turjman, A machine learning-based framework for diagnosis of covid-19 from chest x-ray images, Interdiscipl. Sci. Comput. Life Sci. 13 (2021) 103–117.

[46] C. Rothe, M. Schunk, P. Sothmann, G. Bretzel, G. Froeschl, C. Wallrauch, T. Zimmer, V. Thiel, C. Janke, W. Guggemos, et al., Transmission of 2019-ncov infection from an asymptomatic contact in Germany, N. Engl. J. Med. 382 (2020) 970–971.

[47] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, et al., Imagenet large scale visual recognition challenge, Int. J. Comput. Vis. 115 (2015) 211–252.

[48] F. Sanghavi, K. Panetta, S. Agaian, Covid-19 detection in ct images using custom weighted filter-based cnn, in: Multimodal Image Exploitation and Learning 2021, International Society for Optics and Photonics, 2021, p. 117340L.

[49] J.F.H. Santa Cruz, An ensemble approach for multi-stage transfer learning models for covid-19 detection from chest ct scans, Intell. Based Med. 5 (2021) 100027.

[50] M. Saqib, S. Anwar, A. Anwar, M. Blumenstein, et al., Covid19 Detection from Radiographs: is Deep Learning Able to Handle the Crisis?, TechRxiv, 2020.

[51] S. Sen, S. Saha, S. Chatterjee, S. Mirjalili, R. Sarkar, A bi-stage feature selection approach for covid-19 prediction using chest ct images, Appl. Intell. (2021) 1–16.

[52] N.S. Shaik, T.K. Cherukuri, Lesion-aware attention with neural support vector machine for retinopathy diagnosis, Mach. Vis. Appl. 32 (2021) 1–13.

[53] N.S. Shaik, T.K. Cherukuri, Multi-level attention network: application to brain tumor classification, Signal, Image Video Process. 1–8 (2021), https://doi.org/10.1007/s11760-021-02022-0.

[54] K. Simonyan, A. Zisserman, Very Deep Convolutional Networks for Large-Scale Image Recognition, 2014 arXiv preprint arXiv:1409.1556.

[55] C. Szegedy, S. Ioffe, V. Vanhoucke, A. Alemi, Inception-v4, inception-resnet and the impact of residual connections on learning, in: Proceedings of the AAAI Conference on Artificial Intelligence, 2017, pp. 1–12.

[56] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna, Rethinking the inception architecture for computer vision, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 2818–2826.

[57] M. Toğaçar, B. Ergen, Z. Cömert, Covid-19 detection using deep learning models to exploit social mimic optimization and structured chest x-ray images using fuzzy color and stacking approaches, Comput. Biol. Med. 121 (2020) 103805.

[58] N. Veeranjaneyulu, A. Raghunath, B.J. Devi, V.N. Mandhala, Scene classification using support vector machines with lda, J. Theor. Appl. Inf. Technol. 63 (2014).

[59] S. Wang, L. Zhang, W. Zuo, B. Zhang, Class-specific reconstruction transfer learning for visual recognition across domains, IEEE Trans. Image Process. 29 (2019) 2424–2438.

[60] Z. Zhang, B. Chen, J. Sun, Y. Luo, A bagging dynamic deep learning network for diagnosing covid-19, Sci. Rep. 11 (2021) 1–15.

[61] J. Zhao, Y. Zhang, X. He, P. Xie, Covid-ct-dataset: a Ct Scan Dataset about Covid-19, 2020 arXiv preprint arXiv:2003.13865 490.

[62] N. Zhu, D. Zhang, W. Wang, X. Li, B. Yang, J. Song, X. Zhao, B. Huang, W. Shi, R. Lu, et al., A novel coronavirus from patients with pneumonia in China, 2019, N. Engl. J. Med. (2020).