Transfer learning with fine-tuned deep CNN ResNet50 model for classifying COVID-19 from chest X-ray images

Md. Belal Hossain [a], S.M. Hasan Sazzad Iqbal [a], Md. Monirul Islam [b], Md. Nasim Akhtar [c], Iqbal H. Sarker [d,*]

[a] *Department of Computer Science and Engineering, Pabna University of Science and Technology, Pabna 6600, Bangladesh*
[b] *Department of Textile Engineering, Uttara University, Dhaka 1230, Bangladesh*
[c] *Department of Computer Science and Engineering, Dhaka University of Engineering Technology, Gazipur, 1707, Bangladesh*
[d] *Department of Computer Science and Engineering, Chittagong University of Engineering & Technology, Chittagong 4349, Bangladesh*

A B S T R A C T

COVID-19 cases are putting pressure on healthcare systems all around the world. Due to the lack of available testing kits, it is impractical for screening every patient with a respiratory ailment using traditional methods (RT-PCR). In addition, the tests have a high turn-around time and low sensitivity. Detecting suspected COVID-19 infections from the chest X-ray might help isolate high-risk people before the RT-PCR test. Most healthcare systems already have X-ray equipment, and because most current X-ray systems have already been computerized, there is no need to transfer the samples. The use of a chest X-ray to prioritize the selection of patients for subsequent RT-PCR testing is the motivation of this work. Transfer learning (TL) with fine-tuning on deep convolutional neural network-based ResNet50 model has been proposed in this work to classify COVID-19 patients from the COVID-19 Radiography Database. Ten distinct pre-trained weights, trained on varieties of large-scale datasets using various approaches such as supervised learning, self-supervised learning, and others, have been utilized in this work. Our proposed $iNat2021\_Mini\_SwAV\_1k$ model, pre-trained on the iNat2021 Mini dataset using the SwAV algorithm, outperforms the other ResNet50 TL models. For COVID instances in the two-class (Covid and Normal) classification, our work achieved 99.17% validation accuracy, 99.95% train accuracy, 99.31% precision, 99.03% sensitivity, and 99.17% F1-score. Some domain-adapted ($ImageNet\_ChestX-ray14$) and in-domain (ChexPert, ChestX-ray14) models looked promising in medical image classification by scoring significantly higher than other models.

1. Introduction

In December 2019, Wuhan, China, was the first person to develop a novel corona viral disease (called COVID-19). So far, millions and thousands of deaths have been confirmed worldwide [1]. A 2019 coronavirus disease (COVID-19) due to serious coronavirus acute syndrome 2 (SARS-CoV-2) has become a pandemic and widely spread around the world. The COVID-19 outbreak has affected many aspects, such as everyday life, public health, and the world economy. The World Health Organization announced on 28 June 2020 that more than 10 million confirmed COVID-19 cases have been reported in the world and over 499,000 persons have died. Furthermore, in China, approx. 6.47 (range 1.66–10) [2], in South Korea 2.6 [3], and in Iran 4.7 [4] in the basic reproductive (R0), are defined as the average number of secondary cases born by an infected individual, which indicates that there is a serious spread of COVID-19. Due to the non-availability of specific therapeutic

drugs or COVID-19 vaccines [5], the main priority is to stop the spread of COVID-19 through the testing of many suspicious cases and isolation from the community of the infected. As per Chinese government's most recent guidelines, the diagnosis of COVID-19 should be validated by a reverse transcription-polymerase chain reaction (RT-PCR). In terms of sensitivity, however, RT-PCR may not be high enough. False negatives can also occur when there is a lack of virus quantities in the sample; therefore, several times before the test is finally confirmed, it may be necessary [6–8]. As a result, rapid and precise diagnostic procedures or instruments are required to battle SARS-CoV-2 as soon as possible.

Chest CT is a routine pneumonia imagery tool that provides an advantage for COVID-19 diagnosis. Most COVID-19 patients show similar characteristics on CT images, such as opacity of the ground glass, pulmonary consolidation, and/or peripheral-lung changes [9,10]. Whilst chest CT may be an early screening method for COVID-19, this and

other kinds of infectious and inflammatory lung illness have significant imaging similarities. It is therefore not easy to distinguish between COVID-19 and other viral pneumonia. Radiologists can also take a long time to identify the characteristics. In addition, manual reading of CT images is a long and tiring task, leading to human error in turn. Thus, technology using automated analysis based on artificial intelligence (AI) can assist radiologists in the analysis of COVID-19 from CT images.

The important achievement in AI is deep learning (DL) [11]. The convolutional neural network (CNN) is one of the typical DL architectures [12]. Because of its strong characteristics [13–17], CNN has been widely applied in the health sector. CNN techniques can help in the accurate detection and classification of COVID-19 along with the implementation of radiological imaging [18]. Recent CNN images of COVID-19 and Non-COVID-19 have been classified using a CT-Classification approach [19–22].

Our experiment utilized the transfer learning (TL) [12] method on a well-known CNN model named ResNet50 to build our COVID-19 detection model. We used 10 different pre-trained weights and all of those ResNet50 model's weights trained on different image datasets. We utilized those pre-trained weights and applied the transfer learning approach on a different chest X-ray images dataset for classifying Covid or Normal patients. Some pre-trained models performed very well and we found some insights on the TL approach for medical images classification system especially on chest X-ray images. We applied to fine-tune on ResNet50 model for better performance. This work achieves over 99% accuracy for classifying Covid and Normal patients on chest X-ray images. The contribution of our work is as follows:

- We propose a fine-tuned ResNet50 model applying transfer learning technique with ten different pre-trained weights for classifying COVID-19 from chest X-ray images.
- We have modified ResNet50 model by adding extra two fully connected layers than the default ResNet50 model for applying fine tuning in our task.
- We conduct experiments on the COVID-19 Radiography dataset utilizing pre-trained weights trained on varieties of large-scale datasets and comparison with existing models to show the effectiveness of the proposed model.

The remainder of the paper is structured as follows. In Section 2, we presented a quick literature review. The proposed methodology is stated in Section 3, which also explains experimental Setup, Model Training, and Evaluation. The result analysis and discussions are presented in Section 4. Finally, Section 5 concludes the paper.

2. Literature review

Recently, many researchers did the research in medical sector especially medical image processing techniques. They used machine learning (ML) techniques and DL techniques and so on. ML tools are widely accepted as a prominent tool to improve the prediction and diagnosis of many diseases [23,24]. However, to obtain better ML models, efficient extraction techniques are needed. In medical imaging systems, DL models are therefore widely accepted as they can automatically extract features or use some pre-trained networks like ResNet [25]. In [26], the VGG16 model is proposed for classifying COVID-19 pneumonia and non-COVID-19 pneumonia on the chest X-ray images. In [27], the authors used the pulmonary nodules in CT images through a multifaceted convolution network. The deep networks [28] are used to segment abdominal CT images by deep opponents. The authors used 3-D CNN in Chest CT for the detection of chest nodules [29]. In order to classify the coronary artery plaque and stenosis in the Coronary CT, authors applied a classification procedure using recurrent CNN [30]. A new CNN model was proposed for chest X-ray image classification [31]. As pre-trained CNN models are known to present practical problems, the authors developed a shallow CNN architecture. The authors have used a 12-class chest X-ray image dataset with a measured accuracy

Table 1
COVID-19 radiography database.

Covid	Lung opacity	Viral pneumonia	Normal	Total
3616	6012	1345	10 192	21 165

of 86%. Nardelli et al. [32] used CNN 3-D to categorize chest CT pulmonary artery-vein. To classify CT-images of interstitial lung disease a deep CNN was applied [33]. A TL approach was applied on VGG16 and ResNet50 models for classifying COVID-19, pneumonia infection and, no infection [34]. The study showed 97.67% accuracy using the VGG16 TL model on the Covid-19 radiography dataset. A deep CNN CoroNet [35] was proposed and showed 89% accuracy on chest X-ray images for 4-class classification including COVID-19 class. DarkCovidNet [18], a deep model was proposed on chest X-ray images. A deep CNN DeTraC [36] model was proposed to classify COVID-19 chest X-ray images and showed 93.1% accuracy.

Unlike the above studies, in this work, we propose a fine-tuned ResNet50 model applying transfer learning technique for effectively classifying COVID-19 from chest X-ray images, where we have modified ResNet50 model by adding extra two fully connected layers than the default ResNet50 model.

3. Proposed methodology

In this work, we proposed a transfer learning method to classify Covid vs Normal patients from chest X-ray images by utilizing various pre-trained ResNet50 TL models weights. A diagram of the proposed method is depicted in Fig. 1.

Chest X-ray images dataset is used and we split the dataset randomly into train and test (validation) sets. Data transformation is applied without data augmentation. Ten different pre-trained ResNet50 model weights are used which are trained on different types of natural and medical images dataset. ResNet50 [37] model has loaded with these pre-trained weights with fine-tuning the model for better performance. Then training and evaluation are applied on these ResNet50 fine-tuned TL models. After training and evaluation are performed, we compare these TL models and propose the best ResNet50 fine-tuned TL model. Finally, chest X-ray images are input into the proposed model and, it provides the predicted output as either Covid or Normal using binary classification.

3.1. Dataset description

The researchers used the COVID-19 Radiography Database to carry out experiments [38,39]. A team of researchers and collaborators with medical doctors created this database from multiple international sources at different timescales [14,40–42]. There are four categories (Covid, Lung Opacity, Viral Pneumonia, and Normal) of chest X-ray (CXR) images in this database. All images are of 299 × 299 pixel resolution and in Portable Network Graphics (PNG) file format. The images are in gray-scale having three channels that contain repeated RGB values.

Table 1 presents the number of images of different classes in the COVID-19 Radiography Database. The numbers of images are not balanced for all four concept classes. Two classes out of four classes comprising Covid and Normal are used in our work. To create a balanced dataset, we select a subset of the Normal class.

Table 2 presents the number of images used for the train and test (validation) set of two different classes. We divided the datasets into two sets train set and test or validation set by selecting images randomly with proportions 80% and 20%, respectively. The train set is used for model training and the test set is used for validating the model. Some Covid and Normal patients CXR images of this dataset are shown in Figs. 2 and 3.
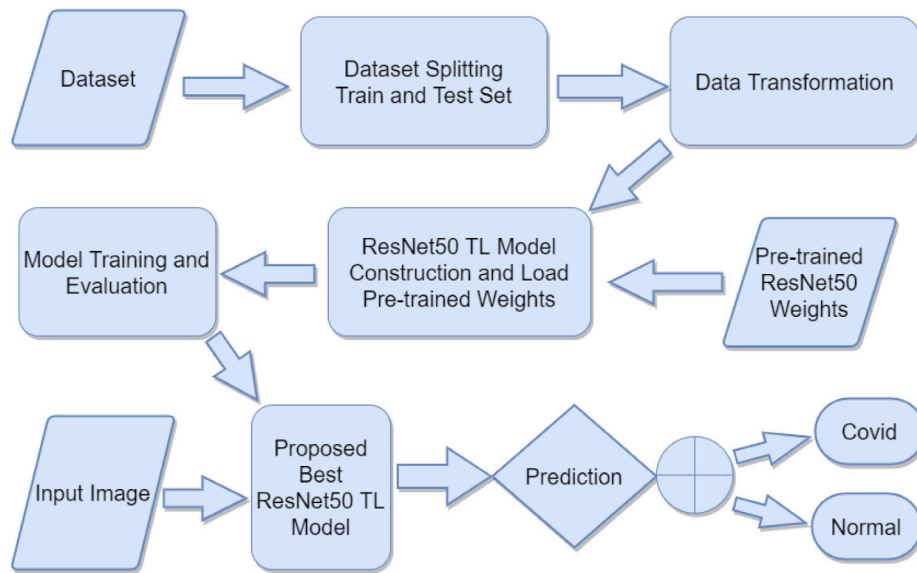
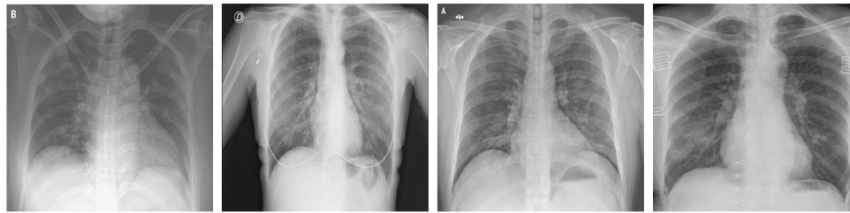**Fig. 1.** A block diagram of the proposed methodology.

Fig. 2. Example of Covid CXR images of COVID-19 Radiography Database.
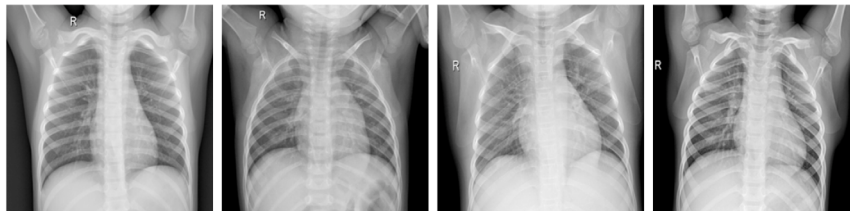

Fig. 3. Example of Normal CXR images of COVID-19 Radiography Database.

Table 2
Dataset splitting of Covid and Normal images into train and test set.

Set	Covid	Normal	Total
Train	2892	2917	5809
Test	723	730	1453
Total	3615	3647	7262

Data preprocessing, augmentation, and transformation are important parts of some computer vision models. In our experiment, no data augmentation techniques are applied. We only applied a transformation to resize the images to 224×224 pixel resolution from 299×299 pixel resolution for the ResNet50 model.

3.2. ResNet50 TL model and pre-trained weights

One of the well-known models which perform very well in different computer vision problems is ResNet [37]. Some of the others are VGG [43], DenseNet [44], Inception v3 [45], AlexNet [46], MobileNet [47], GoogLeNet [48] etc. These models are trained on a vast amount of datasets with various categories of images. Transfer learning techniques can utilize these pre-trained models weights on different computer vision problems with limited resources (dataset and computing resources). In this work, we utilized 10 different pre-trained weights of the ResNet50 model to carry out transfer learning on the limited medical images dataset (COVID-19 Radiography Database). In the following parts, we have described the architecture of the ResNet50 TL model and 10 different pre-trained weights. ResNet50 model is a CNN model consisting of 50 layers. The architecture of the ResNet50 model including fine-tuning configuration for ResNet50 TL is depicted in Fig. 4.

Also the architecture for proposed fine-tuned ResNet50 TL presented in Table 3.

There are a series of convolutional (conv) layers in ResNet50 architecture. First conv layer is made of 7×7 kernel size and 64 different kernels with a stride size of 2. Then 3×3 max pooling with the stride of size 2 is applied. In the next convolution, there are three conv layers $(1 \times 1, 64$ kernels$)$, $(3 \times 3, 64$ kernels$)$ and $(1 \times 1, 256$ kernels$)$ respectively and these three layers are repeated in total 3 times. In the same process, three conv layers $(1 \times 1, 128$ kernels$)$, $(3 \times 3, 128$ kernels$)$ and $(1 \times 1, 512$ kernels$)$ respectively are repeated four times; three conv layers $(1 \times 1, 256$ kernels$)$, $(3 \times 3, 256$ kernels$)$ and $(1 \times 1,$
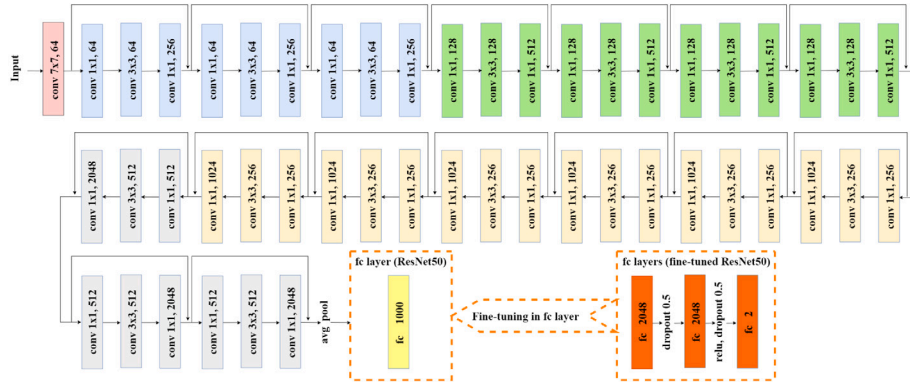
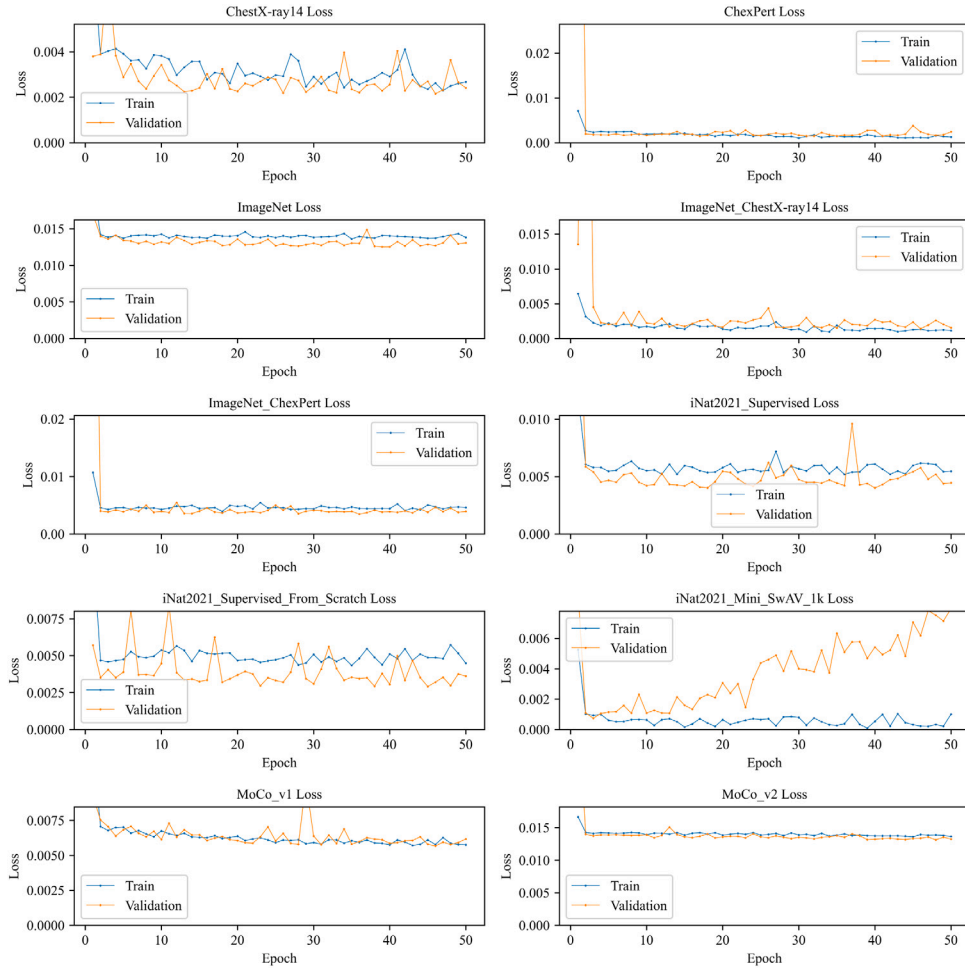**Fig. 4.** Fine tuned ResNet50 TL architecture.

Fig. 5. Train and validation loss of different ResNet50 TL models.

1024 kernels) respectively are repeated six times and another three conv layers (1 × 1, 512 kernels), (3 × 3, 512 kernels) and (1 × 1, 2048 kernels) respectively are repeated 3 times. Then average polling (avg pool) is applied. Most of the hidden layers use Batch Normalization, and ReLU followed by a conv layer. The final layer of the original ResNet50 model is a fully connected (fc) layer with 1000 out-features (for 1000 class). We fine-tune the ResNet50 model by replacing this fc layer with a set of fc layers. The first fc layer has 2048 out-features, then dropout with a probability of 0.5 is applied. The second fc layer is the same as the first fc layer. After the second fc layer, ReLU and dropout

with probability 0.5 are applied. The final fc layer consists of 2048 in-features and only 2 out-features for two-class classification (Covid vs Normal).

In this work, we experimented with transfer learning of 10 different pre-trained weights of the ResNet50 model. These pre-trained weights are generated from different datasets. These datasets have many variations. As we are experimenting with medical images datasets. Medical images (chest X-rays) are very different than other real-world images. Researchers are experimenting on various aspects of transfer learning for medical images like chest X-rays and CT scan images. Supervised
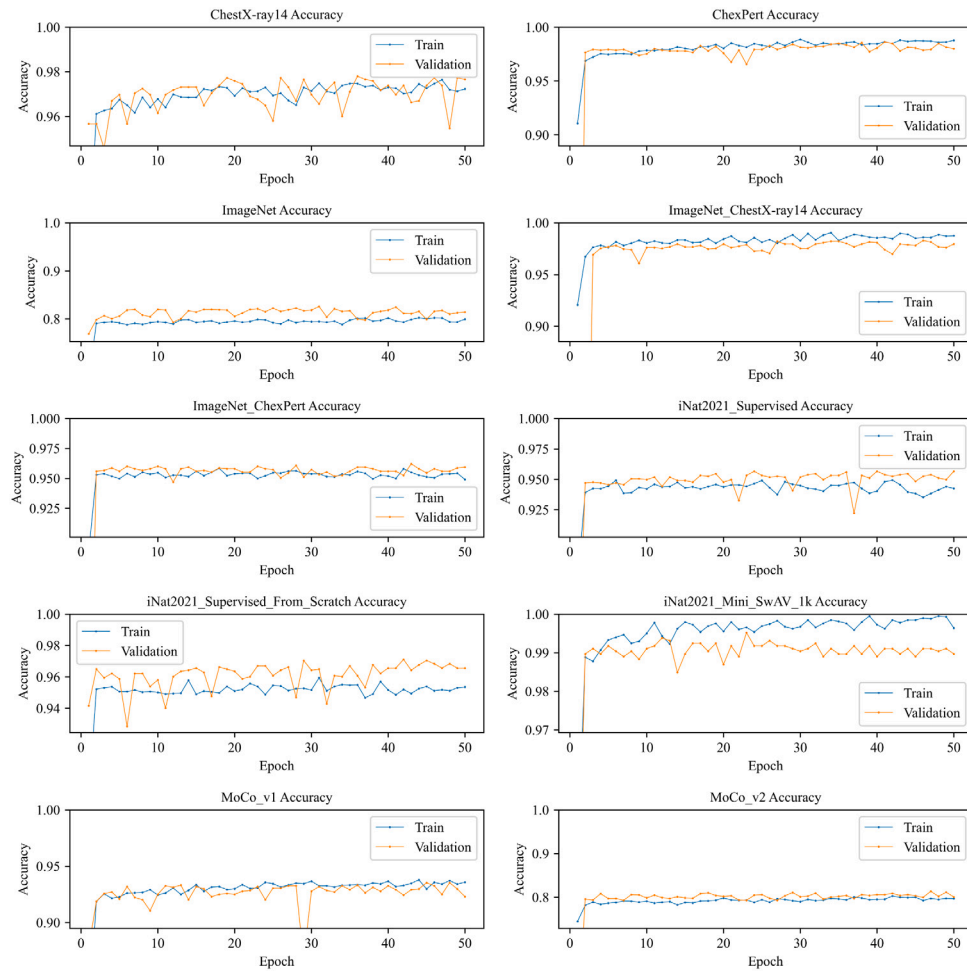
Fig. 6. Train and validation accuracy of different ResNet50 TL models.

Table 3
Architecture for proposed fine-tuned ResNet50 TL. Building blocks are shown in brackets, with the numbers of blocks stacked. Downsampling is performed by conv3_1, conv4_1, and conv5_1 with a stride of 2.

Layer name	Output size	Layer
conv1	112×112	7×7, 64, stride 2
conv2_x	56×56	3×3 max pool, stride 2
		$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
conv3_x	28×28	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$
conv4_x	14×14	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$
conv5_x	7×7	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$
fc1	1×1	Average pool in_features = 2048, out_features = 2048
fc2	1×1	Dropout 0.5 in_features = 2048, out_features = 2048
fc3	1×1	relu, dropout 0.5 in_features = 2048, out_features=2

transfer learning model trained on ImageNet dataset has been frequently used in medical image analysis. The recent study [49] proposed that pre-trained models on fine-grained data are suitable for medical analysis and continual pre-training can reduce the domain gap between natural and medical images. The iNat and NeWT are two large-scale fine-grained natural image data collections. Researchers explored different questions and carried out various analyses on iNat2021, NeWT, and ImageNet datasets as well as showed the strengths and weaknesses of various methods for transferability of supervised and self-supervised learning [50]. A brief description of pre-trained weights used in our experiments are listed below:

3.2.1. ChestX-ray14

This pre-trained weight is an in-domain ResNet50 model trained on the ChestX-ray14 dataset [51] consists of 112K frontal-view chest X-ray images of 30K unique patients. The detailed information on this pre-trained weight can be found in Taher et al. [49].

3.2.2. ChexPert

This pre-trained weight is an in-domain ResNet50 model trained on a large-scale publicly available ChexPert dataset [52] consisting of 224K chest X-ray images of 65K patients. The detail about this pre-trained weight can be found in Taher et al. [49].

3.2.3. ImageNet

This pre-trained weight is the ResNet50 model trained on a well-known ImageNet dataset [53]. These pre-trained weights are loaded from the PyTorch framework's pre-trained models collection (*torch.utils.model_zoo*).
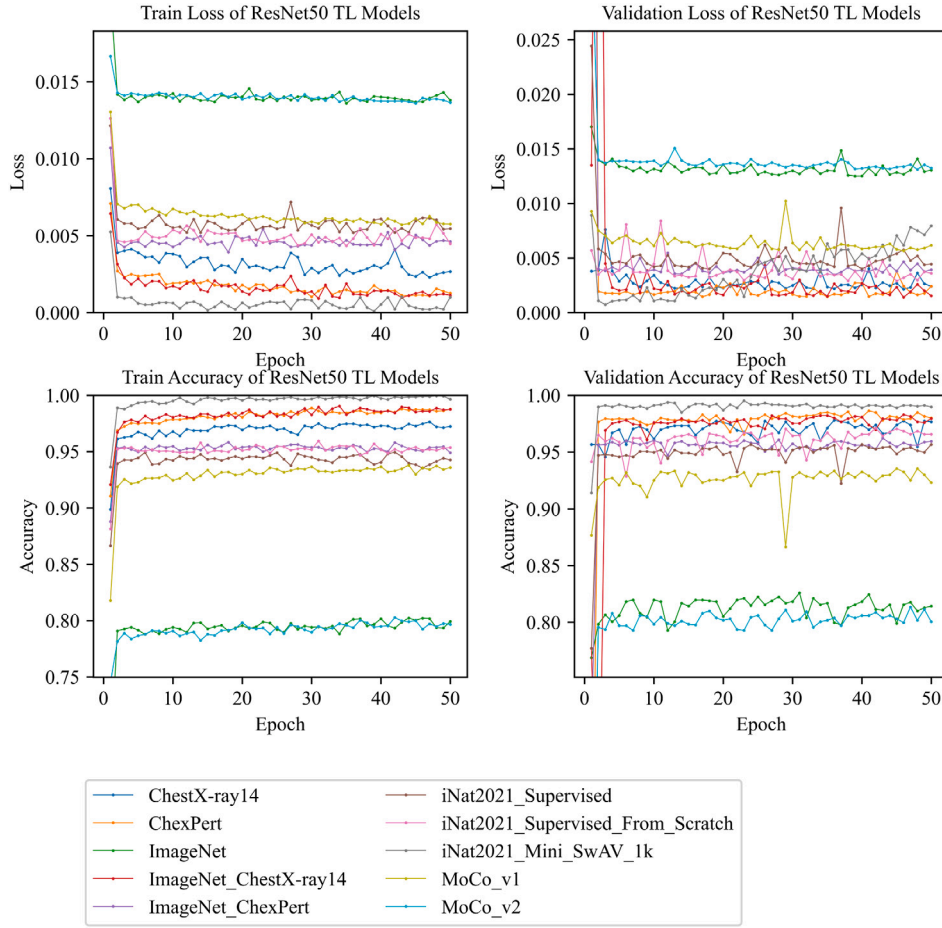
Fig. 7. Train and validation losses and accuracies of all ResNet50 TL models.

3.2.4. $ImageNet\_ChestX-ray14$

A domain-adapted pre-trained ResNet50 model's weight. This model was trained on two different datasets. Firstly, the model was initialized with pre-trained ImageNet weight and then again trained on the ChestX-ray14 dataset [49].

3.2.5. $ImageNet\_ChexPert$

This is also a domain-adapted pre-trained ResNet50 model's weight trained on two different datasets. The model was initialized with pre-trained ImageNet weight and then again trained on the ChexPert dataset [49].

3.2.6. $iNat2021\_Supervised$

The iNat2021 [50] is a large-scale natural world dataset consisting of 2.7M images of 10K different species. This pre-trained supervised ResNet50 model's weight was collected. This pre-trained weight was downloaded manually from the given source on that paper and the wight file name was $inat2021\_supervised\_large.pth.tar$. The model was initialized with pre-trained ImageNet weight and then trained on the iNat2021 dataset.

3.2.7. $iNat2021\_Supervised\_From\_Scratch$

This pre-trained supervised ResNet50 model's weight was also collected from the same source [50] as $iNat2021\_Supervised$. But this model only trained on the iNat2021 dataset instead of initializing with pre-trainedImageNet weight.$inat2021\_supervised\_large\_from\_scratch.pth.tar$ was the name of the pre-trained weight file.

3.2.8. $iNat2021\_Mini\_SwAV\_1k$

The iNat2021 Mini [50] is a smaller (500K images) version of the iNat2021 dataset that contains 50 training images per species. This pre-trained ResNet50 model utilized a self-supervised SwAV algorithm [54] while training on the iNat2021 Mini dataset. The pre-trained weight is collected from the same source paper as $iNat2021\_Supervised$ and $iNat2021\_Supervised\_From\_Scratch$. The file name of the pre-trained weight was $inat2021\_swav\_mini\_1000\_ep.pth$.

3.2.9. $MoCo\_v1$

One of the self-supervised learning (SSL) methods is MoCo (Momentum Contrast) v1. This is a contrastive learning method. The pre-trained ResNet50 model's weight which is trained on the ImageNet dataset utilizing MoCo v1 was collected [55]. This weight was trained over 200 epochs.

3.2.10. $MoCo\_v2$

The SSL method $MoCo\_v2$ is the improved version of $MoCo\_v1$. The pre-trained ResNet50 TL model's weight which is trained on the ImageNet dataset utilizing $MoCo\_v2$ was collected from these [55,56] research paper. This weight was trained over 800 epochs.

All these pre-trained ResNet50 model weights are used to initialize our ResNet50 TL models. After initializing these pre-trained weights, we trained our models on the COVID-19 Radiography dataset.

3.3. Experimental setup, training, and evaluation

Model training and evaluation tasks are performed on Google Colaboratory with GPU runtime. We make use of PyTorch, an open-source
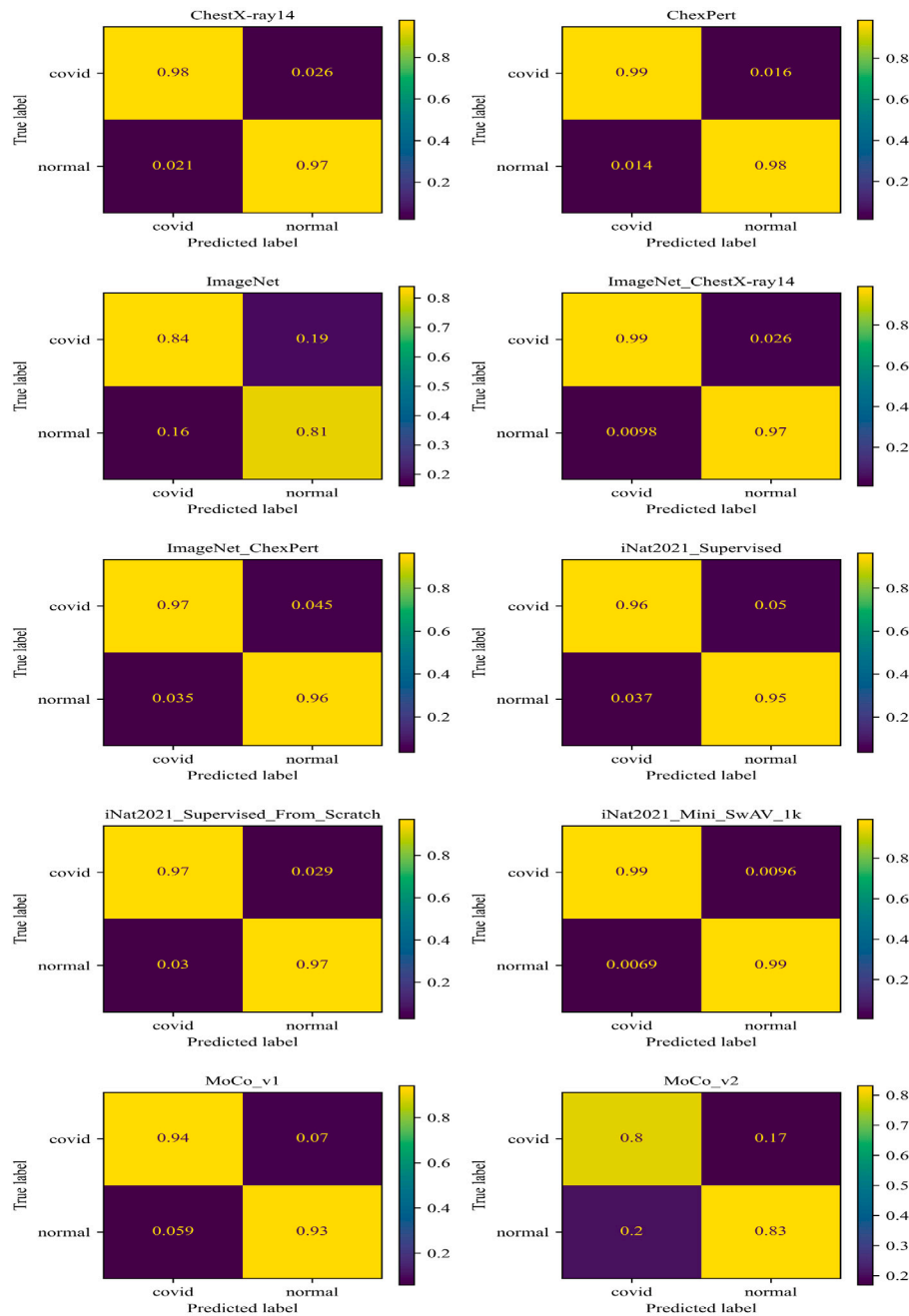
Fig. 8. Confusion Matrix of ResNet50 TL models.

machine learning package created largely by Facebook's AI Research department. PyTorch data loader is used to load the data with a batch size of 32. We use the same configuration and same dataset for all models. Datasets are downloaded from Kaggle datasets (COVID-19 Radiography Database) using Kaggle API. Google drive is used for storing the dataset's file names as well as trained models checkpoints.

The repeated training technique is taken for several epochs. We trained all models 50 epochs. The training dataset is used for model training, and the test dataset is used for model evaluation. In each epoch, we calculate the cross-entropy loss on the train and test set. We use Adam [57] optimizer with a learning rate of 0.001, betas of (0.9, 0.999), eps of 1e−08, weight decay of 0. Table 4 presents hyperparameters of our experimental setup.

The train and validation loss depicted in Fig. 5 gives us an overview of how the model trained over a series of epochs. We see the train and validation loss of most of the models are very close to and overlap each

Table 4
Hyperparameters of ResNet50 TL model.

Parameters	Parameters value
Batch size	32
Optimizer	Adam
Learning rate	0.001
Betas	(0.9, 0.999)
Eps	1e−08
Weight decay	0
Criterion	Cross Eentropy Loss

other. But $iNat2021\_Mini\_SwAV\_1k$ model shows a different picture. The validation loss is increased over epochs while the training loss decreased slightly. Although the initial loss is very little compared to other models. On the other hand, $MoCo\_v2$ and ImageNet do not

show improvement over a higher number of epochs. Another thing we noticed is that decreasing rate of loss was very high in the initial stages (fewer epochs) and this rate becomes almost constant on higher epochs. This could be because of the transfer learning technique with pre-trained weights.

Train and validation accuracy of different models are depicted in Fig. 6. It shows that most of the models reaches their highest accuracy very quickly because of transfer learning with pre-trained weights. No overfitting is shown up in any models. $MoCo\_v1$ shows a fluctuation of validation accuracy at the 29th epoch where it sharply decreases the validation accuracy. We can also see a similar pattern in $Chest\_X-ray14$ but the accuracy difference is smaller than $MoCo\_v1$. On the other hand, we see an interesting pattern on ImageNet and $MoCo\_v2$ models that validation accuracy is higher than train accuracy but does not overlap like $iNat2021\_Supervised$ and $iNat2021\_Supervised\_From\_Scratch$. The opposite picture we can see in ChexPert, $ImageNet\_ChestX-ray14$, and $iNat2021\_Mini\_SwAV\_1k$ models where train accuracy is higher than validation accuracy but the difference of those accuracies are little.

Fig. 7, shows a comparative summary of train and validation losses as well as train and validation accuracies of all models. If we compare train loss of different models, it shows that $iNat2021\_Mini\_SwAV\_1k$ comes first in terms of lowest loss and $MoCo\_v2$ and ImageNet come last over 50 epochs of training. In terms of validation loss, we can see $iNat2021\_Mini\_SwAV\_1k$ achieves the lowest loss in the fewer number of epochs but the loss increases gradually in higher epochs. ChexPert and $ImageNet\_ChestX-ray14$ compete with each other for the lowest loss and $MoCo\_v2$ and ImageNet competes for the highest loss. In terms of training and validation accuracy, $iNat2021\_Mini\_SwAV\_1k$ comes first with the highest accuracy, and $MoCo\_v2$ and, ImageNet come last with lower accuracies. Hence, it shows that $iNat2021\_Mini\_SwAV\_1k$, $ImageNet\_ChestX-ray14$, and ChexPert models perform very well with higher training and validation accuracy.

4. Result analysis and discussion

Confusion Matrix, Precision, Recall, F1 Score, and Accuracy are the standard evaluation method for the classification model. As the name suggests, the Confusion Matrix gives us a matrix as output and describes the complete performance of the model [58]. It has four terminologies including True Positive (TP), False Positive (FP), False Negative (FN), and True Negative (TN).

Precision is the number of correctly identified cases among all the identified cases.

$$Precision, P = TP/(TP + FP) \qquad (1)$$

Recall is the number of correctly identified cases from all the positive representations.

$$Recall, R = TP/(TP + FN) \qquad (2)$$

F1 Score is the harmonic average of precision and recall.

$$F1 Score, F1 = 2 \times (P \times R)/(P + R) \qquad (3)$$

Accuracy, on the other hand, can be define as the ratio of corrected predictions to the total input samples.

$$Accuracy = (Number of correct predictions)$$
$$/(Total number of predictions made) \qquad (4)$$

The confusion matrix of the evaluation of all models is shown in Fig. 8. It shows that $iNat2021\_Mini\_SwAV\_1k$ correctly identified Covid and Normal (without COVID-19) with 99% samples. ChexPert classified Covid with 99% and Normal with 98% samples correctly. $ImageNet\_ChestX-ray14$ classified Covid with 99% and Normal with 97% samples correctly. And ChestX-ray14 classified Covid with 98% and Normal with 97% samples correctly. On the other hand, $MoCo\_v2$

Table 5
Various scores calculated in test dataset for different ResNet50 TL model where Pre = Precision, Re = Recall, F1 = F1-score, Sup = Support, Acc = Accuracy, INCXR14 = $ImageNet\_ChestX-ray14$, INCxP = $ImageNet\_ChexPert$, iNSup = $iNat2021\_Supervised$, iNSupFS = $iNat2021\_Supervised\_From\_Scratch$, iNMSwAV = $iNat2021\_Mini\_SwAV\_1k$.

Model	Class	Pre	Re	F1	Sup	Acc
ChestX-ray14	Covid	0.9791	0.9737	0.9764	723	0.9766
	Normal	0.9741	0.9795	0.9768	730	
ChexPert	Covid	0.9861	0.9834	0.9848	723	0.9849
	Normal	0.9836	0.9863	0.9850	730	
ImageNet	Covid	0.8394	0.7953	0.8168	723	0.8224
	Normal	0.8073	0.8493	0.8278	730	
INCXR14	Covid	0.9902	0.9737	0.9819	723	0.9821
	Normal	0.9744	0.9904	0.9823	730	
INCxP	Covid	0.9650	0.9544	0.9597	723	0.9601
	Normal	0.9553	0.9658	0.9605	730	
iNSup	Covid	0.9635	0.9488	0.9561	723	0.9566
	Normal	0.9501	0.9644	0.9572	730	
iNSupFS	Covid	0.9696	0.9710	0.9703	723	0.9704
	Normal	0.9712	0.9699	0.9705	730	
iNMSwAV	Covid	0.9931	0.9903	0.9917	723	0.9917
	Normal	0.9904	0.9932	0.9918	730	
$MoCo\_v1$	Covid	0.9411	0.9281	0.9345	723	0.9353
	Normal	0.9297	0.9425	0.9361	730	
$MoCo\_v2$	Covid	0.7974	0.8382	0.8173	723	0.8135
	Normal	0.8312	0.7891	0.8096	730	

and ImageNet scores are very low compared to other models. We can conclude that the $iNat2021\_Mini\_SwAV\_1k$ achieves the highest score among all other models.

Various standard evaluation scores are shown in Table 5, where column Support represents the number of samples is used during evaluation. The evaluation has been done on the test dataset which contains 1453 chest X-ray images where 723 Covid class and 730 Normal class. The precision, recall, and f1-score are calculated both for the Covid class and Normal class.

Table 5 shows that the $iNat2021\_Mini\_SwAV\_1k$ model achieves the highest precision, recall, f1-score, and accuracy score among all models. Also, the scores are very promising. The precision, recall and f1-score are 99.31%, 99.03% and 99.17% respectively for Covid class and 99.04%, 99.32% and 99.18% respectively for Normal class including 99.17% validation accuracy. That means all the evaluation scores are above 99%. On the other hand, ImageNet and $MoCo\_v2$ have gotten minimum scores amongst other models. Although $ImageNet\_ChestX-ray14$, ChexPert and ChestX-ray14 have achieved above 97% f1-score, the $ImageNet\_ChestX-ray14$ identified covid class over complete samples 99% correctly.

Table 6 shows the maximum, minimum, and average accuracy as a summary of all TL models during training and validation. From the table, we can summarize that the $iNat2021\_Mini\_SwAV\_1k$ model has reached the highest position by achieving maximum train (99.95%) and validation (99.52%) accuracy among other models.

4.1. Comparison result

Several numbers of research have been done for the identification and classification of COVID-19 in recent years. Researchers are trying to find out a better way to solve this problem. We examined and compared our suggested model with an existing similar work in this part shown in Table 7.

The above-related studies proposed different approaches to carry out the solution for classifying COVID-19 from medical images. Our work has achieved 99.17% validation accuracy on the COVID-19 Radiography dataset. Ten different pre-trained ResNet50 are used in this work. Among them, the $iNat2021\_Mini\_SwAV\_1k$ model shows excellent performance.

Table 6

Accuracy summary of different ResNet50 TL models during 50 epochs of training where Acmx = Maximum Accuracy, Acmn = Minimum Accuracy, Avacy = Average Accuracy, E = epoch, INCXR14 = $ImageNet\_ChestX-ray14$, INCxP = $ImageNet\_ChexPert$, iNSup = $iNat2021\_Supervised$, iNSupFS = $iNat2021\_Supervised\_From\_Scratch$, iNMSwAV = $iNat2021\_Mini\_SwAV\_1k$.

Model	Set	Acmx	Acmx at E	Acmn	Acmn at E	Avacy of 50E
ChestX-ray14	Train	0.9764	47	0.8988	1	0.9688
	Test	0.9780	36	0.9456	3	0.9696
ChexPert	Train	0.9886	30	0.9107	1	0.9810
	Test	0.9862	41	0.5038	1	0.9701
ImageNet	Train	0.8024	44	0.6729	1	0.7925
	Test	0.8259	31	0.7688	1	0.8123
INCXR14	Train	0.9904	34	0.9206	1	0.9823
	Test	0.9828	46	0.5561	2	0.9644
INCxP	Train	0.9583	18	0.8879	1	0.9519
	Test	0.9621	43	0.6965	1	0.9514
iNSup	Train	0.9494	42	0.8666	1	0.9419
	Test	0.9566	24	0.7770	1	0.9467
iNSupFS	Train	0.9594	31	0.8816	1	0.9506
	Test	0.9711	42	0.9284	6	0.9611
iNMSwAV	Train	**0.9995**	39	0.9363	1	0.9951
	Test	**0.9952**	23	0.9140	1	0.9892
$MoCo\_v1$	Train	0.9379	44	0.8179	1	0.9289
	Test	0.9353	45	0.8665	29	0.9256
$MoCo\_v2$	Train	0.8027	42	0.7447	1	0.7917
	Test	0.8135	47	0.5045	1	0.7961

Table 7

Comparison of the proposed model using similar existing studies.

SN	Reference	Method	Accuracy (%)
01	[59]	COVIDX-Net	90
02	[60]	MADE-based CNN	94.65 ± 2.1
03	[61]	VGG16	80
04	[62]	Deep CNN	93
05	[63]	UNet3D Deep Network	90.80
06	[21]	ResNet + Location Attention	86.70
07	[36]	DeTraC Deep CNN	93.1
08	[34]	VGG16	97.67
09	Proposed	$iNat2021\_Mini\_SwAV\_1k$ (Fine-tuned ResNet50)	99.17

5. Conclusion and future work

In this paper, we have presented a fine-tuned ResNet50 model applying transfer learning technique for effectively classifying COVID-19 from chest X-ray images. For this, we have modified the ResNet50 model by adding extra two fully connected layers than the default ResNet50 model. We have utilized ten different pre-trained weights, trained on varieties of large-scale datasets applying various methods such as supervised learning, self-supervised learning, etc. Among these ResNet50 TL models, the $iNat2021\_Mini\_SwAV\_1k$ model performs excellent that was pre-trained on the iNat2021 Mini dataset with the SwAV algorithm. Our work achieved 99.17% validation accuracy, 99.95% train accuracy, the precision of 99.31%, recall of 99.03%, and an F1-score of 99.17% for covid cases in the two-class classification (Covid and Normal). Some domain-adapted ($ImageNet\_ChestX-ray14$) and in-domain (ChexPert, ChestX-ray14) models showed promising importance in medical image classification by achieving relatively better scores than models only trained on ImageNet dataset. As future work, we can conduct further studies to present a qualitative representation of transfer knowledge to our model and detect the region of interest of COVID-19 in medical images.

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.

References

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

[2] Tang B, Wang X, Li Q, Bragazzi NL, Tang S, Xiao Y, Wu J. Estimation of the transmission risk of the 2019-nCoV and its implication for public health interventions. J Clin Med 2020;9(2):462.

[3] Tang B, Xia F, Bragazzi NL, McCarthy Z, Wang X, He S, Sun X, Tang S, Xiao Y, Wu J. Lessons drawn from China and South Korea for managing COVID-19 epidemic: insights from a comparative modeling study. ISA Trans 2021.

[4] Ahmadi A, Fadaei Y, Shirani M, Rahmani F. Modeling and forecasting trend of COVID-19 epidemic in Iran until may 13, 2020. Med J Islamic Repub Iran 2020;34:27.

[5] Gao Q, Bao L, Mao H, Wang L, Xu K, Yang M, Li Y, Zhu L, Wang N, Lv Z, et al. Development of an inactivated vaccine candidate for SARS-CoV-2. Science 2020;369(6499):77–81.

[6] Corman VM, Landt O, Kaiser M, Molenkamp R, Meijer A, Chu DK, Bleicker T, Brünink S, Schneider J, Schmidt ML, et al. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. Eurosurveillance 2020;25(3):2000045.

[7] Chu DK, Pan Y, Cheng SM, Hui KP, Krishnan P, Liu Y, Ng DY, Wan CK, Yang P, Wang Q, et al. Molecular diagnosis of a novel coronavirus (2019-nCoV) causing an outbreak of pneumonia. Clin Chem 2020;66(4):549–55.

[8] Zhang N, Wang L, Deng X, Liang R, Su M, He C, Hu L, Su Y, Ren J, Yu F, et al. Recent advances in the detection of respiratory virus infection in humans. J Med Virol 2020;92(4):408–17.

[9] Ai T, Yang Z, Hou H, Zhan C, Chen C, Lv W, Tao Q, Sun Z, Xia L. Correlation of chest CT and RT-PCR testing for coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases. Radiology 2020;296(2):E32–40.

[10] Wang S, Kang B, Ma J, Zeng X, Xiao M, Guo J, Cai M, Yang J, Li Y, Meng X, et al. A deep learning algorithm using CT images to screen for corona virus disease (COVID-19). Eur Radiol 2021;1–9.

[11] Sarker IH. AI-based modeling: Techniques, applications and research issues towards automation, intelligent and smart systems. SN Comput Sci 2022;1–20.

[12] Sarker IH. Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions. SN Comput Sci 2021;2(6):1–20.

[13] Choe J, Lee SM, Do K-H, Lee G, Lee J-G, Lee SM, Seo JB. Deep learning–based image conversion of CT reconstruction kernels improves radiomics reproducibility for pulmonary nodules or masses. Radiology 2019;292(2):365–73.

[14] Kermany DS, Goldbaum M, Cai W, Valentim CC, Liang H, Baxter SL, McKeown A, Yang G, Wu X, Yan F, et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. Cell 2018;172(5):1122–31.

[15] Negassi M, Suarez-Ibarrola R, Hein S, Miernik A, Reiterer A. Application of artificial neural networks for automated analysis of cystoscopic images: a review of the current status and future prospects. World J Urol 2020;1–10.

[16] Wang P, Xiao X, Brown JRG, Berzin TM, Tu M, Xiong F, Hu X, Liu P, Song Y, Zhang D, et al. Development and validation of a deep-learning algorithm for the detection of polyps during colonoscopy. Nat Biomed Eng 2018;2(10):741–8.

[17] Yan Q, Wang B, Gong D, Luo C, Zhao W, Shen J, Shi Q, Jin S, Zhang L, You Z. COVID-19 chest CT image segmentation–a deep convolutional neural network solution. 2020, arXiv preprint arXiv:2004.10987.

[18] Ozturk T, Talo M, Yildirim EA, Baloglu UB, Yildirim O, Acharya UR. Automated detection of COVID-19 cases using deep neural networks with X-ray images. Comput Biol Med 2020;121:103792.

[19] Ardakani AA, Kanafi AR, Acharya UR, Khadem N, Mohammadi A. 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 2020;121:103795.

[20] Li L, Qin L, Xu Z, Yin Y, Wang X, Kong B, Bai J, Lu Y, Fang Z, Song Q, et al. Artificial intelligence distinguishes COVID-19 from community acquired pneumonia on chest CT. Radiology 2020.

[21] Xu X, Jiang X, Ma C, Du P, Li X, Lv S, Yu L, Ni Q, Chen Y, Su J, et al. A deep learning system to screen novel coronavirus disease 2019 pneumonia. Engineering 2020;6(10):1122–9.

[22] Chen J, Wu L, Zhang J, Zhang L, Gong D, Zhao Y, Chen Q, Huang S, Yang M, Yang X, et al. Deep learning-based model for detecting 2019 novel coronavirus pneumonia on high-resolution computed tomography. Sci Rep 2020;10(1):1–11.

[23] Shukla PK, Shukla PK, Sharma P, Rawat P, Samar J, Moriwal R, Kaur M. Efficient prediction of drug-drug interaction using deep learning models. IET Syst Biol 2020;14(4):211–6.

[24] Kaur M, Gianey HK, Singh D, Sabharwal M. Multi-objective differential evolution based random forest for e-health applications. Modern Phys Lett B 2019;33(05):1950022.

[25] Yu Y, Lin H, Meng J, Wei X, Guo H, Zhao Z. Deep transfer learning for modality classification of medical images. Information 2017;8(3):91.

[26] Nishio M, Noguchi S, Matsuo H, Murakami T. Automatic classification between COVID-19 pneumonia, non-COVID-19 pneumonia, and the healthy on chest X-ray image: combination of data augmentation methods. Sci Rep 2020;10(1):1–6.

[27] Setio AAA, Ciompi F, Litjens G, Gerke P, Jacobs C, Van Riel SJ, Wille MMW, Naqibullah M, Sánchez CI, Van Ginneken B. Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks. IEEE Trans Med Imaging 2016;35(5):1160–9.

[28] Xia K, Yin H, Qian P, Jiang Y, Wang S. Liver semantic segmentation algorithm based on improved deep adversarial networks in combination of weighted loss function on abdominal CT images. IEEE Access 2019;7:96349–58.

[29] Pezeshk A, Hamidian S, Petrick N, Sahiner B. 3-D convolutional neural networks for automatic detection of pulmonary nodules in chest CT. IEEE J Biomed Health Inf 2018;23(5):2080–90.

[30] Zreik M, Van Hamersvelt RW, Wolterink JM, Leiner T, Viergever MA, Išgum I. A recurrent CNN for automatic detection and classification of coronary artery plaque and stenosis in coronary CT angiography. IEEE Trans Med Imaging 2018;38(7):1588–98.

[31] Kesim E, Dokur Z, Olmez T. X-ray chest image classification by a small-sized convolutional neural network. In: 2019 scientific meeting on electrical-electronics & biomedical engineering and computer science (EBBT). IEEE; 2019, p. 1–5.

[32] Nardelli P, Jimenez-Carretero D, Bermejo-Pelaez D, Washko GR, Rahaghi FN, Ledesma-Carbayo MJ, Estépar RSJ. Pulmonary artery–vein classification in CT images using deep learning. IEEE Trans Med Imaging 2018;37(11):2428–40.

[33] Shin H-C, Roth HR, Gao M, Lu L, Xu Z, Nogues I, Yao J, Mollura D, Summers RM. Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. IEEE Trans Med Imaging 2016;35(5):1285–98.

[34] Das AK, Kalam S, Kumar C, Sinha D. TLCoV-An automated Covid-19 screening model using transfer learning from chest X-ray images. Chaos Solitons Fractals 2021;144:110713.

[35] Khan AI, Shah JL, Bhat MM. CoroNet: A deep neural network for detection and diagnosis of COVID-19 from chest x-ray images. Comput Methods Programs Biomed 2020;196:105581.

[36] Abbas A, Abdelsamea MM, Gaber MM. Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network. Appl Intell 2021;51(2):854–64.

[37] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2016, p. 770–8.

[38] Chowdhury ME, Rahman T, Khandakar A, Mazhar R, Kadir MA, Mahbub ZB, Islam KR, Khan MS, Iqbal A, Al Emadi N, et al. Can AI help in screening viral and COVID-19 pneumonia? IEEE Access 2020;8:132665–76.

[39] Rahman T, Khandakar A, Qiblawey Y, Tahir A, Kiranyaz S, Kashem SBA, Islam MT, Al Maadeed S, Zughaier SM, Khan MS, et al. Exploring the effect of image enhancement techniques on COVID-19 detection using chest X-ray images. Comput Biol Med 2021;132:104319.

[40] Vayá MdlI, Saborit JM, Montell JA, Pertusa A, Bustos A, Cazorla M, Galant J, Barber X, Orozco-Beltrán D, García-García F, et al. Bimcv covid-19+: a large annotated dataset of RX and CT images from covid-19 patients. 2020, arXiv preprint arXiv:2006.01174.

[41] Cohen JP, Morrison P, Dao L, Roth K, Duong TQ, Ghassemi M. Covid-19 image data collection: Prospective predictions are the future. 2020, arXiv preprint arXiv:2006.11988.

[42] Haghanifar A, Majdabadi MM, Choi Y, Deivalakshmi S, Ko S. Covid-cxnet: Detecting covid-19 in frontal chest x-ray images using deep learning. 2020, arXiv preprint arXiv:2006.13807.

[43] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. 2014, arXiv preprint arXiv:1409.1556.

[44] Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2017, p. 4700–8.

[45] Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2016, p. 2818–26.

[46] Krizhevsky A. One weird trick for parallelizing convolutional neural networks. 2014, arXiv preprint arXiv:1404.5997.

[47] Sandler M, Howard A, Zhu M, Zhmoginov A, Chen L-C. Mobilenetv2: Inverted residuals and linear bottlenecks. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2018, p. 4510–20.

[48] Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A. Going deeper with convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2015, p. 1–9.

[49] Hosseinzadeh Taher MR, Haghighi F, Feng R, Gotway MB, Liang J. A systematic benchmarking analysis of transfer learning for medical image analysis. In: Domain adaptation and representation transfer, and affordable healthcare and ai for resource diverse global health. Springer; 2021, p. 3–13.

[50] Van Horn G, Cole E, Beery S, Wilber K, Belongie S, Mac Aodha O. Benchmarking representation learning for natural world image collections. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition; 2021, p. 12884–93.

[51] Wang X, Peng Y, Lu L, Lu Z, Bagheri M, Summers RM. Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2017, p. 2097–106.

[52] Irvin J, Rajpurkar P, Ko M, Yu Y, Ciurea-Ilcus S, Chute C, Marklund H, Haghgoo B, Ball R, Shpanskaya K, et al. Chexpert: A large chest radiograph dataset with uncertainty labels and expert comparison. In: Proceedings of the AAAI conference on artificial intelligence, Vol. 33; 2019, p. 590–7.

[53] Deng J, Dong W, Socher R, Li L-J, Li K, Fei-Fei L. Imagenet: A large-scale hierarchical image database. In: 2009 IEEE conference on computer vision and pattern recognition. Ieee; 2009, p. 248–55.

[54] Caron M, Misra I, Mairal J, Goyal P, Bojanowski P, Joulin A. Unsupervised learning of visual features by contrasting cluster assignments. 2020, arXiv preprint arXiv:2006.09882.

[55] He K, Fan H, Wu Y, Xie S, Girshick R. Momentum contrast for unsupervised visual representation learning. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition; 2020, p. 9729–38.

[56] Chen X, Fan H, Girshick R, He K. Improved baselines with momentum contrastive learning. 2020, arXiv preprint arXiv:2003.04297.

[57] Kingma DP, Ba J. Adam: A method for stochastic optimization. 2014, arXiv preprint arXiv:1412.6980.

[58] Islam MM, Kashem MA, Uddin J. Fish survival prediction in an aquatic environment using random forest model. Int J Artif Intell ISSN 2021;2252(8938):8938.

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

[60] Singh D, Kumar V, Yadav V, Kaur M. Deep neural network-based screening model for COVID-19-infected patients using chest X-ray images. Int J Pattern Recognit Artif Intell 2021;35(03):2151004.

[61] Sahinbas K, Catak FO. Transfer learning-based convolutional neural network for COVID-19 detection with X-ray images. In: Data science for COVID-19. Elsevier; 2021, p. 451–66.

[62] Jamil M, Hussain I, et al. Automatic detection of COVID-19 infection from chest X-ray using deep learning. 2020, MedRxiv.

[63] Zheng C, Deng X, Fu Q, Zhou Q, Feng J, Ma H, Liu W, Wang X. Deep learning-based detection for COVID-19 from chest CT using weak label. 2020, MedRxiv.