International Journal on
"Technical and Physical Problems of Engineering"
(IJTPE)
Published by International Organization of IOTPE

ISSN 2077-3528
IJTPE Journal
www.iotpe.com
ijtpe@iotpe.com

June 2022 | Issue 51 | Volume 14 | Number 2 | Pages 104-110

SIAMESE NEURAL NETWORKS FOR PANDEMIC DETECTION USING CHEST RADIOGRAPHS

V. Jain [1] Ap. Jain [1] V. Garg [1] Ac. Jain [1] M. Demirci [2] M.C. Taplamacioglu [2]

1. Information Technology Department, Bharati Vidyapeeth's College of Engineering, New Delhi, India
vanita.jain@bharatividyapeeth.edu, jainapoorv341@gmail.com, vinayakk.garg99@gmail.com, achin.mails@gmail.com
2. Electrical and Electronics Engineering Department, Gazi University, Ankara, Turkey
mervedemirci@gazi.edu.tr, taplam@gazi.edu.tr

Abstract- The recent developments in the field of deep learning have enabled the efficient diagnosis of medical imaging for determining a broad set of diseases. To reduce the spread and impact of the pandemic (COVID virus), machine learning techniques can be used to diagnose and predict the disease using chest X-ray images. In this research, we present an approach using Siamese Convolutional Neural Network (SCNN) to classify chest x-ray images into four classes, namely pandemic, Severe-COVID, Pneumonia and Normal. We present a comparative study between the performance of our Siamese network and other pre-trained CNN architectures i.e. VGG-16 and ResNet50 in this research. The model performance is tested by merging two publicly available datasets: COVID-Chest-Xray dataset and Chest X-Ray Images (Pneumonia). We achieved an accuracy of 98% on Siamese ResNet50 which gives the best performance in contrast to 95% on VGG-16, 93% on ResNet50 and 96% on Siamese VGG-16.

Keywords: Convolutional Neural Network, Siamese Convolutional Neural Network, Chest X-Ray, Pandemic, Corona Virus Disease 2019, Pneumonia.

1. INTRODUCTION

The coronavirus COVID-19 pandemic is the defining global health crisis of present time and is the greatest challenge being faced by the world since World War Two. At present, there are 17,43,94,212 confirmed cases of COVID-19 worldwide with 37,52,693 deaths as reported to WHO [1]. The outbreak was declared as Public Health Emergency of International Concern (PHEIC) by the World Health Organization (WHO) on January 30, 2020 [2].

Most of the symptoms of COVID-19 resemble that of the common flu, hence, it becomes challenging to make a clinical decision [3]. The Real-Time RT-PCR (Reverse Transcriptase Polymerase Chain Reaction) test is suitable to detect the COVID-19 virus, but this test produces high false-negative cases [4]. Moreover, deploying such a test on a large scale is very expensive, especially for developing countries like India. The solution to this problem involves using medical imaging techniques including radiography and tomography. Image processing, artificial intelligence and deep learning algorithms can play a pivotal role to assist medical professionals in making more accurate decisions in lesser time [5, 6].

Convolutional Neural Networks (CNN) has gained much popularity due to their ability to solve computer vision problems in a variety of fields ranging from military, education, robotics and medical imaging [7].

One such variant of CNN is a Siamese CNN (SCNN) [8]. This neural network architecture takes two images as inputs and computes a similarity score for the pair. This type of network has previously been used for tasks like object tracking, facial recognition etc. and yields good results in these tasks [9].

In this research, we present an approach to distinguish between both the diseases using X-ray images. We use an SCNN to classify the input images into four classes, namely COVID, Severe-COVID, Pneumonia and Normal, by comparing similarities between a randomly picked pair of images from different classes. The separately introduced Severe-COVID class helps to identify patients which are suffering from other respiratory diseases in addition to COVID-19 and might require immediate medical attention.

We provide a comparative study for the performance of two CNN architectures i.e., VGG-16 [10] and ResNet50 [11]. We further use these CNN architectures as backbone networks of Siamese CNN architecture and present the improvement in performance for the task. The decent accuracy and differentiating ability of Siamese CNN model also enables us to detect newly discovered diseases and classify them using fewer image samples.

2. RELATED WORK

Chest X-ray and computed tomography are one of the most used diagnosing and monitoring technique for diseases affecting the lungs or nearby structures the chest. A survey by Litjens et al. [12] shows how Artificial Intelligence techniques, especially deep learning, have permeated the entire field of medical image analysis over the last decade. Several studies [13], [14], [15] suggest that the medical community can rely on x-rays COVID testing and monitoring the severity of the infection.

In the current COVID-19 pandemic, many studies have shown that deep learning applications can be used as a faster and efficient alternative for diagnosing the virus through x-rays and CT images [16], [17].

There are several studies which separate COVID-19 patients from non-COVID-19 cases. Chen et al. [18] use CT scan images from 51 COVID-19 patients and 55 other patients. They developed a segmentation model with UNet++ [19] architecture. They classified the patients using segmented lesions and obtained an accuracy of 95.2%. Jin et al., [20] also developed a similar model with UNet++ for segmentation combined with ResNet50 for classification on a much larger dataset and achieved a sensitivity of 97.4% and specificity of 92.2%.

Several studies use transfer learning technique for training state-of-the-art image classification models pre-trained on ImageNet for COVID-19 detection.

Most of these techniques either have limited data or imbalanced dataset due to lesser number of organized, labelled COVID-19 patient radiographs. To address this problem, some of the researchers have made use of Siamese networks for training and testing with a limited dataset.

3. METHODOLOGY

3.1. Data Description

Chest X-Ray Images (Pneumonia) dataset contains 5863 chest x-ray (CXR) images from patients at Guangzhou Women and Children's Medical Center, publicly available on Kaggle. Diagnoses of all CXR images have been reviewed by two radiologists. The images for two classes i.e., Pneumonia and Normal CXR were taken from this dataset for our dataset.

COVID-Chest-Xray is an open dataset publicly available on GitHub containing a total of 542 chest x-ray (CXR) and CT images of patients who are positive or suspected of COVID-19 or other viral pneumonia. The metadata.csv contains 759 rows; each row has information regarding an image.

The columns of the CSV uniquely identify an image with patient-id, sex, findings, offset, view of the x-ray etc. We split the data into two parts COVID and Severe-COVID using this metadata.csv.

The final dataset used for this research by merging the above two dataset comprises of four classes: COVID, Normal, Pneumonia and Severe-COVID. The class wise split of the images in the dataset is mentioned in Table 1.

Table 1. Dataset split statistics

Class	Training	Testing
Normal	149	350
Pneumonia	150	350
COVID	130	304
Severe COVID	3	8

To test the performances of our models, we merged two datasets: Chest X-Ray Images (Pneumonia) and COVID-Chest-Xray.

The images consist of CXR from various angles such as frontal view, axial view, flipped frontal view. The data was pre-processed to restore only the frontal view images that were reshaped to a size of 100×100×3 pixels.

3.2. Model Description

3.2.1. Siamese Network

The Siamese CNN [8] described in Figure 1 is used in this research. The network takes two images of CXR as input to two parallel CNNs to generate a fixed-length feature vector for each input image, as shown by $V(x_1)$ and $V(x_2)$. The twin convolution networks share the same weights; thus, they produce similar feature vectors for images belonging to the same class and two different feature vectors for images belonging to different classes. Hence, a layer is added to calculate the element-wise absolute difference between the two feature vectors, and then a similarity score between 0 and 1 is generated by the output sigmoid layer.

3.2.2. Convolution Neural Network

Two CNN architectures that are trained to encode the images are VGG16 [10] and Residual Networks (ResNet) [11] which were originally designed for the ImageNet challenge.

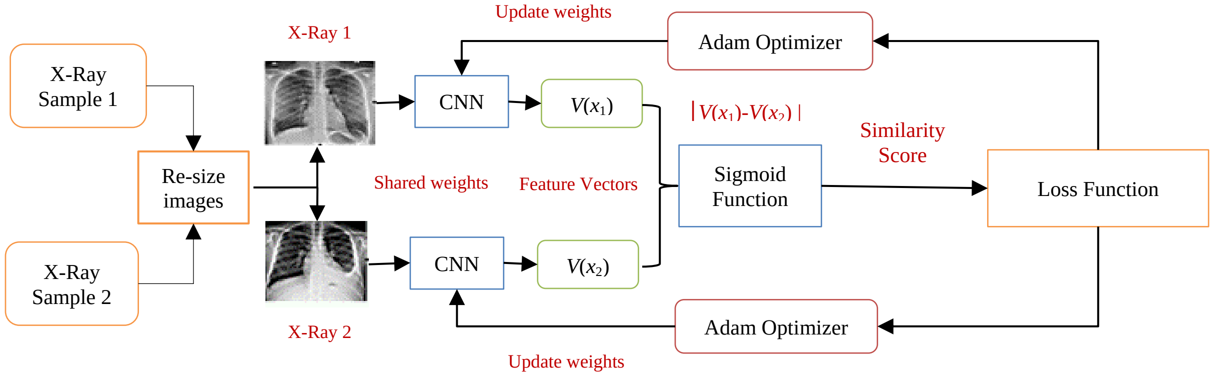

Figure 1. Siamese Convolutional Neural Network (SCNN)

VGG stands for Visual Geometry Group. The VGG-16 architecture consists of 16 layers of blocks of 2D Convolution and Max Pooling layers as depicted in Figure 3. Although, it performs well on the ImageNet challenge and other classification tasks, we found that the network was hard to train and as number of pairs increased network started overfitting and performed poorly on larger batches.

The ResNet-50 architecture which was also used for the CNN units of the Siamese Network distinguished itself from the VGG because of the addition of Batch Normalization and periodic identity shortcut connections that reduces the problem of overfitting and vanishing gradients. The 50 layered architecture was described in [11] and depicted in Figure 2. The top layer was removed that is used for classification to get desired feature vector.

The Siamese network requires a pair of images for input, so a generator method was used to generate batches of pairs of random x-ray images from the dataset for training, validation and testing. Table 2 shows the example of pairs that are to fed into the network.

Table 2. Pairs and their respective labels

Image 1	Image 2	Label
COVID	Normal	0
COVID	Severe COVID	0
COVID	Pneumonia	0
Severe COVID	Pneumonia	0
Severe COVID	Normal	0
Pneumonia	Normal	0
COVID	COVID	1
Normal	Normal	1
Pneumonia	Pneumonia	1
Severe COVID	Severe COVID	1

Label 1 denotes a positive pair, i.e., the pair contains images from the same class while label 0 denotes negative pair, i.e., the images are from different classes. The generator also ensures that each batch contains an equal number of positive and negative pairs.

3.3. Training

In general, it is rare for people to train the entire Convolutional Network from scratch as it is extremely difficult to have a large enough dataset. A much more common practice is to use a CNN that is trained on a very large dataset and then use this pre-trained CNN to extract features for the task. Hence, in order to reduce the training time and to get better performance, we used weights from pre-trained VGG-16 and ResNet50 models. ImageNet dataset offers a broad spectrum of objects that are not fit for this research. Thus, the models were not directly trained on the ImageNet dataset weights, rather we fine-tuned the models using our CXR dataset.

We used VGG-16 and ResNet50 models and added a multilayer perceptron at the end to create a simple image classifier with two dense layers for the four classes used in this research. To further train the twin CNN architecture for our Siamese network, we loaded it with pre-trained weights from the above process.

We used the Siamese network with both VGG-16 and ResNet50 models as the backbone network and found that these weights not only produced better results but also helped in faster training of the Siamese network (Figure 3).

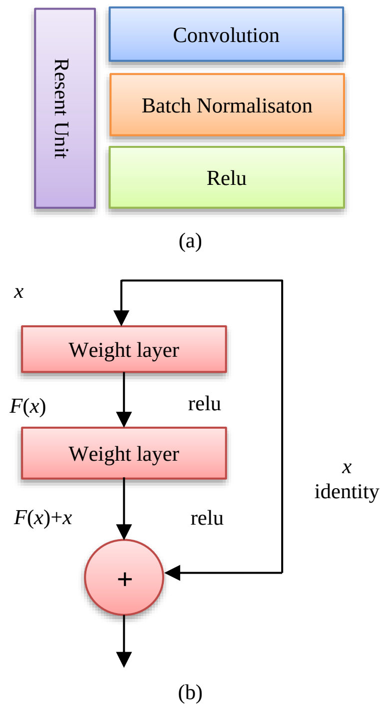

Figure 2. (a) Overview of a single ResNet unit, (b) Skip connections in the ResNet architecture [22]

3.4. Testing

For the testing process, we used the N-way one-shot learning validation process. In each task of N-way evaluation, the method first randomly chooses a test image x from the test dataset, and then randomly selects N different images from different classes in the dataset to form the support set $X = \{x_c\}$ with $c = 1$ to N, where one of the images is from the same class as the test image. The method then calculates the similarity score for the pair of the test image and each of the image in the support set, then predict the positive pair or the correct prediction using:

$$C = \arg\max(preds) \quad (1)$$

where, $preds$ is the list of predictions for N pairs and C is the position of the correct pair.

Now, this process is repeated for K number of trials and for P correct predictions, the accuracy of the model can be calculated by the formula:

$$Accuracy = \left(\frac{P}{K}\right) \times 100 \quad (2)$$

4. EVALUATION METRICS AND RESULTS

We evaluated the Siamese VGG and Siamese ResNet models using N-way One-shot Evaluation, which is the best evaluation approach for these models for classification task. Therefore, we conducted 4-way one-shot evaluation, in which, for each image in the testing dataset, we paired it with an image from each class in the training set. From this batch of 4-pairs, we generated the similarity score for each pair. The pair with the highest similarity score is the predicted class for the test image.

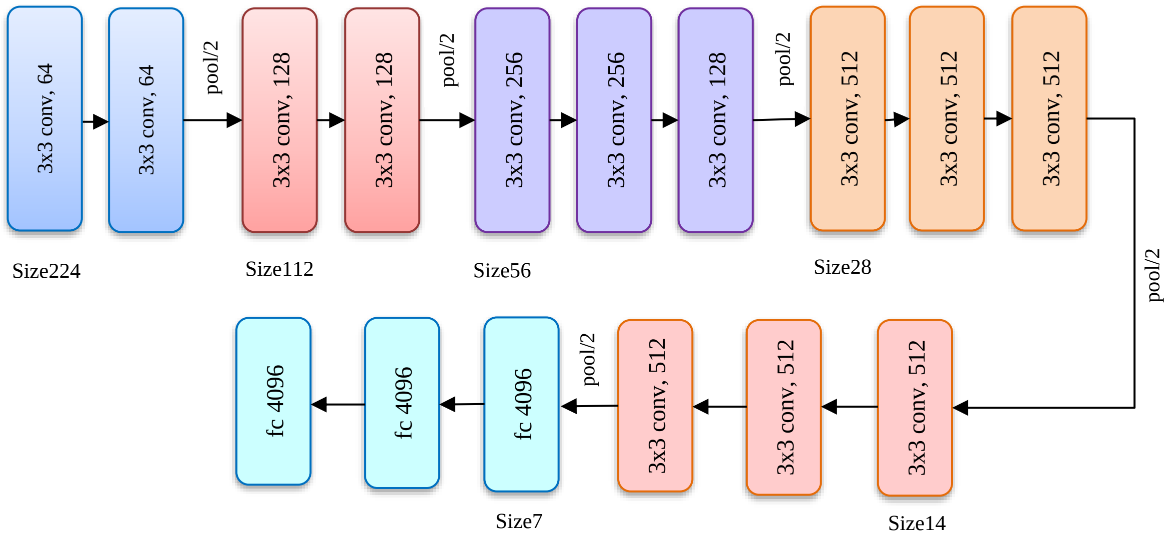

Figure 3. VGG architecture [21]

Along with these Siamese models, we also compared the performance of the pre-trained state-of-the-art CNN models i.e., VGG-16 and ResNet50 classifiers using the same dataset. Figure 4 depicts the graphs showing accuracy and loss of all the models. Table 3 shows the performance comparisons of the four models using the classification results.

Tables 4 and 5 show the performance of the four models for each of the four classes including Normal, Pneumonia, COVID, and Severe COVID. The performance comparisons clearly show that the Siamese network performs much better than the state-of-the-art models, and Siamese ResNet50 gives the best result with an accuracy of 98%.

5. CONCLUSION AND FUTURE SCOPE

This research focuses on using chest X-ray images to classify them into four classes: COVID, Pneumonia, Severe-COVID and Normal. The classification is performed using a Siamese Convolutional Neural Network architecture. The performance of Siamese CNNs is compared against two pre-trained conventional CNN architectures viz. the VGG-16 and ResNet50 models.

These architectures are used as the backbone models for the parallel CNNs of the Siamese network. We evaluate the performance of the four models on our dataset created by merging two publicly available datasets. An N-way one-shot evaluation criteria is used to validate the performance of the Siamese networks and finally all the models are compared against their precision, recall and accuracy.

Siamese ResNet50 gives an accuracy of 98% which is the best performance in contrast to 95% on VGG-16 model, 93% on ResNet50 model and 96% on Siamese VGG-16 model. This work can further be extended to testing the model to quantitatively determine the severity of COVID lung disease as it will help to treat the critical patients appropriately.

Table 3. Classification results

Model	Accuracy	Precision	Recall
VGG-16 Classifier	0.95	0.95	0.95
ResNet50 Classifier	0.93	0.97	0.96
Siamese VGG-16	0.96	0.96	0.96
Siamese ResNet50	0.98	0.98	0.98

Table 4. Classification report of all models

Model	Label	Precision	Recall	F1 - score	Support
VGG-16	Pneumonia	1.00	0.89	0.094	165
	Severe_COVID	0.00	0.00	0.00	4
	COVID19	0.89	1.00	0.94	144
	Normal	0.97	1.00	0.99	165
ResNet50	Pneumonia	0.97	0.83	0.90	165
	Severe_COVID	0.00	0.00	0.00	4
	COVID19	0.76	0.99	0.86	144
	Normal	0.98	0.90	0.94	165
Siamese VGG16	Pneumonia	0.93	0.97	0.95	350
	Severe_COVID	0.55	0.75	0.63	8
	COVID19	0.96	0.99	0.98	304
	Normal	1.00	0.92	0.96	350
Siamese ResNet50	Pneumonia	0.96	1.00	0.98	350
	Severe_COVID	0.57	1.00	0.73	8
	COVID19	0.99	0.98	0.99	304
	Normal	1.00	0.96	0.98	350

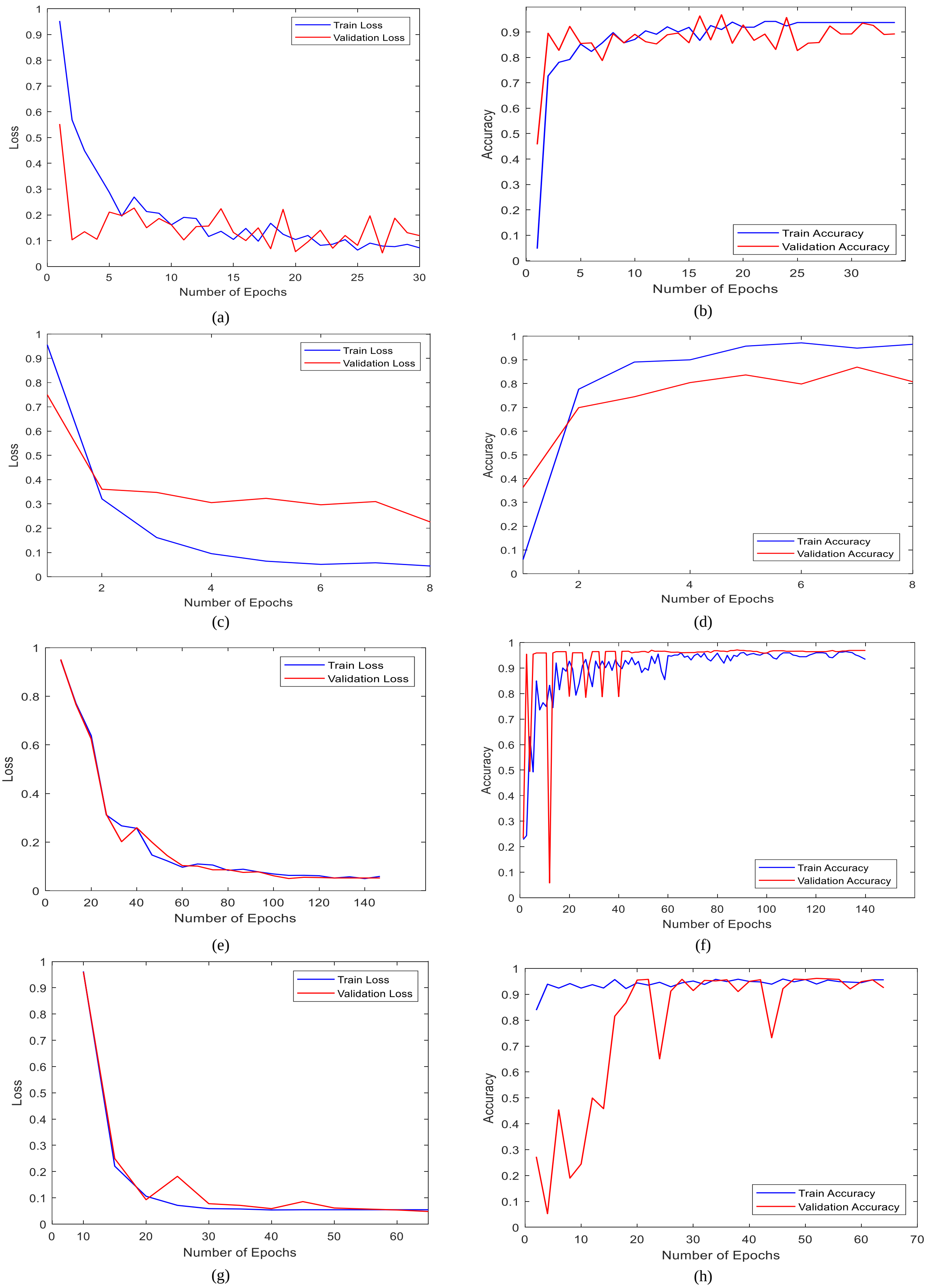

Figure 4. Training and validation loss: (a, b) VGG-16 model, (c, d) Siamese VGG-16 model, (e, f) ResNet50 model (g, h) Siamese ResNet50 model

Table 5. Confusion matrix of all model

Model	Label	Pneumonia	Severe_COVID	COVID19	Normal
VGG-16	Pneumonia	147	0	13	5
	Severe_COVID	0	0	4	0
	COVID19	0	0	144	0
	Normal	0	0	0	165
ResNet50	Pneumonia	140	0	24	2
	Severe_COVID	0	0	4	0
	COVID19	3	3	138	0
	Normal	2	0	12	151
Siamese VGG16	Pneumonia	341	2	7	0
	Severe_COVID	0	6	2	0
	COVID19	0	3	301	0
	Normal	25	0	2	323
Siamese ResNet50	Pneumonia	350	0	0	0
	Severe_COVID	0	8	0	0
	COVID19	1	6	297	0
	Normal	13	0	2	335

REFERENCES

[1] "Who Coronavirus Disease (COVID-19) Dashboard", 2020, https://covid19.who.int/.

[2] "Coronavirus Disease (COVID-19)-World Health Organization", 2020, www.who.int/emergencies/diseases/novel-coronavirus2019.

[3] "Coronavirus", 2020, www.who.int/healthtopics/coronavirus.

[4] B. Udugama, P. Kadhiresan, H.N. Kozlowski, A. Malekjahani, M. Osborne, V.Y.C. Li, H. Chen, S. Mubareka, J.B. Gubbay, W.C.W. Chan, "Diagnosing COVID-19: the Disease and Tools for Detection", American Chemical Society Nano, Vol. 14, No. 4, pp. 3822-3835, April 2020.

[5] M. Mbida, A. Ezzati, "Artificial Intelligence Auscultation System for Physiological Diseases", International Journal on Technical and Physical Problems of Engineering (IJTPE), Issue 49, Vol. 13, No. 4, pp. 97-103, December 2021.

[6] E.H. Hasayni, M. Ettaounil, "Generalization Ability Augmentation and Regularization of Deep Convolutional Neural Networks Using $l^{1/2}$ Pooling", International Journal on Technical and Physical Problems of Engineering (IJTPE), Issue 48, Vol. 13, No. 3, pp. 1-6, September 2021.

[7] L.C. Yann, P. Haffner, L. Bottou, Y. Bengio, "Object Recognition with Gradient-Based Learning", Shape, Contour and Grouping in Computer Vision, pp. 319-345. Springer, Berlin, Heidelberg, 1999.

[8] J. Bromley, J.W. Bentz, L. Bottou, I. Guyon, Y. LeCun, C. Moore, E. Sackinger, R. Shah, "Signature: Verification Using a "Siamese" Time Delay Neural Network", International Journal of Pattern Recognition and Artificial Intelligence, Vol. 7, No. 4, pp. 669-688, August 1993.

[9] F. Schroff, D. Kalenichenko, J. Philbin, "Facenet: A Unified Embedding for Face Recognition and Clustering", IEEE Conference on Computer Vision and Pattern Recognition, pp. 815-823. 2015.

[10] S. Karen, A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition", International Conference on Learning Representations, 2015.

[11] H. Kaiming, X. Zhang, S. Ren, J. Sun, "Deep Residual Learning for Image Recognition", IEEE Conference on Computer Vision and Pattern Recognition, pp. 770-778, 2016.

[12] L. Geert, T. Kooi, B.E. Bejnordi, A.A.A. Setio, F. Ciompi, M. Ghafoorian, J.A.V.D. Laak, B.V. Ginneken, C. I. Sanchez, "A Survey on Deep Learning in Medical Image Analysis", Medical Image Analysis, Vol. 42, pp. 60-88, December 2017.

[13] H.Y.F. Wong, H.Y.S. Lam, A.H.T. Fong, S.T. Leung, T.W.Y. Chin, C.S.Y. Lo, M.M.S. Lui, "Frequency and Distribution of Chest Radiographic Findings in COVID-19 Positive Patients", Radiology, Vol. 296, No. 2, August 2020.

[14] A. Jacobi, M. Chung, A. Bernheim, C. Eber, "Portable Chest X-Ray in Coronavirus Disease-19 (COVID-19): A Pictorial Review", Clinical Imaging, Vol. 64, pp. 35-42, April 2020.

[15] A. Borghesi, R. Maroldi, "COVID-19 Outbreak in Italy: Experimental Chest X-Ray Scoring System for Quantifying and Monitoring Disease Progression", Medical Radiology, Vol. 125, No. 5, pp. 509-513, May 2020.

[16] J. Bullock, K.H. Pham, C.S.N. Lam, M. Luengo-Oroz, "Mapping the Landscape of Artificial Intelligence Applications Against COVID-19", Journal of Artificial Intelligence Research, Vol. 69, pp. 807-845, November 2020.

[17] K. Murphy, H. Smits, A.J.G. Knoops, M.B.J.M. Korst, T. Samson, E.T. Scholten, S. Schalekamp, "COVID-19 on the Chest Radiograph: A Multi-Reader Evaluation of an AI System", Radiology, Vol. 296, No. 3, pp. 166-172, May 2020.

[18] J. Chen, L. Wu, J. Zhang, L. Zhang, D. Gong, Y. Zhao, S. Hu, "Deep Learning-Based Model for Detecting 2019 Novel Coronavirus Pneumonia on High-Resolution Computed Tomography: A Prospective Study", MedRxiv November 2020.

[19] Z. Zongwei, M.M.R. Siddiquee, N. Tajbakhsh, J. Liang, "Unet++: A Nested U-Net Architecture for Medical Image Segmentation", Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support, pp. 3-11. Springer, Cham, 2018.

[20] S. Jin, B. Wang, H. Xu, C. Luo, L. Wei, W. Zhao, X. Hou, "AI-Assisted CT Imaging Analysis for COVID-19

Screening: Building and Deploying a Medical AI System in Four Weeks", Applied Soft Computing, November 2020.
[21] VGG Architecture, www.kaggle.com/code/blurredmachine/vggnet-16-architecture-a-complete-guide/notebook.
[22] J. Yang, H. Wang, K. Guo, Kexiang, "Natural Language Word Prediction Model Based on Multi-Window Convolution and Residual Network", IEEE Access, Vol. 8, pp. 188036-188043, 2020.

BIOGRAPHIES

Vanita Jain was born in New Delhi, India in 1966. She received her B.E. degree in Electrical Eng. in 1984, M.Tech. degree in Controls in 1990 and Ph.D. Degree from V.J.T.I., Mumbai, India. She is the first woman to have completed Ph.D. in Technology from Mumbai University. She is having more than 30 years of teaching experience, and taught students at both UG and PG level. Currently, she is working as Professor and Head of Information Technology Department at Bharati Vidyapeeth's College of Engineering, New Delhi, India since 2010. Her fields of interest include soft computing, control, optimization techniques and system engineering and is actively involved in the research in these areas.

Apoorv Jain was born in New Delhi, India in 1999. He received his B.Tech. degree in the field of Information Technology from Bharati Vidyapeeth's College of Engineering, New Delhi, India. He is currently working as a software developer in Thoughtworks. His areas of research are Machine Learning, Soft Computing.

Vinayak Garg was born in New Delhi, India in 1999. He received his B.Tech. degree in the field of Information Technology from Bharati Vidyapeeth's College of Engineering, New Delhi, India in 2021. He is currently pursuing his MTech. degree in the field of Computer Science Engineering from the International Institute of Information Technology, Bangalore (IIITB), Karnataka, India. His areas of research are machine learning, and deep learning.

Achin Jain was born in Delhi, India, in 1984. He received his B.Tech. and M.Tech. degrees in Information Technology both in Information Security from GGSIPU, Delhi, India in 2007 and 2013, respectively. He is currently a Ph.D. student of University School of Information, Communication and Technology ate the same university and works at Bharati Vidyapeeth's College of Engineering, New Delhi as Assistant Professor since 2014. His research interests are sentiment classification and machine learning using NLP techniques. He has published more than 15 articles in peer reviewed journals and presented work in more than 7 international conferences.

Merve Demirci was born in Trabzon, Turkey in 1992. She graduated from Departmen of Electrical and Electronics Engineering, Ataturk University, Erzurum, Turkey in 2014 and working as a Research Assistant at Gazi University, Ankara, Turkey. She received the degrees of M.Sc. degree and currently a student of Ph.D. degree in Electrical and Electronics Engineering Department, Gazi University. Her main research interests and subjects are power systems analysis, artificial intelligence.

M. Cengiz Taplamacioglu was born in Ankara, Turkey in 1962. He graduated from Department of Electrical and Electronics Engineering, Gazi University, Ankara, Turkey. He received the M.Sc. degrees in Industrial Engineering from Gazi University and also in Electrical and Electronics Engineering from Middle East Technical University, Ankara, Turkey. He received his Ph.D. degree in Electrical, Electronics and System Engineering from University of Wales, Cardiff, UK. He is a Professor of the Electrical and Electronics Engineering since 2000. His research interests and subjects are high voltage engineering, corona discharge and modelling, electrical field computation, measurement and modelling techniques, optical HV measurement techniques, power systems control and protection, lighting techniques, renewable energy systems and smart grid applications.