2020 IEEE International Women in Engineering (WIE) Conference on Electrical and Computer Engineering (WIECON-ECE)

A Deep Transfer Learning Approach to Diagnose Covid-19 using X-ray Images

Nagifa Ilma Progga
Department of Computer Science and Engineering
University of Chittagong
Chittagong, Bangladesh
ilmaprogga1996@gmail.com

Mohammad Shahadat Hossain
Department of Computer Science and Engineering
University of Chittagong
Chittagong, Bangladesh
hossain_ms@cu.ac.bd

Karl Andersson
Department of Computer Science, Electrical and Space Engineering
Lule University of Technology
Skellefte, Sweden
karl.andersson@ltu.se

***Abstract*—The Covid-19 disease which was caused by novel coronavirus (SARS-CoV-2) has already become a great threat for humans beings. The virus is spreading rapidly around the world. Therefore, we crucially need quick diagnostic tests to identify affected patients and to minimize the spread of the virus. With the advancements of Machine Learning, the detection of Covid-19 in the early stage would facilitate taking precautions as early as possible. However, because of the lack of data-sets, especially chest X-ray images of Covid-19 affected patients, it has become challenging to detect this disease. In this paper, a deep transfer learning-based pre-trained model is named VGG16 along with adapt histogram equalization has been developed to diagnose Covid-19 by using X-ray images. An image processing technique named adaptive histogram equalization has been used to generate more images by using the existing data set. It can be observed that VGG-16 provides the highest accuracy which is 98.75% in comparison to two other pre-trained models such as VGG-19 and Mobilnenet-V2(97% accuracy for VGG-19, 92.65% accuracy for Mobilenet-V2).**

***Keywords*—Adapt Histogram Equalization, Coronavirus, Chest X-ray, Convolutional Neural Network, Transfer Learning.**

I. INTRODUCTION

The first case of pneumonia from an unknown origin was noticed in Wuhan, Hubei Province, China in the year 2019 [1]. Later, it has resulted in an epidemic of Covid-19 which spread around the world through the virus SARS-CoV-2. The total number of reported cases of Covid-19 patients is 49,667,976, the number of recovered patients is 35,254,953, and the total reported death is 15248,785 to date (7/11/2020) [2]. It is presumed that Covid-19 is contaminated from bats to humans [3]. The intermediate source of origin of SARS-CoV-2 and its transmission to humans is still unknown. However, the rapid human to human transmission has been confirmed widely. No vaccine or medicine has not been clinically approved to be used to cure Covid-19. Since most of the Covid-19 patients show pneumonia symptoms, chest X-ray has become vital for the detection of this disease.
In this work, the use of transfer learning is proposed for detecting three different classes of chest X-ray images: Covid-19, Viral Pneumonial, and Normal from a given input image. This can be accomplished by training pre-trained models on our data-set. The significant contributes of this research work are:

- Performance enhancement is achieved by using adapt histogram equalization on the chest x-ray data-set. Adapt histogram equalization improves the intensity of the images and they are fed to models for better accuracy.
- Performance comparison is done between three well-known pre-trained models namely VGG16, MobileNetV2, and VGG19.

The next sections are arranged as accordingly: Section II contains discussion about related works on Covid-19 detection, Section III briefly describes transfer learning, Section IV describes data collection, Section V gives an overview of the methodology of this research, Section VI is about system implementation, Section VII analyses the result, Section VIII is about how can we improve this research in future.

II. RELATED WORK

Since Covid-19 is a very new field for researchers, there haven't been a lot of works on this topic.
In [4], they have used GAN(Generative Adversarial Network) for image preprocessing, and transfer learning is used for the training, testing, and validation phase. Pre-trained models Alexnet, Googlenet [5], Resnet18 are used for the transfer learning phase. They have achieved the highest 80.6% accuracy using Googlenet for four classes, 85.2% accuracy using

978-0-7381-3144-3/20/$31.00

Alexnet for classifying three classes, 99.9% accuracy using Googlenet for two classes(covid-19 and normal). However, the recognition rate for each class especially for three and four classes classification varies significantly.

Narin et al. [6] presented a neural network for two classes (Covid-19,normal) using transfer learning. ResNet-50 model has achieved the highest recognition rate of 98% accuracy. The other two proposed models InceptionV3 and Inception-ResNetV2 have achieved 97% and 87% accuracy accordingly. The model is trained using only 100 images, 50 for each class.

In [7] they proposed a series of Xception [8] and ResNet50V2 [9] networks as their deep learning model. They have obtained overall 91.4% accuracy by using the features extracted by Xception and ResNet50V2 networks.

Asmaa et al. [10] proposed a classification model constructed using convolutional neural network, called Decompose, Transfer(transfer learning), and Compose (DeTraC) for the diagnosis of the disease COVID-19 from SARS(severe acute respiratory syndrome) & normal cases. They have achieved the highest 95.21% accuracy with a dataset consisting of 1764 sample images.

In [11] they proposed a system that is based on parallel-dilated CNN, in order to diagnose COVID-19 affected patients using chest X-ray images. They have used a benchmark dataset [12] for their training. They have achieved the highest 96.58% accuracy.

III. Transfer Learning

CNN(Convolutional Neural Network) is a deep learning algorithm that is able to differentiate between images by learning their characteristics [13] [14] [15] [16]. CNN usually is used for multiple tasks such as image analysis and classification, image and video recognition, etc. [17] [18]

Deep learning is used successfully for analyzing medical images e.g MRI(Magnetic Resonance Imaging), chest X-ray [11]. For analyzing medical images, a limited number of data-set is available for the researchers. A large amount of data is needed for the better performance of the deep neural network model. This problem can be solved using deep transfer learning. Transfer learning allows us to train our model with a limited number of data-set. In this method, the knowledge of the pre-trained models gained from a large data-set can be transferred to the CNN model while training. Even though pre-trained models are designed for different types of images, we can build our model by training it on our new data-set so that it can learn new features. By combining new and previously learned features, transfer learning can be used to classify unknown images efficiently. In our work, we have used transfer learning based three well-known pre-trained models.

- VGG-16
- MobileNetV2
- VGG-19

VGG-16 architecture is build by using 13 convolutional layers and 3 fully connected layers [19]. VGG-16 is trained on ImageNet dataset which consists of over 14 millions images from over 20000 categories [6]. Whereas, VGG-19 architecture is build by using 19 convolutional layers and 3 fully connected layers [19]. It is also trained on ImageNet dataset. MobileNetV2 consists of residual block and another block for downsizing. Each type of blocks is consists of three layers [20].

IV. Dataset

A. Data Collection

We have used the benchmark data-set for our research experiment.

- Covid19-radiography-database consists of 219 sample images of Covid-19, 1345 sample images of viral pneumonia, and 1341 normal chest X-ray images. [12]. Among them, we have used 219 sample images of Covid-19, 400 sample images of viral pneumonia & 400 normal chest X-ray images for our dataset in order to avoid an unbalanced dataset.

B. Adapt Histogram Equalization

Adapt Histogram Equalization is a technique of creating a high contrast image from a low contrast image. This is accomplished by operating on small regions in an image. It enhances the contrast of each small regions in the image [21]. Improving contrast enables convolutional neural network to explore the dataset in a superior way. The result is striking for gray-scale images. The performance and accuracy of CNN highly depend on the size of the dataset. So to increase the size of the dataset and to enhance the model performance, data augmentation can be used [22] [23]. In our work, we have applied adapt histogram equalization on each image using Matlab [24] to generate a new image of better contrast. This process resulted in a total of 2038 sample images.

Fig. 1: Data Preprocessing

V. METHODOLOGY

We have used transfer learning based pre-trained models VGG-16, MobileNetV2, VGG-19 for classifying X-ray images into three different classes: Covid-19, Normal and Viral Pneumonia. Figure 2 shows the schematic representation of our model. After loading a pre-trained model, the fully connected layers at the end of the model are removed. Then those fully connected layers are replaced with freshly initialized dense layers. Earlier convolution layers are frozen and not removed so that any strong features learned by the neural network previously are not destroyed. The softmax function is used as activation function.

$$Softmax(x) = \frac{e^j}{\sum_i e^i} \tag{1}$$

We have trained this new model on our new data-set. Algorithm 1 shows the steps we followed to build our system.

Algorithm 1: Proposed Algorithm

```
Input: Image
Output: Image with predicted class
If (model is not trained) then
    load data-set
    apply histogram equalization for data augmentation
    train(90%),test(10%)
    load pre-trained model
    remove fully connected nodes at the end
    replace them with freshly intialized dense layers
    train the model
    save trained model
else
    load trained model
test model with test data-set
plot graphs and confusion matrix
```

Performance Metrics: In our study, the performances of the pre-trained models are evaluated based on four criteria. These are:

$$Accuracy = \frac{TP(True\,Positive) + TN(True\,Negative)}{TP(True\,Positive) + FP(False\,Positive) + TN(True\,Negative) + FN(False\,Negative)} \tag{2}$$

$$Recall = \frac{TP(True\,Positive)}{TP(True\,Positive) + FN(False\,Negative)} \tag{3}$$

$$Precision = \frac{TP(True\,Positive)}{TP(True\,Positive) + FP(False\,Positive)} \tag{4}$$

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision\, +\, Recall} \tag{5}$$

The proportion of the samples that are predicted correctly since the class of interest is true positive. The proportion of the samples that are not predicted correctly since the class of interest is true negative. False positive is the proportion of the samples that are mislabeled as a class of interest. False negative is the proportion of the samples that are mislabeled as not the class of interest. These are calculated for each class by comparing its label to the remaining labels. Precision calculates that if our system predicts ”yes he is covid-19 positive”, how often is it predicted correctly. Recall indicates how many samples that are predicted as covid-19 positive, are covid positive. F1-score is the average of Recall and Precision [25].

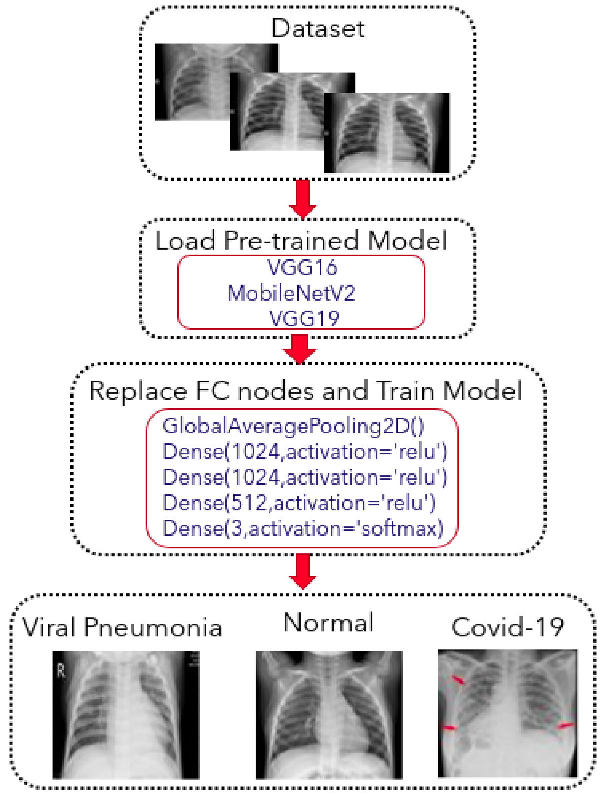

Fig. 2: Schematic representation of our system

VI. SYSTEM IMPLEMENTATION

The system was implemented using the Python programming language is used for system implementation. The libraries used in this system are Matlab, Matplotlib, Numpy, Keras, PIL. As the system back-end, Tensorflow was used. Built-in functions for the system such as optimizers, activation functions, layers, etc are provided by Keras [26]. Matlab is used for applying adapt histogram equalization on images. A system is built to classify images in real-time. The user needs to enter the path of his desired image and the system will classify it in real-time. Some sample screenshots of real-time validations are shown in figure 3. Some chest X-ray images collected from local hospitals have been provided as input to show the system's ability to predict any real-life image.

VII. RESULT AND DISCUSSION

In our research work, X-ray images are used as dataset for the automatic detection of Covid-19 positive patients. Pre-trained models such as VGG16, VGG19, and MobileNetV2 are trained and tested on this dataset. All the pre-rained models are trained up to 20 epochs for avoiding overfitting. Training accuracy values for each model are shown in figure 4. It can be seen from figure 4 that VGG16 has obtained the highest

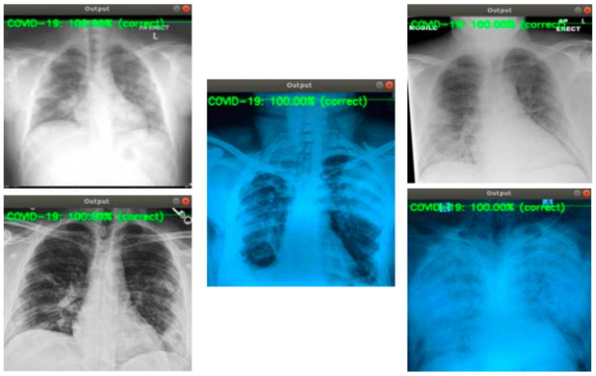

Fig. 3: Real Time Validation

training accuracy. VGG19 and MobileNetV2 models have similar performance. However, VGG16 has a faster training process than VGG19 and MobileNetV2.
In figure 4, the y-axis shows the recognition rate and the x-

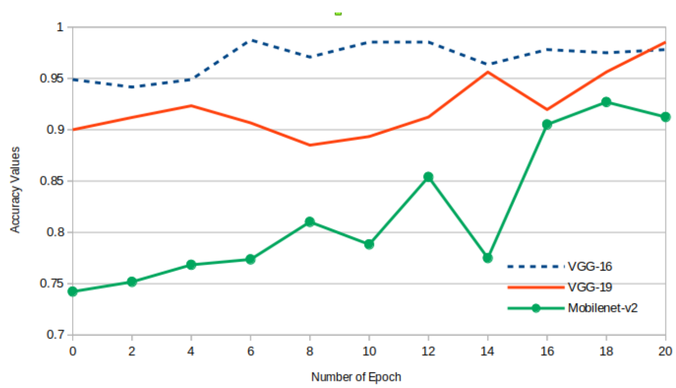

Fig. 4: Comparison between VGG-16, VGG-19 and MobileNet-v2

axis shows the epoch numbers. We have achieved satisfactory accuracy after only 20 epochs. Adapt histogram equalization has played a major role in boosting our system's accuracy.

A. Analysis of Results

The comparison of the performance of the pre-trained models, SVM [27] and Random forest [28] methods are given in Table I. Pre-trained models VGG16 and VGG19 have obtained the highest recall, precision, and f1-score value on average.
In figure 5 and figure 6 the confusion matrix and training loss/accuracy curve for MobileNetV2 are shown respectively. MobileNetV2 has the highest accuracy for Viral Pneumonia class. But it has obtained the lowest 82% accuracy for the normal class.

In figure 7 the confusion matrix and training loss/accuracy curve for VGG19 are shown respectively. VGG19 has obtained the best accuracy for the class Covid-19. It has the lowest accuracy for Normal chest X-rays.

TABLE I: Parameters for the proposed model

		Precision	Recall	F1-score
VGG-19	Covid-19	0.96	0.98	1.00
	Normal	0.98	0.95	0.97
	Viral Pneumonia	0.98	0.98	0.98
	Macro Average	0.98	0.98	0.98
		Precision	**Recall**	**F1-score**
Mobilenet-V2	Covid-19	1.00	0.96	0.98
	Normal	1.00	0.82	0.90
	Viral Pneumonia	0.82	1.00	0.90
	Macro Average	0.92	0.92	0.92
		Precision	**Recall**	**F1-score**
VGG-16	Covid-19	0.98	1.00	0.99
	Normal	0.97	0.97	0.97
	Viral Pneumonia	0.98	0.97	0.97
	Macro Average	0.98	0.98	0.98
		Precision	**Recall**	**F1-score**
Random Forest	Covid-19	0.89	0.90	0.90
	Normal	0.93	0.92	0.93
	Viral Pneumonia	0.95	0.96	0.95
	Macro Average	0.93	0.93	0.93
		Precision	**Recall**	**F1-score**
SVM	Covid-19	0.96	0.96	0.96
	Normal	0.96	0.98	0.97
	Viral Pneumonia	0.98	0.97	0.98
	Macro Average	0.97	0.97	0.97

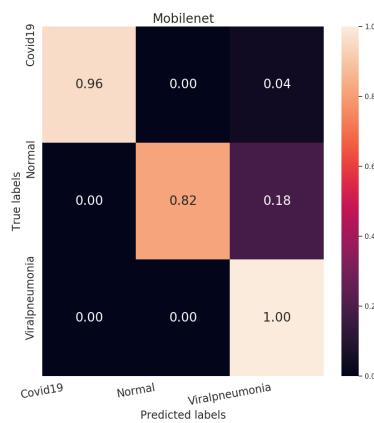

Fig. 5: The confusion matrix for MobileNetV2

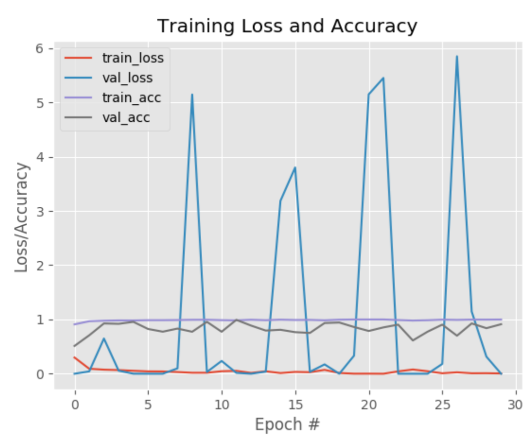

Fig. 6: The Loss/Accuracy curve for MobileNetV2

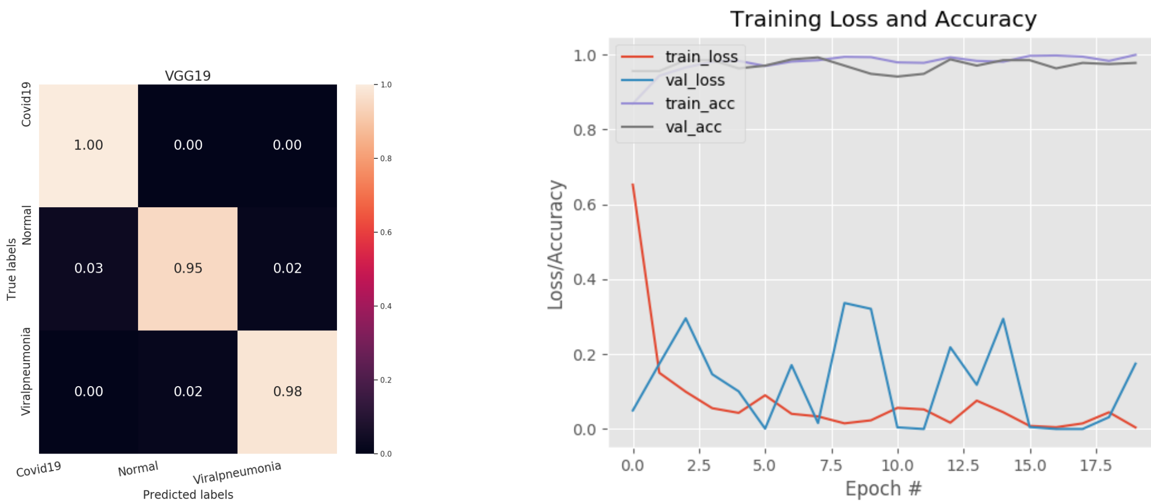

Fig. 7: The confusion matrix & Loss/Accuracy curve for VGG19

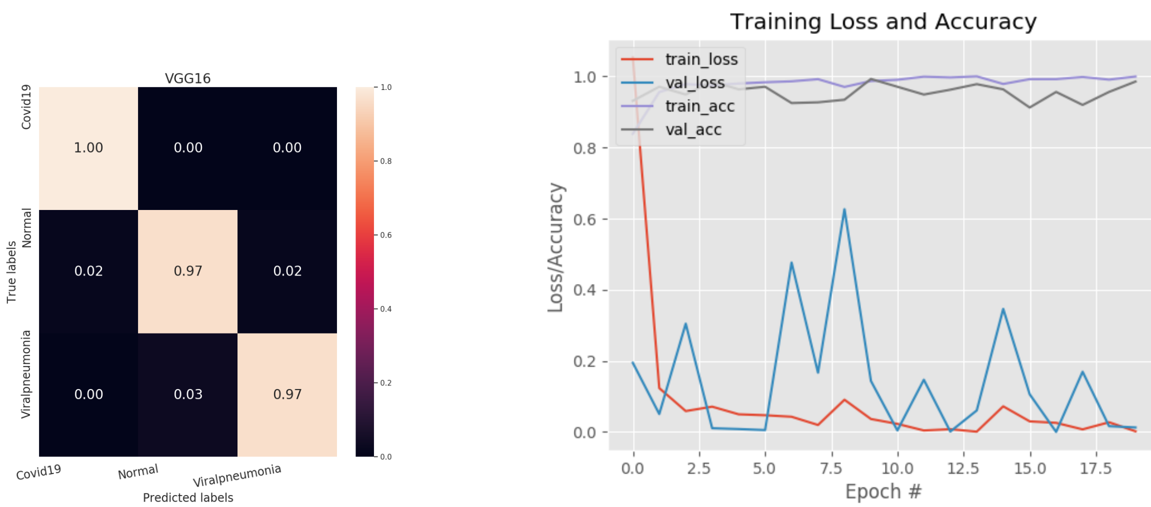

Fig. 8: The confusion matrix & Loss/Accuracy curve for VGG16

In figure 8 the confusion matrix and training loss/accuracy curve for VGG16 are shown respectively. VGG16 has the highest accuracy for the class Covid-19 class and 97% accuracy for both normal and viral pneumonia class. We can observe from the loss/accuracy curve in figure 8 that the training & validation accuracy has gradually increased.

We can see that the loss for MobileNetV2 model has been fluctuating through the epochs. But for other models the loss is more stable throughout the epochs. By using other loss function for training we can solve this problem.

VIII. CONCLUSION AND FUTURE WORK

The result we got from our experiments looks promising. CNN(convolutional neural network) with adapt histogram equalization [29] has achieved more efficiency compared to other machine learning approaches in the case of image processing.

In addition, by using a framework such as Belief Rule Based Expert Systems (BRBES) more robust system can be built [30] [31] [32]. Even though the result looks encouraging, there is room for improvement in some areas, e.g:

- Experiment with a bigger data-set by adding more data in each category and see if there is any improvement. We believe it would increase accuracy.
- Experiment by building a system using two or more network parallelly and explore the result.

REFERENCES

[1] K. Roosa, Y. Lee, R. Luo, A. Kirpich, R. Rothenberg, J. Hyman, P. Yan, and G. Chowell, "Real-time forecasts of the covid-19 epidemic in china from february 5th to february 24th, 2020," *Infectious Disease Modelling*, vol. 5, pp. 256–263, 2020.

[2] Worldometers.info. (2020) The World Meter website. [Online]. Available: https://www.worldometers.info/coronavirus/

[3] M. A. Shereen, S. Khan, A. Kazmi, N. Bashir, and R. Siddique, "Covid-19 infection: origin, transmission, and characteristics of human coronaviruses," *Journal of Advanced Research*, 2020.

[4] M. Loey, F. Smarandache, and N. E. M. Khalifa, "Within the lack of chest covid-19 x-ray dataset: A novel detection model based on gan and deep transfer learning," 2020.

[5] Z. Zhong, L. Jin, and Z. Xie, "High performance offline handwritten chinese character recognition using googlenet and directional feature maps," in *2015 13th International Conference on Document Analysis and Recognition (ICDAR)*. IEEE, 2015, pp. 846–850.

[6] A. Narin, C. Kaya, and Z. Pamuk, "Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks," *arXiv preprint arXiv:2003.10849*, 2020.

[7] M. Rahimzadeh and A. Attar, "A modified deep convolutional neural network for detecting covid-19 and pneumonia from chest x-ray images based on the concatenation of xception and resnet50v2," *Informatics in Medicine Unlocked*, p. 100360, 2020.

[8] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1251–1258.

[9] K. He, X. Zhang, S. Ren, and J. Sun, "Identity mappings in deep residual networks," in *European conference on computer vision*. Springer, 2016, pp. 630–645.

[10] A. Abbas, M. M. Abdelsamea, and M. M. Gaber, "Classification of covid-19 in chest x-ray images using detrac deep convolutional neural network," *arXiv preprint arXiv:2003.13815*, 2020.

[11] N. K. Chowdhury, M. M. Rahman, and M. A. Kabir, "Pdcovidnet: a parallel-dilated convolutional neural network architecture for detecting covid-19 from chest x-ray images," *Health information science and systems*, vol. 8, no. 1, pp. 1–14, 2020.

[12] M. E. Chowdhury, T. Rahman, A. Khandakar, R. Mazhar, M. A. Kadir, Z. B. Mahbub, K. R. Islam, M. S. Khan, A. Iqbal, N. Al-Emadi *et al.*, "Can ai help in screening viral and covid-19 pneumonia?" *arXiv preprint arXiv:2003.13145*, 2020.

[13] Q. Li, W. Cai, X. Wang, Y. Zhou, D. D. Feng, and M. Chen, "Medical image classification with convolutional neural network," in *2014 13th International Conference on Control Automation Robotics & Vision (ICARCV)*. IEEE, 2014, pp. 844–848.

[14] M. Z. Abedin, A. C. Nath, P. Dhar, K. Deb, and M. S. Hossain, "License plate recognition system based on contour properties and deep learning model," in *2017 IEEE Region 10 Humanitarian Technology Conference (R10-HTC)*. IEEE, 2017, pp. 590–593.

[15] M. Z. Islam, M. S. Hossain, R. ul Islam, and K. Andersson, "Static hand gesture recognition using convolutional neural network with data augmentation," in *2019 Joint 8th International Conference on Informatics, Electronics & Vision (ICIEV) and 2019 3rd International Conference on Imaging, Vision & Pattern Recognition (icIVPR)*. IEEE, 2019, pp. 324–329.

[16] R. R. Chowdhury, M. S. Hossain, R. ul Islam, K. Andersson, and S. Hossain, "Bangla handwritten character recognition using convolutional neural network with data augmentation," in *2019 Joint 8th International Conference on Informatics, Electronics & Vision (ICIEV) and 2019 3rd International Conference on Imaging, Vision & Pattern Recognition (icIVPR)*. IEEE, 2019, pp. 318–323.

[17] T. U. Ahmed, M. S. Hossain, M. J. Alam, and K. Andersson, "An integrated cnn-rnn framework to assess road crack," in *2019 22nd International Conference on Computer and Information Technology (ICCIT)*. IEEE, 2019, pp. 1–6.

[18] M. S. Hossain, Z. Sultana, L. Nahar, and K. Andersson, "An intelligent system to diagnose chikungunya under uncertainty," *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications*, vol. 10, no. 2, pp. 37–54, 2019.

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

[20] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "Mobilenetv2: Inverted residuals and linear bottlenecks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 4510–4520.

[21] MathWorks. (2020) Adaptive histogram equalization. [Online]. Available: https://www.mathworks.com/help/images/adaptive-histogram-equalization.html

[22] T. U. Ahmed, S. Hossain, M. S. Hossain, R. ul Islam, and K. Andersson, "Facial expression recognition using convolutional neural network with data augmentation," in *2019 Joint 8th International Conference on Informatics, Electronics & Vision (ICIEV) and 2019 3rd International Conference on Imaging, Vision & Pattern Recognition (icIVPR)*. IEEE, 2019, pp. 336–341.

[23] M. Akter, M. S. Hossain, T. Uddin Ahmed, and K. Andersson, "Mosquito classication using convolutional neural network with data augmentation," in *3rd International Conference on Intelligent Computing & Optimization 2020, ICO 2020*, 2020.

[24] R. C. Gonzalez, R. E. Woods, and S. L. Eddins, *Digital image processing using MATLAB*. Pearson Education India, 2004.

[25] D. M. Powers, "Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation," *arXiv preprint arXiv:2010.16061*, 2020.

[26] N. Ketkar, "Introduction to keras," in *Deep learning with Python*. Springer, 2017, pp. 97–111.

[27] G. M. Foody and A. Mathur, "The use of small training sets containing mixed pixels for accurate hard image classification: Training on mixed spectral responses for classification by a svm," *Remote Sensing of Environment*, vol. 103, no. 2, pp. 179–189, 2006.

[28] A. Bosch, A. Zisserman, and X. Munoz, "Image classification using random forests and ferns," in *2007 IEEE 11th international conference on computer vision*. Ieee, 2007, pp. 1–8.

[29] D. Mungra, A. Agrawal, P. Sharma, S. Tanwar, and M. S. Obaidat, "Pratit: a cnn-based emotion recognition system using histogram equalization and data augmentation," *Multimedia Tools and Applications*, pp. 1–23, 2019.

[30] M. S. Hossain, M. S. Khalid, S. Akter, and S. Dey, "A belief rule-based expert system to diagnose influenza," in *2014 9Th international forum on strategic technology (IFOST)*. IEEE, 2014, pp. 113–116.

[31] R. Karim, K. Andersson, M. S. Hossain, M. J. Uddin, and M. P. Meah, "A belief rule based expert system to assess clinical bronchopneumonia suspicion," in *2016 Future Technologies Conference (FTC)*. IEEE, 2016, pp. 655–660.

[32] S. Rahaman and M. S. Hossain, "A belief rule based clinical decision support system to assess suspicion of heart failure from signs, symptoms and risk factors," in *2013 International Conference on Informatics, Electronics and Vision (ICIEV)*. IEEE, 2013, pp. 1–6.