

Research Article

Face Mask Detection Using Deep Convolutional Neural Network and MobileNetV2-Based Transfer Learning

Dostdar Hussain,¹ Muhammad Ismail,¹ Israr Hussain,² Roobaea Alroobaea ,³
Saddam Hussain ,⁴ and Syed Sajid Ullah ,^{5,6}

¹Department of Computer Sciences, Karakoram International University Gilgit, 15100, Pakistan

²Shenzhen Key Laboratory of Media Security, Shenzhen University, Shenzhen 518060, China

³Department of Computer Science, College of Computers and Information Technology, Taif University, P. O. Box 11099, Taif 21944, Saudi Arabia

⁴Department of Information Technology, Hazara University, Mansehra, 21120 KP, Pakistan

⁵Department of Information and Communication Technology, University of Agder (UiA), N-4898 Grimstad, Norway

⁶Department of Electrical and Computer Engineering, Villanova University, PA 19085, USA

Correspondence should be addressed to Saddam Hussain; saddam_1993@hotmail.com
and Syed Sajid Ullah; sajidullah718@gmail.com

Received 11 February 2022; Revised 13 March 2022; Accepted 22 March 2022; Published 4 May 2022

Academic Editor: Nawab Muhammad Faseeh Qureshi

Copyright © 2022 Dostdar Hussain et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The rapid spreading of Coronavirus disease 2019 (COVID-19) is a major health risk that the whole world is facing for the last two years. One of the main causes of the fast spreading of this virus is the direct contact of people with each other. There are many precautionary measures to reduce the spread of this virus; however, the major one is wearing face masks in public places. Detection of face masks in public places is a real challenge that needs to be addressed to reduce the risk of spreading the virus. To address these challenges, an automated system for face mask detection using deep learning (DL) algorithms has been proposed to control the spreading of this infectious disease effectively. This work applies deep convolution neural network (DCNN) and MobileNetV2-based transfer learning models for effectual face mask detection. We evaluated the performance of these two models on two separate datasets, i.e., our developed dataset by considering real-world scenarios having 2500 images (dataset-1) and the dataset taken from PyImage Search Reader Prajna Bhandary and some random sources (dataset-2). The experimental results demonstrated that MobileNetV2 achieved 98% and 99% accuracies on dataset-1 and dataset-2, respectively, whereas DCNN achieved 97% accuracy on both datasets. Based on our findings, it can be concluded that the MobileNetV2-based transfer learning model would be an alternative to the DCNN model for highly accurate face mask detection.

1. Introduction

Coronavirus (COVID-19) is the latest evolutionary virus that has taken over the world in just a few months. It is a type of pneumonia that was initiated at the beginning of December 2019 near Wuhan City, Hubei Province, China, while, on 11th March 2020, it was declared as a world pandemic by the World Health Organization (WHO) [1]. According to WHO statistics, till 24 February 2021, more than 111 million people were affected by the virus and about 2.46 million deaths were reported [2]. The most common

symptoms of Coronavirus are fever, dry cough, and tiredness among many others. It mainly spreads through close direct contact of people with respiratory drops of an infected person generated through coughs, sneezes, or exhales. Since these droplets are too dense to swing in the air for long distances and quickly fall on floors or surfaces, therefore, it also spreads when individuals touch the impaired surfaces with the virus and touch back to their face (e.g., eyes, nose, and mouth) [3]. The WHO has declared a state of emergency all over the world and has developed some emergency precautionary measures to limit the spread of the virus, i.e.,

washing hands regularly with soap for the 20 s, using sanitizers, keeping distance, regularly disinfecting the surfaces, using disposable tissues while coughing or sneezing, and the most importantly wearing of face masks in public places [4, 5]. Like controlling the community spread of the SARS virus effectively during the SARS epidemic in the year 2003 [6], wearing community-wide face masks has also been proven very effective in controlling the widespread of Coronavirus [7–11]. Due to effective controlling the respiratory droplets, the wearing of masks has become a prominent feature of the COVID-19 response. For instance, the efficiency of N95 and surgical masks in blocking virus transmission (through blocking the respiratory droplets) is 91% and 68%, respectively [12, 13]. Wearing face masks can effectively interrupt airborne viruses and particles so that certain contaminants cannot reach the respiratory system of another person [14].

The worldwide scientific cooperation has improved dramatically due to the outbreak of Coronavirus and is searching for new tools and technologies to fight this virus. One such technology that can be used is artificial intelligence (AI). It can track the spread of the virus quickly, can recognize high-risk patients, and can potentially control the pandemic in real-time [3]. It is also beneficial in early predicting infection by the analysis of the previous patient’s data, which in turn can reduce the mortality risks because of the virus.

As already discussed that wearing face masks is the most effective protective measure against Coronavirus transmission, however, ensuring the wearing of face masks in public places is a difficult task for the government and the relevant authorities. Luckily, AI as a tool (by using machine learning (ML) or deep learning (DL) algorithms) can help ensure the wearing of face masks in public places just by detecting face masks in real-time with the help of an already installed camera network (surveillance camera network or any other). It is an easy method to manage the people in the society, to maintain social distancing, and to make sure that everyone has worn a face mask.

Because of the importance of face mask detection in public places, here in this study, we have demonstrated the use of two popular DL-based architectures, i.e., DCNN and advanced CNN based on “transfer learning,” i.e., MobileNetV2 for effective face mask detection. To evaluate the performance of the employed DL architectures, two different datasets have been used, i.e., (1) our own developed face mask detection dataset having 2500 images (dataset-1) and (2) the dataset taken from PyImage Search Reader Prajna Bhandary and some random sources (dataset-2). Finally, the results achieved by the algorithms on both datasets have been compared. Moreover, dataset-1 was collected from Karakoram International University, Pakistan, keeping in view the limitations of datasets under different real-world scenarios. In the future, this technology can be employed in real-time applications that require face mask detection for safety reasons due to the COVID-19 pandemic. This project can be integrated with embedded systems for applications in offices, schools, airports, and public places to ensure public safety guidelines.

TABLE 1: Dataset description.

Datasets	With mask	Without mask	Total images
Dataset-1	1250	1250	2500
Dataset-2	2220	2216	4436

2. Related Work

Loey et al. [15] introduced a face mask detection model that works on deep transfer learning and classical ML classifiers (classical ML classifiers refer to the ML algorithms that work on handcrafted extracted and engineered features from the input data). They used the Residual Neural Network (ResNet 50) algorithm for feature extraction. The extracted features were then used to train three classical ML algorithms, i.e., Support Vector Machine (SVM), Decision Tree (DT), and Ensemble Learning (EL). Three different face mask datasets have been used in the study for the investigation, i.e., (i) Real-World Masked Face Dataset (RMFD), (ii) Simulated Masked Face Dataset (SMFD), and (iii) Labeled Faces in the Wild (LFW) dataset. Finally, the trained classifiers were tested for possible face mask detection. During the testing experiment, the SVM classifier achieved the highest detection accuracies as compared to DT and EL classifiers. In RMFD and SMFD, it achieved 99.64% and 99.49% detection accuracies, respectively, while, in the case of LFW, it achieved 100% detection accuracy.

Militante and Dionisio [16] developed an automatic system to detect whether a person wears a mask or not and if the person does not wear a mask the system generates an alarm. To develop their system, the authors used the VGG-16 architecture of CNN. Their system achieved overall 96% detection accuracy. In the future, the authors decide to make a system that will not only detect whether a person is wearing a mask or not but will also detect a physical distance between each individual and will sound an alert if the physical distancing is not followed properly.

Rahman et al. [17] built a framework that gathers data from the IoT (Internet of Things) sensors of the smart city network (where all the public places are monitored with Closed-Circuit Television (CCTV) cameras) and detects whether an individual wears a mask or not. Real-time video footage from CCTV cameras of the smart city is collected for extracting facial images from it. The extracted facial images of people with and without masks are then used to train CNN architecture. Finally, the trained CNN architecture is used to distinguish people with and without facial masks. One advantage of using the CCTV network of the smart city is that when the system detects people without wearing masks, the information is sent through the city network to the concerned authority of the smart city to take appropriate action. The overall dataset used in this work is 1539 images including 858 images with masks and 681 images without masks. 80% of data was used for training, and the rest of the data was used for testing. The system achieved overall 98.7% accuracy on previously unseen data for distinguishing people with and without masks.

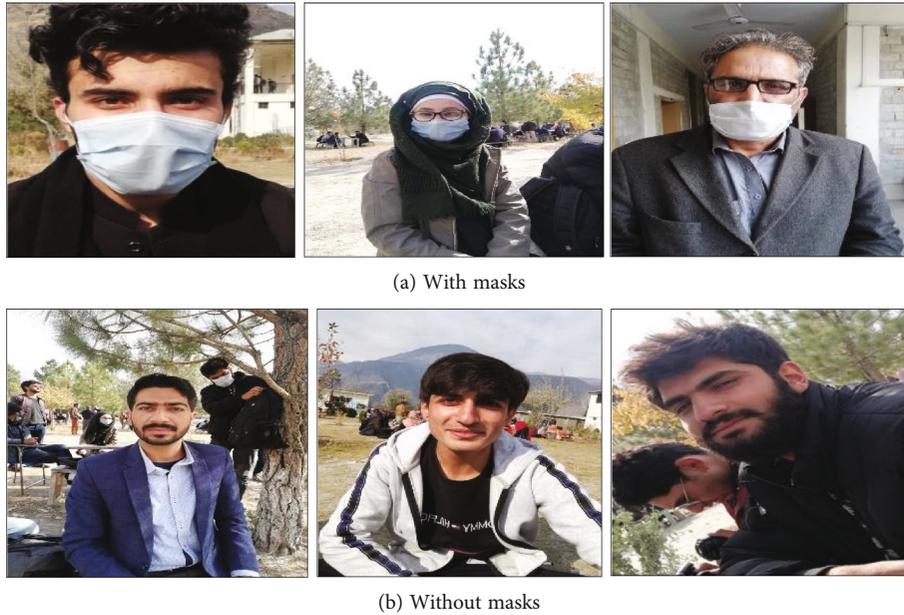


FIGURE 1: Some images with and without face masks taken from our own collected dataset (dataset-1).

Sanjaya and Rakhmawan [18] introduced a model using a DL algorithm to detect whether a person wears a mask or not in public areas. To do this, they used the MobileNetV2 image classification method, which is a pretrained method. In this experiment, the authors used two datasets, i.e., (1) RMFD, taken from Kaggle, and (2) dataset collected from 25 cities of Indonesia using CCTV cameras, traffic lamp cameras, and shop cameras. Both the datasets were used to train their model. The trained model achieved 96% and 85% detection accuracies on the test sets of these two datasets, respectively.

Sandler et al. [19] present a method for automatically detecting whether someone wears a mask or not. They designed a transfer learning-based method with MobileNetV2 for detecting masks in images and also in video streaming. Their system achieved 98% detection accuracy at a dataset of 4095 images. According to the authors, their model can work on devices with minimal computational capability and can process on real-time image data.

Oumina et al. [20] developed a system by combining pretrained DL models such as Xception, MobileNetV2, and VGG19 for the extraction of features from the input images. After extracting features, they used different ML classifiers such as SVM and k-nearest neighbor (k-NN) for the classification of extracted features of images. They used a total of 1376 images of two classes (i.e., with mask and without mask). The experimental results show that the combination of MobileNetV2 with SVM achieved the highest classification accuracy of 97.11%.

Related to the above-mentioned literature, here in this research, we have used two DL architectures for face mask detection, i.e., DCNN and transfer learning-based MobileNetV2. The performance of these architectures is evaluated on its own collected dataset as well as on the dataset collected from PyImage Search Reader Prajna Bhandary and some random sources. The purpose of evaluating the perfor-

mance of these two architectures on two different datasets is to compare their performance and to know how better the models perform on our own collected dataset.

3. Materials and Methods

3.1. Dataset. This research performed its experiments on two different datasets:

The first is our own collected dataset, which was collected at Karakorum International University, Pakistan, considering the real-world scenarios. The images were taken from each individual with and without wearing face masks. To increase the size of the dataset (for achieving better model performance), some augmentation techniques (like rotating, zooming, and blurring) have been performed on the collected images. The final version of our dataset included 2500 images labeled as with mask and without mask. The second dataset used in our experiments was taken from PyImage Search Reader developed by Mikolaj Witkowski and Prajna Bhandary (available at Kaggle) and some from random sources. This dataset contains 4436 images belonging to two classes (i.e., with mask and without mask). Mikolaj and Prajna created this dataset by taking normal images of faces and then by creating a custom computer vision python script to add a face mask to them, thus creating an artificial dataset that is used as a with mask. Details of the images included in both datasets have been provided in Table 1. Furthermore, Figures 1 and 2 show some pictures taken from our own collected dataset and the dataset developed by Prajna Bhandary.

3.2. Methodology of the Proposed Study. In this study, we are considering two DL architectures, i.e., DCNN and transfer learning-based MobileNetV2. To evaluate the performance of these two models, two different datasets have been used. For convenience, these datasets were named dataset-1 and



FIGURE 2: Some images with and without face masks taken from Prajna Bhandary dataset (dataset-2).

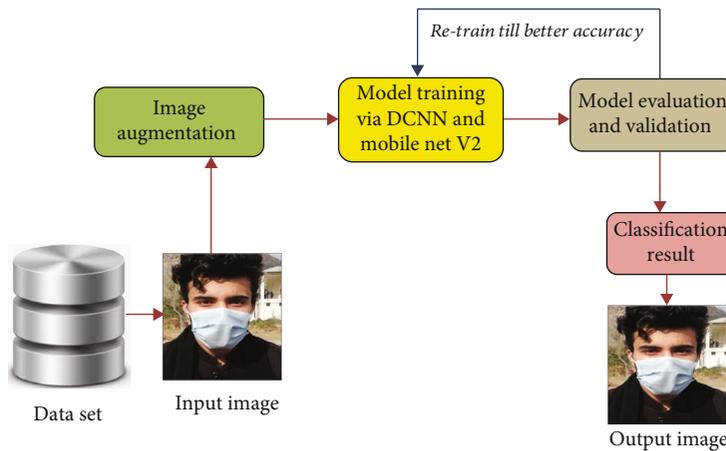


FIGURE 3: Schematic representation of the proposed work.

dataset-2, respectively. Dataset-1 contains 2500 and dataset-2 contains 4000 with and without mask images (refer to Figures 1 and 2 for a few samples of images taken from dataset-1 and dataset-2), respectively. Each dataset is split into two groups, one for training the models, while, the other for testing the models. In the case of training MobileNetV2 architecture, 80% data of each dataset was used, whereas, the remaining 20% data was used for testing the model. In the case of DCNN, 90% data of each dataset was used for training, and the remaining 10% was used for testing the model. Data augmentation technique was used to increase the amount of data by making slight changes like resizing, zooming, and rotating the images. This technique helps to reduce the problem of overfitting during training the model. We resized images to 100×100 , rotated images to 40 degrees, and zoomed images using a 0.2 zoom-in factor. A schematic diagram of DCNN and MobileNetV2 for face mask detection has been presented in Figure 3.

3.2.1. Deep Convolutional Neural Network (DCNN). DCNN is not just a deep neural network with many hidden layers, but actually, it is a deep network that mimics the way the human brain's visual cortex processes and recognizes images [21]. A given input image is processed by assigning relevant weights (learnable parameters) to various parts of the image and then making distinctions between the various characteristics. It is substantially less preprocessing and time-consuming than other classification methods when compared to DCNN. While traditional techniques necessitate the creation of filters by manually, DCNN can learn to create these filters with sufficient training. The DCNN architecture that we used in our research is depicted in Figure 4. These five layers include convolutional layers, average-pooling layers, and one fully connected layer. Convolutional layers and avg-pooling layers are included in this network. For every convolution layer, the layer is convolved with their respective kernel size and after every convolution; Rectified

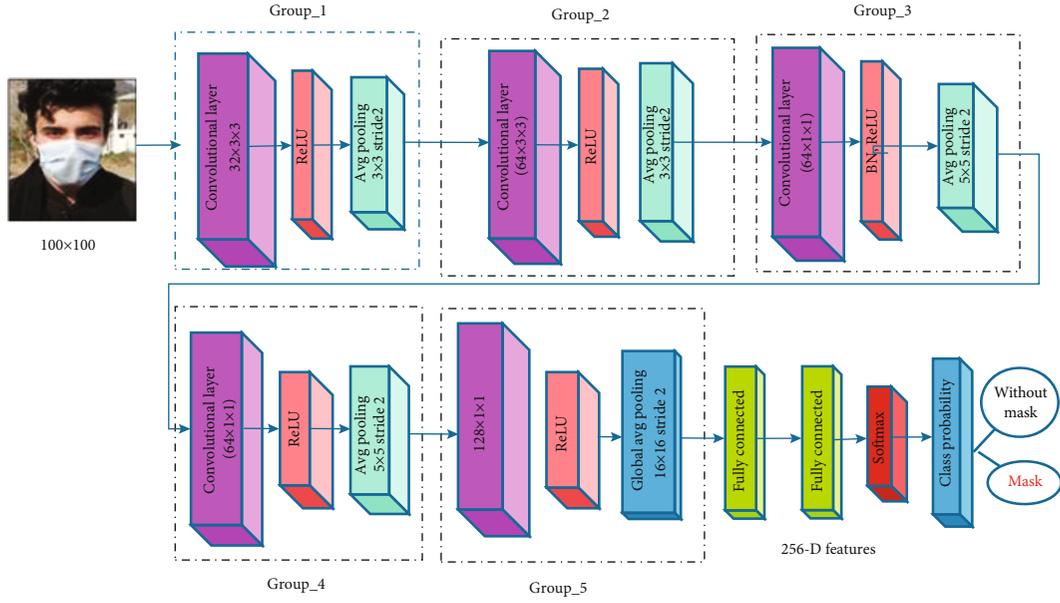


FIGURE 4: Deep Convolutional Neural Network Architecture.

Linear Unit (ReLU) activation function is added. ReLU is used for filtering information that propagates forward through the network. After every convolution, the avg-pooling operation takes place. The fully connected layer also known as the classification layer includes the flattening process. Flattening converts the matrix found after the last avg-pooling into a single column matrix for inputting it to the final output layer. At the last output layer, a “softmax” activation function is used that predicts a multinomial probability distribution.

3.2.2. *MobileNetV2*. MobileNetV2 is a Google-based developed architecture that is pertained on 1.4 million images of 1000 classes [19]. It is an advanced DCNN architecture that performs well on mobile devices. In MobileNetV2, we do not have to train the model from scratch, we only change the last output layers according to our domain. The architecture of MobileNetV2 is based on its previous version (i.e., MobileNetV1). To preserve the information, it introduced a new structure named “inverted residual.” The problem of information destroying in convolution blocks by a nonlinear layer applies the technique of Depthwise Separable Convolution (DSC) by using a linear bottleneck layer [22]. Figure 5 shows the basic architecture of MobileNetV2.

3.3. *Evaluation Metrics*. The performance of the classification models on testing data was evaluated using the accuracy (Equation (1)), precision (Equation (2)), recall (Equation (3)), specificity (Equation (4)), F1-score (Equation (5)), and kappa coefficient (Equation (6)). The F1-score is the harmonic mean of recall and precision. Recall, precision, and accuracy are computed using the True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN), which can be calculated using the confusion metric.

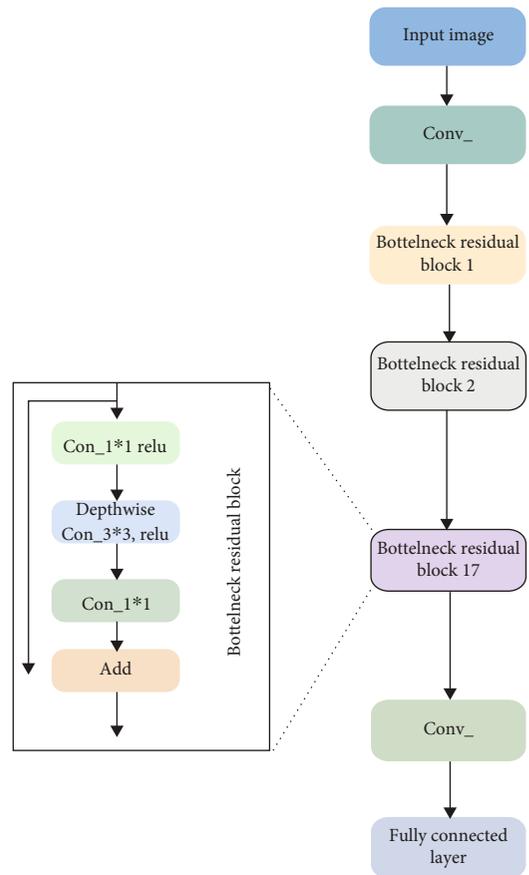


FIGURE 5: MobileNetV2 architecture. Source: Sandler et al. [19].

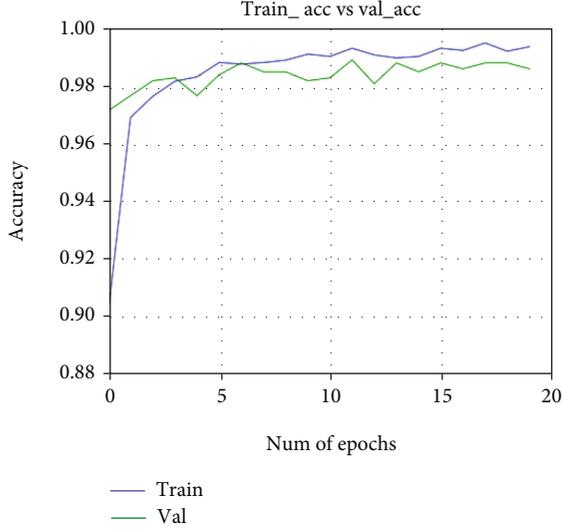


FIGURE 6: Training/validation accuracy graph of dataset-1 using the MobileNetV2.

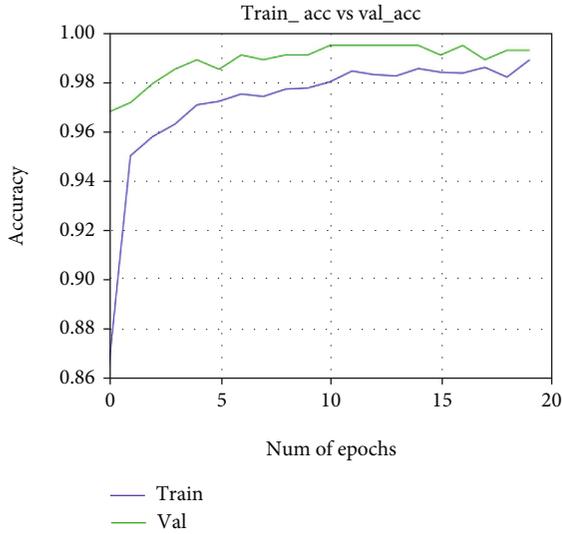


FIGURE 7: Training/validation accuracy graph of dataset-2 using the MobileNetV2.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}, \quad (1)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \quad (2)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \quad (3)$$

$$\text{Secificity} = \frac{\text{TN}}{\text{TN} + \text{FP}}, \quad (4)$$

$$\text{F1-score} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}, \quad (5)$$

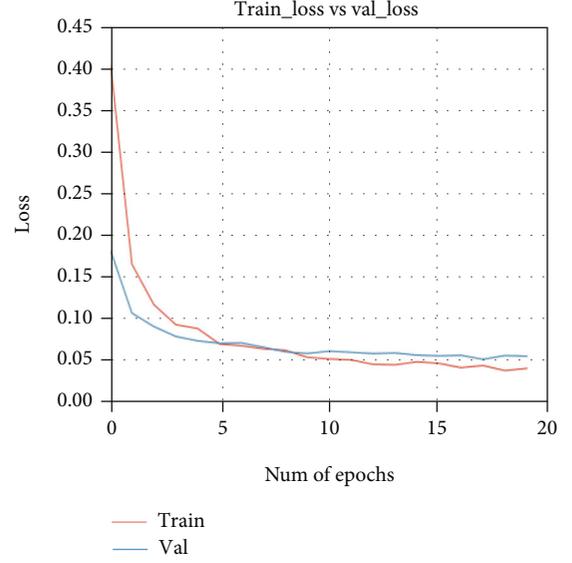


FIGURE 8: Training/validation loss graph of dataset-1 using MobileNetV2.

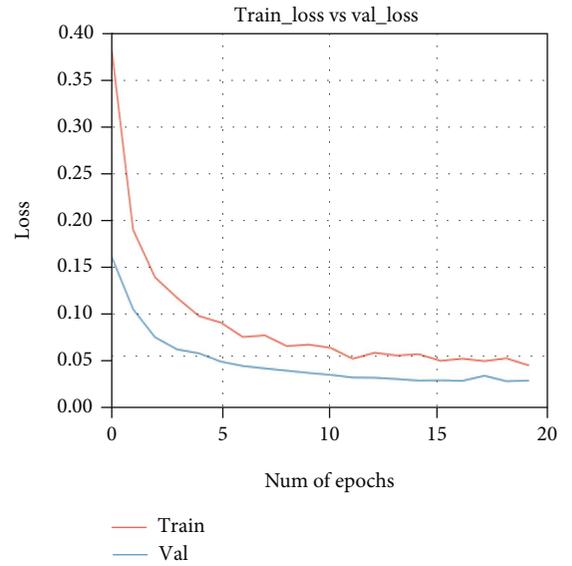


FIGURE 9: Training/validation loss graph of dataset-2 using MobileNetV2.

$$\text{Kappa coefficient} = \frac{(\text{total accuracy} - \text{random accuracy})}{(1 - \text{random accuracy})}. \quad (6)$$

Kappa coefficient is the measure of agreement between predicted and true values in testing datasets. The value of kappa can be 0 to 1. If the value of kappa is 0, there is no agreement between the predicted and actual image, and if the value of kappa is 1, then the predicted and actual image are identical. Thus, the higher the value of kappa, the more accurate the classification. Moreover, the random accuracy for binary classification can be calculated as

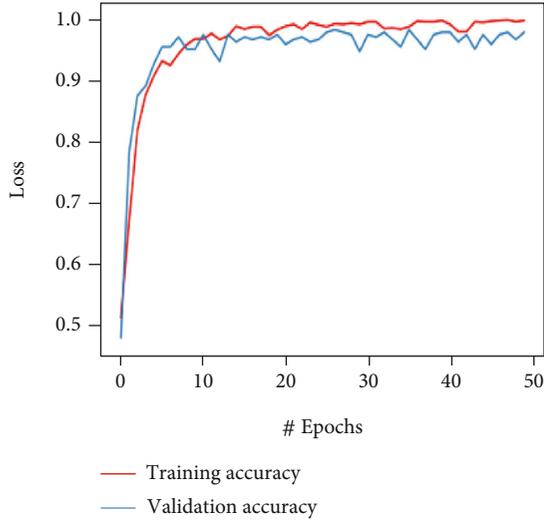


FIGURE 10: Training/validation loss graph of dataset-1 using DCNN.

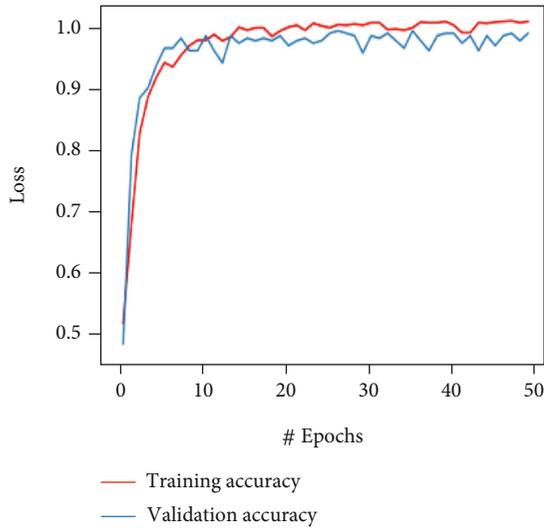


FIGURE 11: Training/validation loss graph of dataset-2 using DCNN.

$$\text{Random accuracy} = \frac{1}{\text{total no. of classes}}. \quad (7)$$

4. Experimental Results and Analysis

On two separate datasets of images, extensive experiments were conducted to evaluate the performance and effectiveness of the suggested models. On dataset-1, Figure 6 shows the MobileNetV2 model’s training and validation curves. It shows that over 20 epochs, the training and validation accuracy achieved by MobileNetV2 are 99% and 98%, respectively. Similarly, Figure 7 provides the training and validation plots of the MobileNetV2 model on dataset-2. Figure 7 indicates that over 20 epochs, the training and validation accuracy achieved by MobileNetV2 are 99% and 99%, respectively. Hence, the MobileNetV2 model achieved

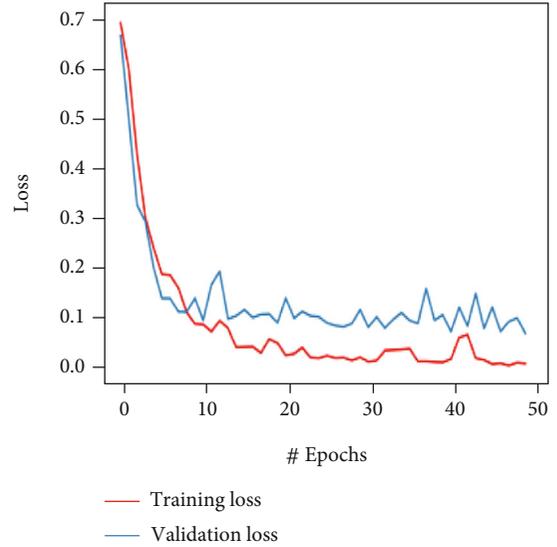


FIGURE 12: Training/validation loss graph of dataset-1 using DCNN.

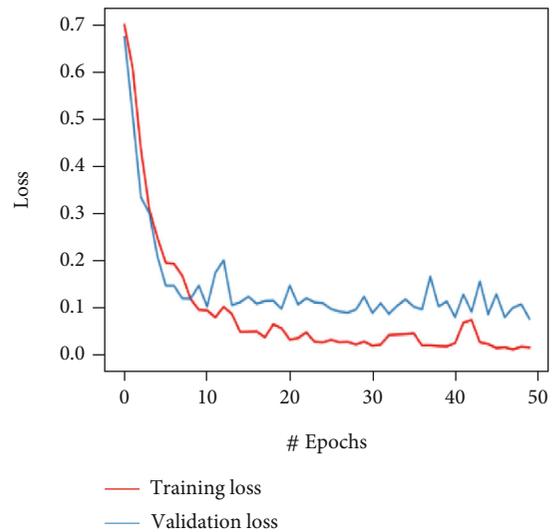


FIGURE 13: Training/validation loss graph of dataset-2 using DCNN.

equal training accuracy on both datasets, however, higher validation accuracy on dataset-2 as compared to dataset-1. Figure 8 presents the training and validation loss curves of the MobileNetV2 model on dataset-1 which shows that over 20 epochs, the training and validation losses are 5%. Similarly, Figure 9 provides the training and validation loss curves of the MobileNetV2 model on dataset-2. It shows that over 20 epochs, the training and validation losses acquired by MobileNetV2 are 5% and 4%, respectively. Hence, the MobileNetV2 model achieved comparatively equal training losses on both the datasets, whereas higher validation losses on dataset-1 as compared to dataset-2.

Figures 8 and 9 further specify that there is less gap in training and validation loss curves, which indicates that the employed model is well converged on the datasets and there

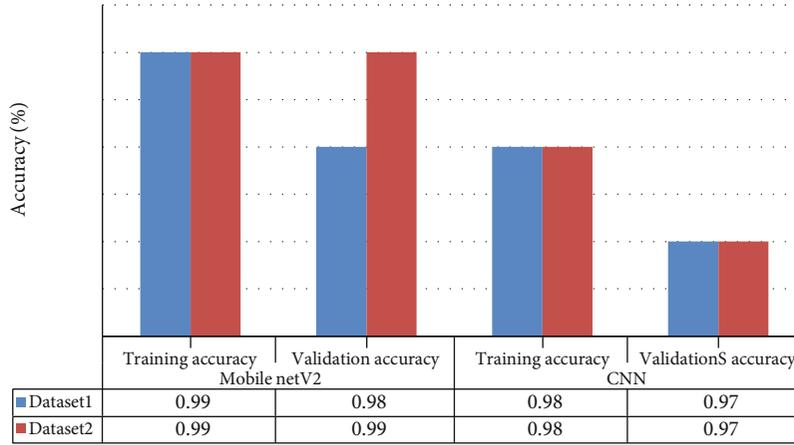


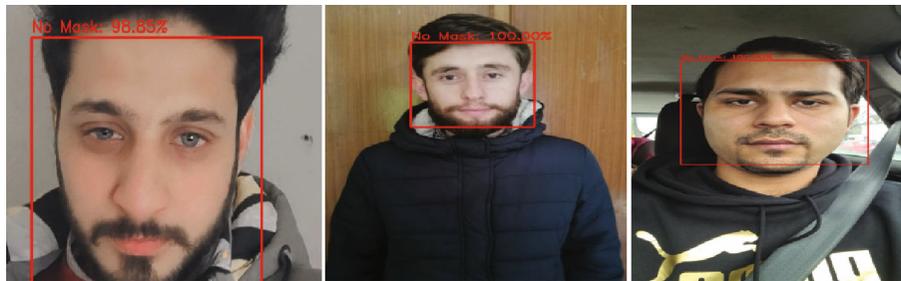
FIGURE 14: Training/validation accuracy graph of dataset-1 and dataset-2 using DCNN and MobilNetV2.

TABLE 2: The performance of MobileNetV2 for the test datasets.

Datasets	Precision	Recall	Specificity	F1-score	Accuracy	Error rate	Kappa coefficient
Dataset-1	0.98	0.98	0.98	0.98	0.98	0.014	0.96
Dataset-2	0.99	0.99	0.99	0.99	0.99	0.017	0.98

TABLE 3: The performance of DCNN for the test datasets.

Datasets	Precision	Recall	Specificity	F1-score	Accuracy	Error rate	Kappa coefficient
Dataset-1	0.96	0.98	0.95	0.97	0.97	0.02	0.94
Dataset-2	0.96	0.98	0.95	0.97	0.97	0.02	0.94



(a) Without masks



(b) With masks

FIGURE 15: Results of the test images with a face mask and without a face mask.

was no problem of overfitting occurring during training and validation. Figures 10 and 11 show the training and validation plots of the DCNN model on dataset-1 and dataset-2,

respectively. Figures 10 and 11 show that during 50 epochs, the training and validation accuracy achieved by the DCNN model on both the datasets are 98% and 97%, respectively.

Hence, the DCNN model performed equally well on dataset-1 and dataset-2.

Figures 12 and 13 describe the training and validation loss curves of the DCNN model achieved during experiments, plotted throughout 50 epochs using dataset-1 and dataset-2, respectively. Both the graphs indicate that in both the datasets, the DCNN model achieved the minimum and equal training and validation losses, i.e., 0.5% and 1%, respectively. Figures 12 and 13 further specify that there is less gap in training and validation loss curves, which indicates that the DCNN model was well converged during training and validation and no problem of overfitting occurred during the training and validation process. By comparing the training and validation accuracy as well as training and validation losses of MobileNetV2 and DCNN models, it can be concluded that the MobileNetV2 model achieved higher accuracy and minimum losses over DCNN in both datasets.

Figure 14 provides the overall comparative summary of the training and validation accuracy achieved by MobileNetV2 and DCNN models. It indicates that for both datasets, the MobileNetV2 model achieved higher accuracy as compared to the DCNN model. Furthermore, in the case of dataset-2, the accuracy achieved by the MobileNetV2 model are better as compared to dataset-1.

Tables 2 and 3 provide the classification reports of MobileNetV2 and DCNN on dataset-1 and dataset-2, respectively. Both models achieved higher evaluation metric scores of accuracy, precision, recall, F1-score, specificity, error rate, and kappa coefficient on both datasets. These higher values indicate that both models performed well on both datasets. However, from the overall experimental results, it can be concluded that even with the less amount of data, the MobileNetV2 model can provide better accuracies than DCNN. It is because MobileNetV2 models are already trained on a large amount of data and we do not need to train them from the scratch. We only have to change the last two layers of the MobileNetV2 model according to our problem. Moreover, results further show that in the case of our collected dataset (dataset-1), both the models performed well. It is because we collected our dataset with real face masks and without face masks and did not create a dataset with masks artificially. However, in the case of dataset-2, the dataset was generated with artificial masks. Furthermore, our models are succeeded in detecting the real-time images with a mask and without a mask with the accuracy displayed on the images as shown in Figure 15.

5. Conclusion

COVID-19 is one of the fast-spreading viruses that have been threatening human health, world trade, and the economy. Its high mutation and spreading rate made the situation difficult to be under control. Taking precautionary measures may reduce the spreading of this virus, and one of the most important measures is to wear a face mask in public places. Therefore, in this study, a deep learning-based approach has been applied to detect the face mask automatically. The learning models, i.e., Deep Convolutional

Neural Network (DCNN) and MobileNetV2 transferred learning-based model, have been evaluated on two different datasets. The datasets consist of our own collected dataset containing 2500 images of individuals with and without masks (dataset-1) and dataset-2 from PyImage Search Reader Prajna Bhandary (dataset-2) and some random sources. The comparative results show that MobileNetV2 achieved 98% and 99% classification accuracy on dataset-1 and dataset-2, respectively, whereas Deep Convolutional Neural Network achieved 97% accuracy on both datasets. The main contribution of the study is the development of our own face mask detection dataset with 2500 images, which were collected from Karakoram International University, Pakistan, keeping in view the limitations of datasets under different real-world scenarios. In the future, we will increase the size of the dataset by embedding real-time video streams into it to detect face masks in real-time.

Data Availability

The data used in this research can be obtained upon request to the corresponding authors.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors are grateful to the Taif University Researchers Supporting Project Number (TURSP-2020/36), Taif University, Taif, Saudi Arabia.

References

- [1] S. J. Fong, N. Dey, and J. Chaki, "An introduction to covid-19," in *In Artificial intelligence for coronavirus outbreak*, pp. 1–22, Springer, Singapore, 2021.
- [2] J. J. O. Canete, "When expressions of faith in the philippines becomes a potential covid-19," *Journal of Public Health*, vol. 43, no. 2, pp. e366–e367, 2021.
- [3] R. Vaishya, M. Javaid, I. H. Khan, and A. Haleem, "Artificial intelligence (ai) applications for covid-19 pandemic," *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, vol. 14, no. 4, pp. 337–339, 2020.
- [4] S. Bhagat, N. Yadav, J. Shah et al., "Novel corona virus (covid-19) pandemic: current status and possible strategies for detection and treatment of the disease," in *Expert Review of Anti-Infective Therapy*, pp. 1–24, Taylor & Francis, 2020.
- [5] Y. Xiao and M. E. Torok, "Taking the right measures to control covid-19," *The Lancet Infectious Diseases*, vol. 20, no. 5, pp. 523–524, 2020.
- [6] P. W. H. Peng, D. T. Wong, D. Bevan, and M. Gardam, "Infection control and anesthesia: lessons learned from the toronto sars outbreak," *Canadian Journal of Anesthesia*, vol. 50, no. 10, pp. 989–997, 2003.
- [7] V. C.-C. Cheng, S.-C. Wong, V. W.-M. Chuang et al., "The role of community-wide wearing of face mask for control of coronavirus disease 2019 (covid-19) epidemic due to sars-cov-2," *Journal of Infection*, vol. 81, no. 1, pp. 107–114, 2020.

- [8] E. Steffen, M. M. Eikenberry, E. Iboi et al., "To mask or not to mask: modeling the potential for face mask use by the general public to curtail the covid-19 pandemic," *Infectious Disease Modelling*, vol. 5, pp. 293–308, 2020.
- [9] T. M. Cook, "Personal protective equipment during the coronavirus disease (covid) 2019 pandemic—a narrative review," *Anaesthesia*, vol. 75, no. 7, pp. 920–927, 2020.
- [10] T. Greenhalgh, M. B. Schmid, T. Czypionka, D. Bassler, and L. Gruer, "Face masks for the public during the covid-19 crisis," *BMJ*, vol. 369, 2020.
- [11] C. Matuschek, F. Moll, H. Fangerau et al., "Face masks: benefits and risks during the covid-19 crisis," *European Journal of Medical Research*, vol. 25, no. 1, pp. 1–8, 2020.
- [12] T. Li, Y. Liu, M. Li, X. Qian, and S. Y. Dai, "Mask or no mask for covid-19: a public health and market study," *PloS one*, vol. 15, no. 8, article e0237691, 2020.
- [13] N. R. Longrich and S. K. Sheppard, *Publicuse of masks to control the coronavirus pandemic*, preprints.org, 2020.
- [14] B. Qin and D. Li, "Identifying facemask-wearing condition using image super-resolution with classification network to prevent covid-19," *Sensors*, vol. 20, no. 18, p. 5236, 2020.
- [15] M. Loey, G. Manogaran, M. H. N. Taha, and N. E. M. Khalifa, "A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic," *Measurement*, vol. 167, article 108288, 2021.
- [16] S. V. Militante and N. V. Dionisio, "Realtime facemask recognition with alarm system using deep learning," in *2020 11th IEEE Control and System Graduate Research Colloquium (ICSGRC)*, pp. 106–110, IEEE, Shah Alam, Malaysia, 2020.
- [17] M. M. Rahman, M. M. H. Manik, M. M. Islam, S. Mahmud, and J. H. Kim, "An automated system to limit covid-19 using facial mask detection in smart city network," in *2020 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS)*, Vancouver, BC, Canada, 2020IEEE.
- [18] S. A. Sanjaya and S. A. Rakhmawan, "Face mask detection using mobilenetv2 in the era of covid19 pandemic," in *2020 International Conference on Data Analytics for Business and Industry: Way Towards a Sustainable Economy (ICDABI)*, pp. 1–5, IEEE, Sakheer, Bahrain, 2020.
- [19] 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*, pp. 4510–4520, Utah, 2018.
- [20] A. Oumina, N. El Makhfi, and M. Hamdi, "Control the covid-19 pandemic: face mask detection using transfer learning," in *In 2020 IEEE 2nd International Conference on Electronics, Control, Optimization and Computer Science (ICECOCS)*, Kenitra, Morocco, 2020IEEE.
- [21] P. Kim, "Convolutional neural network," in *In MATLAB deep learning*, pp. 121–147, Springer, 2017.
- [22] K. Dong, C. Zhou, Y. Ruan, and Y. Li, "Mobilenetv2 model for image classification," in *In 2020 2nd International Conference on Information Technology and Computer Application (ITCA)*, pp. 476–480, IEEE, Guangzhou, China, 2020.