Deep learning based masked face recognition in the era of the COVID-19 pandemic

Ashwan A. Abdulmunem, Noor D. Al-Shakarchy, Mais Saad Safoq
Department of Computer Science, Faculty of Computer Science and Information Technology, Kerbala University, Kerbala, Iraq

Article Info

Article history:

Received Apr 21, 2022
Revised Sep 28, 2022
Accepted Oct 27, 2022

Keywords:

Classification
Convolutional neural network
COVID-19
Deep learning model
Face mask detection

ABSTRACT

During the coronavirus disease 2019 (COVID-19) pandemic, monitoring for wearing masks obtains a crucial attention due to the effect of wearing masks to prevent the spread of coronavirus. This work introduces two deep learning models, the former based on pre-trained convolutional neural network (CNN) which called MobileNetv2, and the latter is a new CNN architecture. These two models have been used to detect masked face with three classes (correct, not correct, and no mask). The experiments conducted on benchmark dataset which is face mask detection dataset from Kaggle. Moreover, the comparison between two models is driven to evaluate the results of these two proposed models.

Corresponding Author:

Ashwan A. Abdulmunem
Department of Computer Science, Faculty of Computer Science and Information Technology, Kerbala University
Kerbala, Iraq
Email: ashwan.a@uokerbala.edu.iq

1. INTRODUCTION

The worldwide spread of coronavirus disease 2019 (COVID-19) pandemic necessitates an immediate commitment to the battle throughout the human population. For this sudden outbreak and abandoned situation, the human health care emergency is restricted. In this case, ingenious automation such as computer vision (machine learning, deep learning, artificial intelligence), and medical imaging (magnetic resonance imaging (MRI) scans, X-Ray) has produced a promising strategy against COVID-19 [1], [2]. Different systems implementation have been introduced to deal with the pandemic as a diagnosis such as lung defected COVID-19 recognition, segmentation of the affected area in the lungs. Other techniques as a prophylactic procedure to prevent affect by the COVID-19 such as monitoring the appropriate distance among persons and masked face recognition [3]. In the masked face field there are two categorizes, the former one can be considered as an emergency and security tackle and many methods designed to overcome and give acceptable recognition solutions [4], [5]. While the second category can be considered as a healthcare requirement to prevent spread the coronavirus. Because of the COVID-19 pandemic, the masked face detection problem becomes the most important area to innovate new methods and algorithms to detect people who are wearing or not wearing mask to reduce and prevent spread of COVID-19 [6], [7]. Resulting of the successful of using artificial intelligence and deep learning in several medical issues and health care systems, these techniques also applied on masked face detection and illustrate the effective success.

The goal of object detection issues is to localize and categorize items in images [8]. Masked face detection is categorized as an object detection problem as it is based on detecting the mask (object) in the images or videos. In the last few years, object detection and face recognition algorithms have progressed [9].

Since 2014, deep-learning applications in object identification have led to significant advancements in accuracy and detection speed [10]. In the next section, the recent works on the masked face will illustrated with brief details.

For the purpose of taking the accurate results of the deep learning algorithms in masked face detection problem, recently numerous face detection algorithms have been specifically implemented for face mask detection using convolutional neural network (CNN) models [11]–[14]. Adhikarla and Davison [11] designed a framework contains of eight object detection and four face detection models. The idea of using many numbers of models is fine-tune them for face mask detection. They use three labels for identification (with-mask, without-mask, and unsure). Although the improvement in accuracy results but still suffer from the cost of time complexity and computation.

Ding *et al.* [12] introduce two-branch CNN, including a global branch for discriminative global feature learning and a partial branch for latent part identification and discriminative partial feature learning. In global branch they employed ResNet-50 model to achieve the highest convolutional feature maps. While in local part, the latent part detection model has been used to localize the most discriminative latent region in masked face images. The new knowledge the author introduced is to fuse these two parts to gain enhancement in the performance of the masked face systems by combine CNN parameters in the two branches are shared to benefit each other and to make the two branches smaller and more informative features can be obtained. Despite of advantages of fusing methods to improve the accuracy rate of detection, some works tend to use the earliest features such as local binary patterns and combine them with CNN models [13]. This combination improves the findings of system but still the features of CNN discriminative than handcraft features.

In addition, many works introduced based on combination of different CNN models [14]–[16]. Two CNN models applied in [14], [15]. Singh *et al.* [14] used two pre-trained CNN models called YOLOv3 and faster regions with convolutional neural networks (R-CNN) to achieve and improve the results of this challenge. These combined models improve the accuracy and the performance of the masked face detection but still suffer from the complexity. Zhu *et al*. [15] used two CNN models the former one is Dilation RetinaNet face location (DRFL) network to detect faces in a crowd places and the next level the SRNet20 network handles classification process of masked face. The use of three models introduced for face mask detection [16], [17]. Hariri [16] used CNN models called VGG-16, AlexNet, and ResNet-50, to extract deep feature maps from the desired areas (mostly eyes and forehead regions).

The effectiveness of YOLOv4 has been verified in determining whether or not people wear a mask, based on a database of crowded images [18]. The results show that YOLOv4 is mistakenly identify masks for people who use elements on the face other than the mask. The database is small and perhaps this is the reason for the errors obtained by YOLOv4.

Kong *et al.* [17] build a combined system composed of three deep learning models, the first one is a single shot detector (SSD) model employed to extract the masked face out of the image and the next model use this as an input which is processed using an Hourglass model to align the eye-brow cropped image based on the points it retrieved, finally the FaceNet model is presented for the sake of recognition of the eye-brow image. Although the system obtained a good accuracy to recognize the face mask, it still needs some development to speed it up. Other CNN combined work from realizing that the robust features play a vital role in the detection results. Peng *et al.* [19] implemented a framework to select and progress only the key features using locally nonlinear feature fusion-based network (LNFF-Net) and fusing the visual saliency map and the heat map by using fully convolutional network (FCN). Remarkable results obtain from this fusion technique but still suffer from the complexity. Kaur *et al*. [20] developed a face mask recognition system using a camera to determine whether a person use a mask or not, the system trained using a CNN model for a dataset from Kaggle. The result was accurate and the model can predict the availability of the face mask.

Many systems have been constructed and developed for the sake of the importance of masked face recognition in different applications, the earliest proposed systems have been illustrated [21] and briefly discussed the time cost and accuracy of them. Some of these systems focused on the features extraction techniques and others are reliable on the deep learning model implemented. As a comprehensive purpose for the recognition of mask type, Song *et al.* [22] constructed a compound system. The composed of different network each one is adopted for a specific reason. The CNN employed for the detection of mask existing, AlexNet for determining the mask type, VGG16 to detect mask position and a facial recognition for identity recognition. The system performed well while implemented using four datasets, although time cost may be improved if some of the proposed models can be merged to handle one task.

A deep convolution neural network (DCNN) and MobileNetv2-based transfer learning models is considered [23] to test the effectivity of the trained models in mask detection, after training the models on two datasets the MobileNetv2 achieved the higher accuracy. Predicting the identity of the person after mask detection is entitled [24], the mask recognition is applied using CNN then the face is mapped with

68 landmarks which is used for the matching of face images with a masked face image. Despite the fact that the system performed well, the efficacy of the system may vary due to the small dataset employed.

Wu [25] proposed a new technique for the recognition process by building a system which is based on subsampling, the technique was implemented using an attention machine neural network and a ResNet for feature extraction, the experiments performed using real world masked face recognition dataset (RMFRD) and synthetic masked face recognition dataset (SMFRD) databases and shows a good performance due to the low cost of time and high accuracy rate. The transfer learning technique is adopted using an InceptionV3 pre-trained model which is implemented using a simulated masked face dataset (SMFD) [26]. Despite the accurate results presented, the time efficiency is unemphasized and the amount of time saved is not reported. Another transfer learning mechanism introduced using ResNet50 deep learning model for the feature extraction phase, then a system composed of three different models support vector machine (SVM), decision trees and ensemble algorithm embedded with ResNet50 to achieve the detection phase process, the system is checked using three datasets (RMFD, SMFD and LFW) [27]. Considering the recent deep learning techniques may improve the performance of the proposed approach and that can lead for time saving and simpler model.

The improvement of feature extraction was employed using the YOLOv4 [28]; the CSPDarkNet53 of YOLOv4 has been developed to enhance the feature extracted from the image. In an attempt to reduce the time and speed up the performance of the CNN, Asghar *et al.* [29] developed a MobileNet convolution neural network composed of depthwise separable filters which is an improvement of the current convolution network. The proposed system implemented using two datasets AIZOO and Moxa3K. In this work two frameworks for masked face detection have been proposed based on using two different CNN models. The first one is based on pre-trained model called MobileNet and the other one is a new CNN structure is proposed to find out the efficient model to detect wearing mask in correct manner or not or either does not wearing the mask. In the next sections the two proposed CNN models will be depict in detail with the dataset that utilize in the experimental phase.

2. PROPOSED WORK

The proposed work in this paper based on employing two CNN models for masked face detection. Firstly, pre-trained was used in the experiments and new CNN model was constructed to achieve the experiments on the same dataset from Kaggle for mask face detection. The general block diagram of the mask face detection problem illustrated in Figure 1.

From Figure 1, it can be clearly seen that the main stages are pre-processing, feature extraction and classification stages and these stages are same in the two proposed models. The main aim of the proposed architecture is predicting (decide) the manner of face mask situation, which it either correct/incorrect/or no mask. Then giving an appropriate alarm to the situation warning the monitoring systems. Following sections will explain the dataset that used in the experiments and the details of the two proposed CNN models.

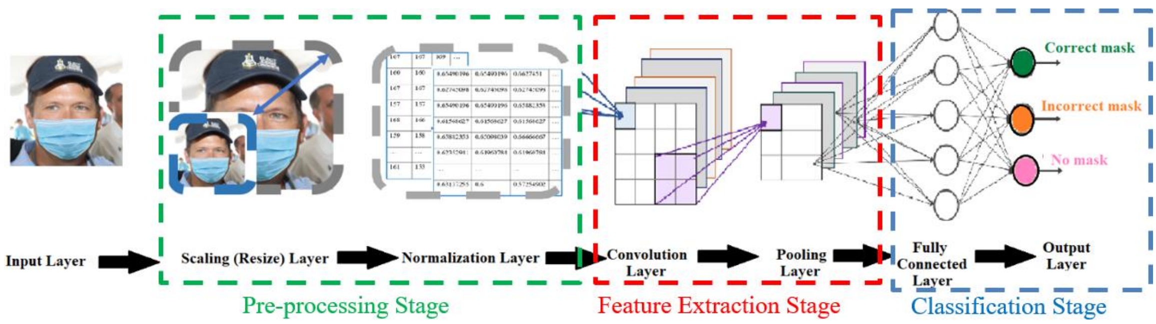


Figure 1. Steps of face mask detection system based on CNN models

2.1. Dataset

In recent trend in worldwide lockdowns due to COVID-19 outbreak, as face mask is becoming mandatory for everyone while roaming outside, approach of deep learning for detecting faces correct, Incorrect, and no mask were a good trendy practice. The whole dataset is split into two sub-datasets; training set represented 85% from the original data set and testing set with 15% used for testing the system on unseen data. The training set is also divided into actual training dataset represented 80% of the training dataset, and validation set represented 20% of the training dataset.

The main dataset used is face mask detection data set from Kaggle. This dataset consists of 7,553 RGB images in 2 folders as with mask and without mask. Images are named as label with mask (correct mask) and without mask (no mask). Images of faces with mask are 3,725 and images of faces without mask are 3828. The third-class incorrect mask is an enrichment to the dataset from MaskedFace-Net which is a dataset of correctly/incorrectly masked face images in the context of COVID-19. The images are added fairly.

2.2. Proposed pre-trained network structure

In order to obtain the benefits of the pre-trained CNN models, there are many applications and systems based on modification on these models to gain accurate results by only modifying some layers in the original architecture. The one experiment in this research is using MobileNetv2 [30] as a backbone architecture to detect three states of masked face (correct, incorrect, no mask). The architecture of modified MobileNetv2 model is revealed in Figure 2.

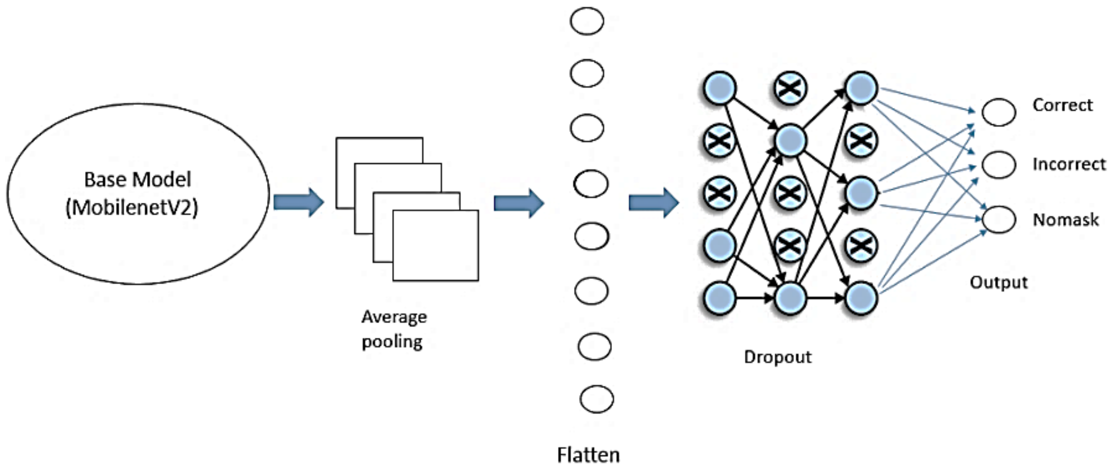


Figure 2. Proposed system method1 (build a CNN based on pre-trained architecture (MobileNetv2)

The figure shows that the base model is MobileNetv2 network, with ensuring the head FC layer sets are left off. The training stage followed by these steps. Firstly, generate class names. Secondly, freeze all layers except the final layers and train for some epochs until plateau (no improvement stage) is reached. Finally unfreeze all the layers and train all the weights while continuously reducing the learning rate until again plateau is reached.

2.3. New CNN model

The CNNs are a strong model that are easy to control and even easier to train. They provide immunity to overfitting at any alarming scales when being used on big data. The proposed face-mask CNN contained eight layers; the first five were two dimensional convolutional layers some of them followed by max-pooling layers, and the last three were fully connected layers. It used the non-saturating rectified linear units (ReLU) activation function, which showed improved training performance over tanh and sigmoid. The summary representation of the proposed architecture has been shown in Table 1 which visualizes the information of the proposed model layers, output shape, values of parameters (weights) for each layer, and finally the total number of parameters (weights) of the proposed model.

The pre-processing stage performs all preparing works on the input data to be suitable for implementation on a proposed classification prediction stage. This stage consists of resizing (scale), and normalization steps. The most important step is the scaling the dimensionality of cropped images in the samples to provide the generality to samples and to be suitable for the prediction deep learning model. The image scaling interprets as an image resizing that involve reconstruction image from the one-pixel grid to another by increasing or decreasing the number of pixels in samples images. For better results, image resizing uses one of image scaling algorithms that employing the interpolation of known data (the values at surrounding pixels) in image to estimate missing values at missing points.

Neural network models deal with small weight values during inputs processed. The learning process can be disrupted or slow down due to large integer values inputs. Normalization process changes the intensity range of input values with normally viewed so that each input value has a value range between 0 and 1.

Normalize the input value is done by dividing all values by the largest value (which is 255). Face-mask CNN model performs two functions respectively, feature extraction function to extract the salient features of the input image then classification function of these features. Each function employed some layers which are using to perform the specific purpose.

Table 1. The summary representation of the proposed system

Block no.	Layer (type)	Output shape	Parameters no.
1	conv2d_1 (Conv2D)	(None, 224, 224, 24)	672
	conv2d_2 (Conv2D)	(None, 224, 224, 24)	5208
	dropout_1 (Dropout)	(None, 224, 224, 24)	0
2	conv2d_3 (Conv2D)	(None, 224, 224, 32)	6944
	conv2d_4 (Conv2D)	(None, 224, 224, 32)	9248
	dropout_2 (Dropout)	(None, 224, 224, 32)	0
	batch_normalization_1 (Batch Normalization)	(None, 224, 224, 32)	128
3	conv2d_5 (Conv2D)	(None, 224, 224, 64)	18496
	conv2d_6 (Conv2D)	(None, 224, 224, 64)	36928
	max_pooling2d_1 MaxPooling2)	(None, 112, 112, 64)	0
4	conv2d_7 (Conv2D)	(None, 112, 112, 128)	73856
	conv2d_8 (Conv2D)	(None, 112, 112, 128)	147584
	dropout_3 (Dropout)	(None, 112, 112, 128)	0
5	conv2d_9 (Conv2D)	(None, 112, 112, 256)	295168
	conv2d_10 (Conv2D)	(None, 112, 112, 256)	590080
	batch_normalization_1 (Batch Normalization)	(None, 112, 112, 256)	1024
6	conv2d_11 (Conv2D)	(None, 112, 112, 512)	1180160
	conv2d_12 (Conv2D)	(None, 112, 112, 512)	2359808
	max_pooling2d_2 MaxPooling2)	(None, 56, 56, 512)	0
7	flatten_1 (Flatten)	(None, 1605632)	0
8	dense_1 (Dense)	(None, 512)	822084096
	batch_normalization_3 (Batch Normalization)	(None, 512)	2048
	dropout_3 (Dropout)	(None, 512)	0
9	dense_2 (Dense)	(None, 3)	1539
	Total parameters: 826,812,987		
	Trainable parameters: 826,811,387		
	Non-trainable parameters: 1,600		

2.3.1. Extraction step

The feature extractor step consists of many two-dimensional convolutions layers each one followed by a nonlinear layer represents the activation function applied in that layer to distinguish the special signal of useful features on each hidden layer. The primary concern with this model is preventing overfitting, in additional to the faster learning has a great influence on the performance of large models trained on large datasets; so that can be provided when using ReLU for all layers of the feature extraction step. The features maps of the input image are extracted and constructed by the convolution layer. In other words, the convolution layer acts the role of local filters on the input data and the filter kernel coefficients determined during the training process. Earlier convolution layer constructs low-level features in the input data as a set of primitive patterns. The next convolution layer can combine these primary features and constructs patterns of patterns and so on these secondary features combine and constructs the higher-level features patterns.

The pooling/subsampling layer is responsible for making the features robust against noise and blurring. This is implemented by reduction the resolution of the features. The activation maps (provided by convolutional layers) are pooling by applying non–linear down-sampling to discard weak information. The outputs of neuron clustered at one-layer combine using max pooling into a single neuron in the next layer which leads to the reduction in features resolution. The proposed model used the max-pooling function with window size=2 which takes a cluster of neurons at the prior layer and uses the maximum value as output. These clusters of neurons are produced by dividing the input image into non-overlapping two-dimensional spaces (2D matrix) and each space considered as cluster and the maximum value is selected. The pooling process presented in Figure 3.

To achieve a regularization approach to avoid overfitting and to effectively control noise during the training process, therefore; the proposed CNN model introduces a stochastic behavior in the neural activations by presents some dropout layers and normalize batch statistics in the feature activations by present some batch normalization layers. As well as batch normalization layers are added to accelerate the training process and coordinate the update of multiple layers in the model.

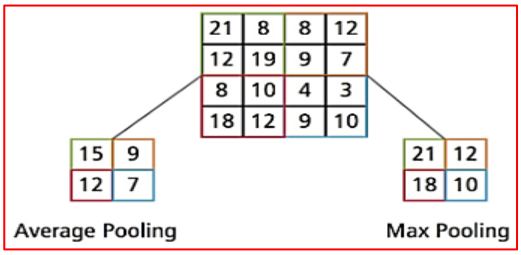


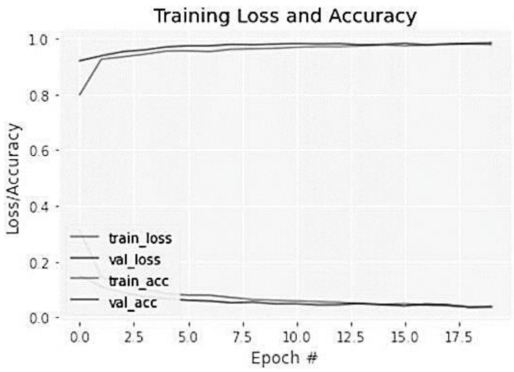
Figure 3. Pictorial representation of max pooling and average pooling

2.3.2. Classification step

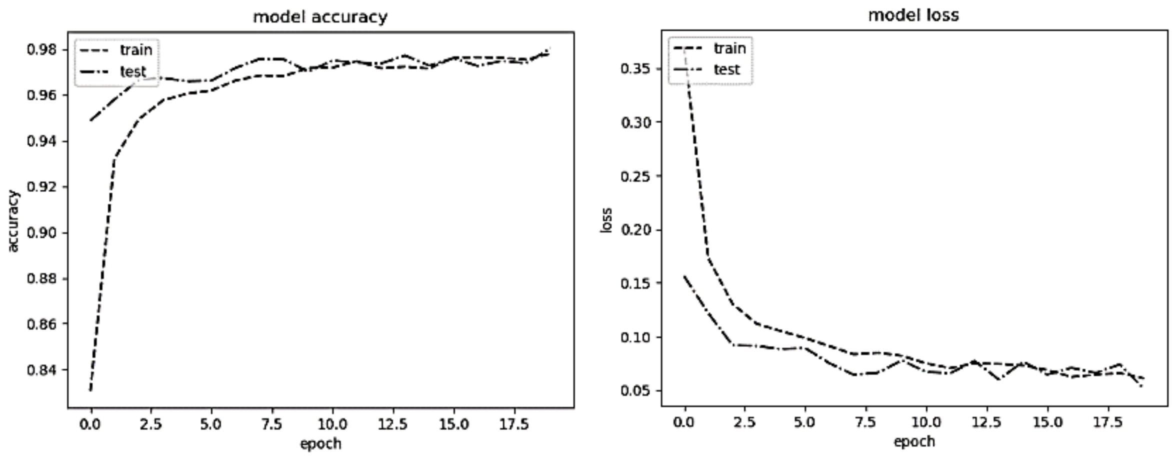
The classification step concern with finding a model relationship and determining the system orders and approximation of the unknown function by using fully connected layers also called dense and apply the ReLU activation function of these layers except the final fully connected layer called output layer or decision-making layer implements a SoftMax function. In this step, the dropout layer also adds with randomly around 25% of neurons selected.

3. RESULTS AND DISCUSSION

The performance of the training and evaluation modes is evaluated with two key points (metrics), which are accuracy and loss functions. A quick way to understanding the behavior for the learning of the proposed models on a specific dataset by evaluating the training and a validation dataset for each epoch and plot the results as can be shown in Figure 4(a) shows the loss and accuracy for pretrained model, while Figure 4(b) shows for the new model.



(a)



(b)

Figure 4. Accuracy and loss function of (a) pre-trained proposed model and (b) new CNN model

4. EVALUATION OF THE PROPOSED ARCHITECTURE

Deep learning models are stochastic, that each time the same model is fit on the same data, it may give different predictions and in turn have different overall skill. The evaluation of the model is based on the procedure to estimating model skill (controlling for model variance); that gives different results when the same model is trained on different data, by using k-fold cross-validation. The second procedure is estimating a stochastic model's skill (controlling for model stability); that different results when the same model is trained on the same data; by repeat the experiment of evaluating a non-stochastic model multiple times then calculate the mean of the estimated mean model skill, the so-called mean. Results of precision, recall and F1-measure of two models have been in Tables 2 and 3 consequently.

Table 2. Metrics for evaluation the pre-trained CNN model

Class	Precision	Recall	F1-measure
Correct	0.99	0.97	0.98
Incorrect	0.97	0.99	0.98
Nomask	1.00	1.00	1.00

Table 3. Metrics for evaluation the proposed CNN model

Class	Precision	Recall	F1-measure
Correct	0.99	0.92	0.95
Incorrect	0.96	1.00	0.98
Nomask	0.94	0.97	0.96

4.1. K-folds cross validation

The cross-validation procedure estimates the skill of a machine learning model on unseen data by evaluating the machine learning models on specific resampling data set based on a single parameter called *k*. In other words, the limited sample is used in order to estimate how the model is generally expected to predict unseen data during the training of the model. The procedure of 10-fold cross-validation with training and testing shown in Figure 5.

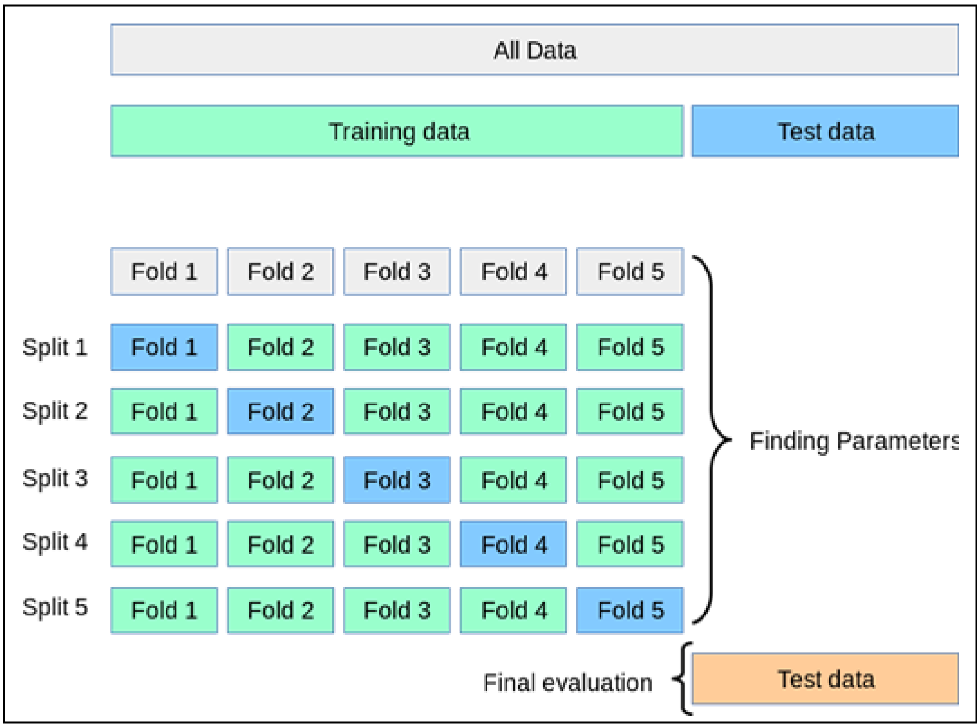


Figure 5. Cross-validation procedure with 5-Folds

The proposed models are given $k=10$, for each value of K, it will split dataset into $T_r(80\%)+V_a(10\%)=90\%$ and $T_e=10\%$, recording the testing performance according to the metric used (adopted accuracy). Finally, the average of the performance is computed to represent the final result. The 10-fold cross validation implementation and the experimental results can be represented through Table 4. A comparison between recent works and our proposed methods explains in Table 5.

Table 4. The 10-fold cross validation for proposed model

No of fold	Accuracy of each fold	Model accuracy
1	92.05%	97.59% (+/-2.48%)
2	97.75%	
3	94.85%	
4	99.77%	
5	100.00%	
6	100.00%	
7	99.56%	
8	98.27%	
9	97.69%	
10	95.99%	

Table 5. Comparison of recent works with our proposed methods

Reference	Proposed technique	No of classes	Result
[15]	Cascade network of two levees, (DRFL) Network and SRNet20 network.	3	90.6% on the wider face dataset and 98.5% on prepared dataset
[17]	SSD model, Hourglass network, FaceNet	3	
[22]	CNN, AlexNet, VGG16, and FaceNet	CNN of 2 classes AlexNet of 4 classes	97% accuracy
[28]	YOLOv4	3	98.3%
[24]	CNN	2	97.14%
[29]	MobileNet	2	93.14%
[26]	InceptionV3	2	100%
[27]	Resnet50, decision trees, SVM and ensemble algorithm	2	99.64% in RMFD and 99.49% in SMFD
Proposed pre-trained CNN	Based on MobileNetv2	3	99.05%
New CNN	New architecture of CNN	3	97.59%

4. CONCLUSION

The classification system using deep neural networks can be considered the best approach to achieve high accuracy and give better results than other traditional approaches in terms of accuracy and loss functions. During the period of this study, there are several notes that can be concluded as an outcome of this work. The proposed two models achieve success results to detect the status of face mask by correct/incorrect or no mask. Using 2D CNN provided the possibility of dealing with raw data, and this saved the time and effort required for the annoying pre-processing. The ReLU activation function integrated with the CNN is used to extract salient features and neglect the weak features, which leads to deal with noise associated with samples. The ReLU does that by removing all the noise elements from the sequence and keeping only those carrying a positive value. Adding the batch normalization layers after the convolutional layers to obey the convolutional property, which decreases the pose variation problem in the sample. In the batch normalization layer different elements of the same feature map, at different locations, are normalized in the same way independently of its direct special surroundings.

ACKNOWLEDGEMENTS

This work is supported by University of Kerbala.

REFERENCES

[1] A. Bhargava and A. Bansal, “Novel coronavirus (COVID-19) diagnosis using computer vision and artificial intelligence techniques: a review,” *Multimedia Tools and Applications*, vol. 80, no. 13, pp. 19931–19946, May 2021, doi: 10.1007/s11042-021-10714-5.

[2] C. Shorten, T. M. Khoshgoftaar, and B. Furht, “Deep learning applications for COVID-19,” *Journal of Big Data*, vol. 8, no. 1, Dec. 2021, doi: 10.1186/s40537-020-00392-9.

[3] M. Razavi, H. Alikhani, V. Janfaza, B. Sadeghi, and E. Alikhani, “An automatic system to monitor the physical distance and face mask wearing of construction workers in COVID-19 pandemic,” *SN Computer Science*, vol. 3, no. 1, Jan. 2022, doi: 10.1007/s42979-021-00894-0.

[4] H. Deng, Z. Feng, G. Qian, X. Lv, H. Li, and G. Li, “MFCosface: a masked-face recognition algorithm based on large margin cosine loss,” *Applied Sciences*, vol. 11, no. 16, Aug. 2021, doi: 10.3390/app11167310.

[5] S. Lin, L. Cai, X. Lin, and R. Ji, “Masked face detection via a modified LeNet,” *Neurocomputing*, vol. 218, pp. 197–202, Dec. 2016, doi: 10.1016/j.neucom.2016.08.056.

[6] M. Kuzu Kumcu, S. Tezcan Aydemir, B. Ölmez, N. Durmaz Çelik, and C. Yücesan, “Masked face recognition in patients with relapsing–remitting multiple sclerosis during the ongoing COVID-19 pandemic,” *Neurological Sciences*, vol. 43, no. 3,

pp. 1549–1556, Mar. 2022, doi: 10.1007/s10072-021-05797-9.
[7] S. Sethi, M. Kathuria, and T. Kaushik, "Face mask detection using deep learning: an approach to reduce risk of Coronavirus spread," *Journal of Biomedical Informatics*, vol. 120, Aug. 2021, doi: 10.1016/j.jbi.2021.103848.
[8] L. Liu *et al.*, "Deep learning for generic object detection: a survey," *International Journal of Computer Vision*, vol. 128, no. 2, pp. 261–318, Feb. 2020, doi: 10.1007/s11263-019-01247-4.
[9] Y. Wei, "Review of face recognition algorithms," in *5th International Conference on Biomedical Imaging, Signal Processing*, Sep. 2020, pp. 28–31, doi: 10.1145/3436349.3436367.
[10] Z.-Q. Zhao, P. Zheng, S.-T. Xu, and X. Wu, "Object detection with deep learning: a review," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 11, pp. 3212–3232, Nov. 2019, doi: 10.1109/TNNLS.2018.2876865.
[11] E. Adhikarla and B. D. Davison, "Face mask detection on real-world Webcam images," in *Proceedings of the Conference on Information Technology for Social Good*, Sep. 2021, pp. 139–144, doi: 10.1145/3462203.3475903.
[12] F. Ding, P. Peng, Y. Huang, M. Geng, and Y. Tian, "Masked face recognition with latent part detection," in *Proceedings of the 28th ACM International Conference on Multimedia*, Oct. 2020, pp. 2281–2289, doi: 10.1145/3394171.3413731.
[13] H. N. Vu, M. H. Nguyen, and C. Pham, "Masked face recognition with convolutional neural networks and local binary patterns," *Applied Intelligence*, vol. 52, no. 5, pp. 5497–5512, Mar. 2022, doi: 10.1007/s10489-021-02728-1.
[14] S. Singh, U. Ahuja, M. Kumar, K. Kumar, and M. Sachdeva, "Face mask detection using YOLOv3 and faster R-CNN models: COVID-19 environment," *Multimedia Tools and Applications*, vol. 80, no. 13, pp. 19753–19768, May 2021, doi: 10.1007/s11042-021-10711-8.
[15] R. Zhu, K. Yin, H. Xiong, H. Tang, and G. Yin, "Masked face detection algorithm in the dense crowd based on federated learning," *Wireless Communications and Mobile Computing*, vol. 2021, pp. 1–8, Oct. 2021, doi: 10.1155/2021/8586016.
[16] W. Hariri, "Efficient masked face recognition method during the COVID-19 pandemic," *Signal, Image and Video Processing*, vol. 16, no. 3, pp. 605–612, Apr. 2022, doi: 10.1007/s11760-021-02050-w.
[17] Y. Kong, S. Zhang, X. Li, K. Zhang, Y. Qi, and Z. Zhao, "Recognition system for masked face based on deep learning," in *2nd International Conference on Computer Vision, Image, and Deep Learning*, Oct. 2021, vol. 11911, doi: 10.1117/12.2604714.
[18] M. Harahap, L. Kusuma, M. Suryani, C. E. Situmeang, and J. F. Purba, "Identification of face mask with YOLOv4 based on outdoor video," *SinkrOn*, vol. 6, no. 1, pp. 127–134, Oct. 2021, doi: 10.33395/sinkron.v6i1.11190.
[19] X.-Y. Peng, J. Cao, and F.-Y. Zhang, "Masked face detection based on locally nonlinear feature fusion," in *Proceedings of the 2020 9th International Conference on Software and Computer Applications*, Feb. 2020, pp. 114–118, doi: 10.1145/3384544.3384568.
[20] G. Kaur *et al.*, "Face mask recognition system using CNN model," *Neuroscience Informatics*, vol. 2, no. 3, Sep. 2022, doi: 10.1016/j.neuri.2021.100035.
[21] M. Kumar and R. Mann, "Masked face recognition using deep learning model," in *3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N)*, Dec. 2021, vol. 10, no. 21, pp. 428–432, doi: 10.1109/ICAC3N53548.2021.9725368.
[22] Z. Song, K. Nguyen, T. Nguyen, C. Cho, and J. Gao, "Spartan face mask detection and facial recognition system," *Healthcare*, vol. 10, no. 1, Jan. 2022, doi: 10.3390/healthcare10010087.
[23] D. Hussain, M. Ismail, I. Hussain, R. Alroobaea, S. Hussain, and S. S. Ullah, "Face mask detection using deep convolutional neural network and MobileNetv2-based transfer learning," *Wireless Communications and Mobile Computing*, vol. 2022, pp. 1–10, May 2022, doi: 10.1155/2022/1536318.
[24] M. S. Mazli Shahar and L. Mazalan, "Face identity for face mask recognition system," in *11th IEEE Symposium on Computer Applications and Industrial Electronics (ISCAIE)*, Apr. 2021, pp. 42–47, doi: 10.1109/ISCAIE51753.2021.9431791.
[25] G. Wu, "Masked face recognition algorithm for a contactless distribution cabinet," *Mathematical Problems in Engineering*, vol. 2021, pp. 1–11, May 2021, doi: 10.1155/2021/5591020.
[26] G. Jignesh Chowdary, N. S. Punn, S. K. Sonbhadra, and S. Agarwal, "Face mask detection using transfer learning of InceptionV3," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 12581, Springer, Cham, 2020, pp. 81–90.
[27] 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, Jan. 2021, doi: 10.1016/j.measurement.2020.108288.
[28] J. Yu and W. Zhang, "Face mask wearing detection algorithm based on improved YOLO-v4," *Sensors*, vol. 21, no. 9, May 2021, doi: 10.3390/s21093263.
[29] M. Z. Asghar *et al.*, "Facial mask detection using depthwise separable convolutional neural network model during COVID-19 pandemic," *Frontiers in Public Health*, vol. 10, Mar. 2022, doi: 10.3389/fpubh.2022.855254.
[30] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: inverted residuals and linear bottlenecks," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Jun. 2018, pp. 4510–4520, doi: 10.1109/CVPR.2018.00474.

BIOGRAPHIES OF AUTHORS

Ashwan A. Abdulmunem received her Ph.D. degree in computer vision, artificial intelligence from Cardiff University, UK. Recently, she is an asst. prof. at College of Computer Science and Information Technology, University of Kerbala. Her research interests include computer vision and graphics, computational imaging, pattern recognition, artificial intelligence, machine learning, and deep learning. She can be contacted at email: Ashwan.a@uokerbala.edu.iq.

Noor D. Al-Shakarchy is a faculty member of Computer Science and Information Technology College, Kerbala University, Karbala, Iraq. She received Ph.D. from The University of Babylon, and B.Sc. and M.Sc. at University of Technology, Computer Science and Information Systems/Information Systems Department, Bagdad, Iraq in 2000 and 2003 respectively. Here research interests include computer vision, pattern recognition, Information Security, Artificial intelligent and deep neural networks. She can be contacted at email: noor.d@uokerbala.edu.iq.

Mais Saad Safoq received the B.Sc. and M.Sc. degrees in computer science from Beirut Arab University, Beirut, Lebanon, in 2008 and 2011, respectively. She has been a lecturer of computer science with Kerbala University, since 2012. She has authored or coauthored four refereed journal and conference papers. Her research interests include the applications of artificial intelligence, regression, classification models. Data Mining, pattern recognition and computer vision. She can be contacted at email: mais.s@uokerbala.edu.iq.