

Deep Learning Based Face Mask Detection in Religious Mass Gathering During COVID-19 Pandemic

Abdullah S. AL-Malaise AL-Ghamdi^{1,2,3}, Sultanah M. Alshammari^{3,4} and Mahmoud Ragab^{3,5,6,*}

¹Information Systems Department, Faculty of Computing and Information Technology King Abdulaziz University, Jeddah, 21589, Saudi Arabia

²Information Systems Department, HECI School, Dar Alhekma University, Jeddah, Saudi Arabia

³Center of Excellence in Smart Environment Research, King Abdulaziz University, Jeddah, 21589, Saudi Arabia

⁴Department of Computer Science, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, 21589, Saudi Arabia

⁵Information Technology Department, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, 21589, Saudi Arabia

⁶Department of Mathematics, Faculty of Science, Al-Azhar University, Naser City, Cairo, 11884, Egypt

*Corresponding Author: Mahmoud Ragab. Email: mragab@kau.edu.sa

Received: 07 September 2022; Accepted: 08 December 2022

Abstract: Notwithstanding the religious intention of billions of devotees, the religious mass gathering increased major public health concerns since it likely became a huge super spreading event for the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). Most attendees ignored preventive measures, namely maintaining physical distance, practising hand hygiene, and wearing face-masks. Wearing a face mask in public areas protects people from spreading COVID-19. Artificial intelligence (AI) based on deep learning (DL) and machine learning (ML) could assist in fighting covid-19 in several ways. This study introduces a new deep learning-based Face Mask Detection in Religious Mass Gathering (DLFMD-RMG) technique during the COVID-19 pandemic. The DLFMD-RMG technique focuses mainly on detecting face masks in a religious mass gathering. To accomplish this, the presented DLFMD-RMG technique undergoes two pre-processing levels: Bilateral Filtering (BF) and Contrast Enhancement. For face detection, the DLFMD-RMG technique uses YOLOv5 with a ResNet-50 detector. In addition, the face detection performance can be improved by the seeker optimization algorithm (SOA) for tuning the hyperparameter of the ResNet-50 module, showing the novelty of the work. At last, the faces with and without masks are classified using the Fuzzy Neural Network (FNN) model. The stimulation study of the DLFMD-RMG algorithm is examined on a benchmark dataset. The results highlighted the remarkable performance of the DLFMD-RMG model algorithm in other recent approaches.

Keywords: Religious mass gathering; Hajj and Umrah; covid-19 pandemic; face mask; computer vision; deep learning



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1 Introduction

Religious events involving the gathering of an enormous number of practitioners of the religion are significant ritualized features. A massive gathering event (MGE) can be no exception, an assembly of many people attending an organized event within a definite space [1]. There exist large public health consequences of MGE. Public health concerns during those events could encompass a range of health problems: spreading food, air, and water, exacerbating non-communicable diseases, vector-borne infectious disease, accidents, health consequences of alcohol and substance abuse, stampedes, and worsening of mental health problems [2]. The mass gathering is considered by the concentration of people at a particular place for a particular goal over a key planning recommendation for mass gathering during COVID19: Interim guidance-2-set time that can potentially strain the response and planning resources of the host community or country.

The mass gathering occurs as a single event or a grouping of numerous events at diverse places [3]. They might be private or public, spontaneous or planned, one-off or recurrent, and of different times and sizes. The range of mass gatherings is wider, from music, sports, religious, entertainment, or business events, to huge meetings and conferences. Some health interferences, like mass drug administrations or immunization campaigns, are considered mass gatherings [4]. The mass gathering includes high-visibility events, frequently related to international travel, great participation, extended media coverage, prolonged duration, and many venues (multiple host countries). Also, high-visibility events are often related to the improved frequency of small private gatherings (at home, in streets, restaurants or bars, etc.), which could characterize further difficulty since they are less controlled [5]. During the COVID19 pandemic, mass gatherings could be closely related to the high risk of spreading SARS-CoV-2; also, they can potentially strain the response and planning resources of the host community or country and be related to destructive effects on health services [6].

Organizing crowded event frequently includes complicated tasks, however, organizing them during pandemics or ongoing crises, namely COVID19 which needs adherence to numerous constraints, is increasingly sophisticated [7]. Also, many out-of-control crowded events exist with or without the participation of regulatory bodies (government of the region or country). Besides washing hands frequently and maintaining physical distancing, the appropriate usage of facemasks has now developed as the pillar to preventing community transmission of the disease [8]. The purpose is to protect oneself from getting infected and spreading the virus. Now, policymaker faces several risks and challenges while facing the transmission and spreading of COVID19. The laws and rules emerged as an action to considerably expand cases and deaths in several regions [9,10]. But monitoring massive groups of people has become increasingly complex.

This study introduces a new Deep Learning Based Face Mask Detection in Religious Mass Gathering (DLFMD-RMG) technique during the COVID-19 pandemic. The DLFMD-RMG technique focuses mainly on detecting face masks in religious mass gatherings. To accomplish this, the presented DLFMD-RMG technique undergoes two pre-processing levels: Bilateral Filtering (BF) and Contrast Enhancement. For face detection, the DLFMD-RMG technique uses YOLOv5 with a ResNet-50 detector. In addition, the face detection performance can be improved by the seeker optimization algorithm (SOA) for tuning the hyperparameter of the ResNet-50 model. At last, the faces with and without masks are classified using the Fuzzy Neural Network (FNN) model. The stimulation study of the DLFMD-RMG model is examined on benchmark datasets.

2 Literature Review

In [11], a single shot multibox detector and MobileNetV2 (SSDMNV2) proposed for the detection of face masks utilizing MobileNetV2, TensorFlow, Keras, and OpenCV Deep Neural Network (DNN)

structure are utilized as image classification. OpenCV DNN employed in SSDMNV2 has SSD, including ResNet-10, as the backbone and can identify faces from many angles. While MobileNetV2 offers lightweight and precise estimations for classification. In [12], a hybrid method utilizing deep and traditional ML for face mask detection is presented. This presented technique method has 2 elements, namely classification and feature extraction.

Gupta et al. [13] introduce a mask detector that utilizes an ML facial classification mechanism to determine whether the mask is worn. It is linked to a closed-circuit television (CCTV) mechanism verifying that only individuals wearing masks are allowable. Asif et al. [14] propose automatically utilizing DL to identify face masks in the video. The presented structure has 2 elements. Initially, it is devised to track and detect faces through ML and OpenCV; then, facial frames are processed into modelled deep transfer learning (DTL) MobileNetV2 to identify the mask region. In [15], an IoT-based smart door utilizes an ML technique for detecting face masks and monitoring body temperature. This method is utilized in apartment entrances, shopping malls, hotels, etc.

Loey et al. [16] intend to localize and annotate the face mask object in real-time images. Wearing a face mask in public places would protect individuals from the spread of transmission among them. The presented method has 2 elements. The first element was devised for the feature-extracting process related to the ResNet50 DTL method. At the same time, the second element was devised for detecting face masks related to YOLO v2. Singh et al. [17] suggest a method that would draw bounding boxes (green or red) around the faces of individuals, whether an individual was wearing a mask or not. The authors have also associated the performance of both methods: inference time and precision rate.

3 The Proposed Model

In this study, new DLFMD-RMG technology has been developed. The DLFMD-RMG technique focuses mainly on detecting face masks in religious mass gatherings. The DLFMD-RMG technique undergoes two pre-processing levels: the BF technique and Contrast Enhancement. Then, it involves two processes: face detection and face mask classification. Fig. 1 defines the overall flow of the DLFMD-RMG system.

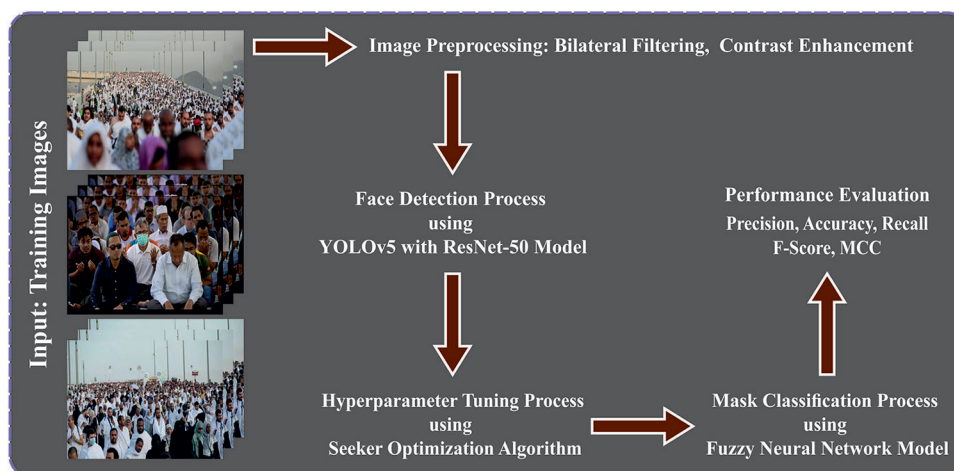


Figure 1: Overall process of DLFMD-RMG approach

3.1 Face Detection Module

For face detection, the DLFMD-RMG technique uses YOLOv5 with a ResNet-50 detector. YOLO is a novel target detection technique with higher accuracy and faster detection. The base YOLO model process image is at forty-five frames every second [18]. The YOLO network has 2 FC layers and twenty-four convolution layers. Interchanging 1×1 convolution layers reduces the feature space from previous layers. YOLO could pre-train the convolution layer on ImageNet classifier tasks at half the resolution (224×224 input image) and doubles the resolution for recognition. Several YOLO versions were generated. The YOLO v3 approach involves three techniques: YOLO v3-SPP, YOLO v3, and YOLO v3-tiny; the YOLO v4 approach involves four techniques: YOLO v4m-mish, YOLO v4s-mish, YOLO v4x-mish, and YOLO v4l-mish; YOLO v5 architecture involves YOLO v5s, YOLO v5n, YOLO v5x, and YOLO v5l.

The study chooses the YOLO v4s-mish, YOLO v5s, and YOLO v3 approaches for the study. The major reason is: firstly, the three modules could produce a test outcomes graph with a similar indicator that is appropriate for analysis and comparison; next, the three methods are light weighted, appropriate for target recognition in smaller scenes and smaller and medium datasets; Then, the 3 approaches depend on PyTorch DL architecture proposed by Facebook, and have better outcomes in target recognition.

The detection network is a convolution neural network (CNN) that encompasses output, transform, and convolution layers. The transform layer extracts activation of the convolution layer and increases the stability of the DNN. The location of pure bounding boxes can be generated using the output layer.

In this work, ResNet50 is employed as a deep transfer mechanism for extracting features [19]. A ResNet is a type of DTL that uses the residual network. ResNet50 has sixteen residual bottleneck blocks; every block has a convolutional size of 1×1 , 3×3 and 1×1 with a feature map (64, 128, 256, 512, 1024). To calculate MSE loss among the target and predicted bounding boxes, the loss function of YOLO v5 is computed by:

$$\text{YOLO v5 loss} = \text{Localization loss} + \text{Confidence loss} + \text{Classification loss} \quad (1)$$

The coefficient for computing the localization loss includes the grid cells' height (h) and width (w), which is given below.

$$\text{Localization loss} = q_1 \sum_{a=0}^{g^2} \sum_{b=0}^v 1_{ij}^{obj} \left[(x_j - \hat{x})^2 + (y_j - \hat{y})^2 \right] + q_1 \sum_{a=0}^{g^2} \sum_{b=0}^v 1_{ij}^{obj} \left[\left(\sqrt{w_i} - \sqrt{\hat{w}} \right)^2 + \left(\sqrt{h_i} - \sqrt{\hat{h}} \right)^2 \right] \quad (2)$$

From the expression, q_1 refers to a weight, g indicates the number of grid cells, v denotes the number of bounding boxes in g , (x_i, y_i) represents a center of v in g , (w_i, h_i) implies a width and height of v in g , (\hat{x}_i, \hat{y}_i) symbolizes a center of the target in every g , (\hat{w}_i, \hat{h}_i) signify a center of the target in every g , 1_{ij}^{obj} means 1 when there is an object in v in every g or else 0. Once the object is identified in v bounding boxes of g , Confidence loss calculates the confidence score of error. It is evaluated in the following expression.

$$\text{Confidence loss} = q_2 \sum_{a=0}^{g^2} \sum_{b=0}^v 1_{ij}^{obj} (cs_j - \hat{cs}_i)^2 + q_3 \sum_{i=0}^{g^2} \sum_{j=0}^v 1_{ij}^{noobj} (cs_j - \hat{cs}_i)^2 \quad (3)$$

In Eq. (3), q_2, q_3 represents a weight of confidence error, cs_i shows the confidence score of v in g , 1_{ij}^{obj} is 1 when there is an object in v in every g or else 0, 1_{ij}^{noobj} is 1 when there is no object in v in every g or else 0 is determined as follows.

$$Classification\ loss = q_4 \sum_{a=0}^{g^2} 1_i^{obj} \sum_{z \in classes} (p_i(z) - \hat{p}_i(z))^2 \tag{4}$$

In Eq. (4), q_4 indicates the weight of classifier error, $p_i(z)$, and $\hat{p}_i(z)$ denotes the likelihood of the estimated object and actual conditional classes in a grid cell.

In addition, the face detection performance can be improved by the SOA for the hyperparameter tuning of the ResNet-50 model. The SOA performs extensive analysis of human search behaviours [20]. Through “experience gradient”, the search direction has been determined, and uncertain reasoning is used to solve the search step measurement.

The SOA has three major updating steps. Here, i refers to i -th individual searchers, and j characterizes the dimension of the individual. s shows the overall number of individuals; D represents the overall number of dimensions; t indicates the present algebra, and $iter_{max}$ signifies the maximal optimization algebra. $x_{ij}(t)$ and $x_{ij}(t + 1)$, correspondingly, epitomize the searcher site at t and $(t + 1)$ algebra.

The forward direction of a search can be determined using the experience gradient attained from the individual movement and the assessment of other individuals searching previous locations. The pre-emptive direction $\vec{f}_{i,p}(t)$, egoistic direction $\vec{f}_{i,e}(t)$, and altruistic direction $\vec{f}_{i,a}(t)$ of i -th individuals in other dimensions could be attained.

$$\begin{aligned} \vec{f}_{i,e}(t) &= \vec{p}_{i,best} - \vec{x}_i(t), \\ \vec{f}_{i,a}(t) &= \vec{g}_{i,best} - \vec{x}_i(t), \\ \vec{f}_{i,p}(t) &= \vec{x}_i(t_1) - \vec{x}_i(t_2). \end{aligned} \tag{5}$$

The searcher uses a random weighted average to attain the search direction.

$$\vec{f}_i(t) = sign\left(\omega \vec{f}_{i,p}(t) + \phi_1 \vec{f}_{i,e}(t) + \phi_2 \vec{f}_{i,a}(t)\right), \tag{6}$$

In Eq. (6), $t_1, t_2 \in \{t, t - 1, t - 2\}$; $\vec{x}_i(t_1)$ and $\vec{x}_i(t_2)$ denotes the best advantage of $\{\vec{x}_i(t - 2), \vec{x}_i(t - 1), \vec{x}_i(t)\}$; $g_{i,best}$ refers to the previous optimum position in the locality where i -th search factors are positioned; $p_{i,best}$ denotes the optimal neighborhood from i -th search factors for the present locality; ψ_1 and ψ_1 indicate random numbers between zero and one; ω indicates the weight of inertia.

SOA describes the reasoning of fuzzy approximation capability. Through computer language, natural human language stimulates the behaviors of human intelligence. When the algorithm expressed fuzzy rules, it adapted to a better estimate of the optimization problem. But the smallest fitness corresponds to the smallest search step length.

$$\mu(\alpha) = e^{-\alpha^2/2\delta^2}, \tag{7}$$

In Eq. (7), α and δ represent parameters of the membership function. The likelihood of the output parameter exceeding $[-3\delta, 3\delta]$ is lesser than 0.0111. Then, $\mu_{min} = 0.0111$. Usually, the optimum location of the individual has $\mu_{max} = 1.0$, and the worst location is 0.0111. Choose the following function as a fuzzy parameter with “smaller” target function values:

$$\mu_i = \mu_{max} - \frac{s - I_i}{s - I} (\mu_{max} - \mu_{min}), \quad i = 1, 2, \dots, s, \tag{8}$$

$$\mu_{ij} = rand(\mu_i, 1), j = 1, 2, \dots, D, \quad (9)$$

From the expression, μ_{ij} can be defined using the above formula. I_i refers to the amount of the sequence $X(t)$ of the present individual ordered from higher to lower values. The function $rand(\mu_i, 1)$ represents the real number within $[\mu_i, 1]$. It is noted Eq. (8) simulate the random search behaviors of human. Step measurement of j -th dimension search interspace is defined as follows.

$$\alpha_{ij} = \delta_{ij} - \sqrt{-\ln(\mu_{ij})}, \quad (10)$$

In Eq. (10), δ_{ij} denotes a variable of the Gaussian distribution function as follows.

$$\omega = \frac{(iter_{max} - t)}{iter_{max}}, \quad (11)$$

$$\delta_{ij} = \omega |\vec{x}_{min} - \vec{x}_{max}|, \quad (12)$$

Now, ω indicates the weight of inertia. ω linearly reduced from 0.9 to 0.1 as the evolutionary algebra surges. \vec{x}_{min} and \vec{x}_{max} are the variates of minimal and maximal values.

After attaining the scout step measurement and direction, the position update is denoted as

$$x_{ij}(t+1) = x_{ij}(t) + \alpha_{ij}(t)f_{ij}(t), i = 1, 2, \dots, s; j = 1, 2, \dots, D. \quad (13)$$

$f_{ij}(t)$ and $\alpha_{ij}(t)$, correspondingly, the searcher search direction and step size at t time.

The SOA algorithm derives a fitness function from accomplishing better classification performance. It describes a positive integer to characterize the improved candidate solution performance. The reduction of the classification error rate is regarded as the fitness function in the following.

$$fitness(x_i) = ClassifierErrorRate(x_i) = \frac{\text{number of misclassified samples}}{\text{Total number of samples}} * 100 \quad (14)$$

3.2 Face Mask Classification Model

In this study, the faces with and without masks are classified using the FNN model. NN is an effective and dynamic approach to understanding supervised learning [21]. The FNN includes output, input, rule, and membership function (MF) layers. Fig. 2 represents the framework of FNN. The dimensional of the input vector signifies the count of neurons from the input layer $X = [x_1, x_2, \dots, x_n]$. The MF layer calculates the membership degree of input modules. All the neurons in this layer define the linguistic variable. The MF F_{ij} is determined as follows:

$$F_{ij} = e^{-\frac{(x_i - c_{ij})^2}{\sigma_{ij}^2}} \quad (15)$$

The MF F_{ij} illustrates the membership degree of input members ' i ' of fuzzy set ' j '. c_{ij} and σ_{ij}^2 define the center and width of membership Gaussian functions of x_i . The Gaussian MFs are the base for linking betwixt fuzzy systems and RBFN. The outcome of this function has extremely smooth and continuously takes a maximal value of one. The multivariate Gaussian function was designed for the production of univariate sets. During the rule layer of FNN, all the neurons illustrate a fuzzy rule, and their resultant activation degree A_j , can be demonstrated as:

$$A_j = F_{1j}(x_1)F_{2j}(x_2) \dots F_{nj}(x_n) \tag{16}$$

$$\bar{A}_j = \frac{A_j}{\sum_{j=1}^r A_j} \tag{17}$$

At this point, \bar{A}_j illustrates the normalization value, and r represents the count of neurons from the rule layer. The resultant layer identifies the outcome computation Y , which can be demonstrated as:

$$Y = \sum_{j=1}^r w_j \bar{A}_j \tag{18}$$

At this point, w_j signifies the weighted connecting resultant and rule layers. During the fuzzy NN learning procedure, the main function is for developing and training the weighted connecting the resultant layer w_j and rule layer, width σ_{ij}^2 and center c_{ij} of membership Gaussian function of x_i . The gradient descent is the usual famous technique employed for enhancing these parameters. It can be an optimized technique employed for minimizing a function with iteratively subsequent the direction of steepest descents, as demonstrated by the gradients negative. It can be employed for updating the model parameter. Conversely, the DBN is BP-based fine-tuned, and FNN takes several benefits from supervised learning infrastructure. A most important benefit is that it improves the weight betwixt the rule and resultant layers by employing the BP technique one time, and the parameter present from the membership layer is also improved one time. Specifically, the BP technique in FNN could not be controlled in a repeating approach that effectively eliminates the gradient diffusion issue of DBN fine-tuned.

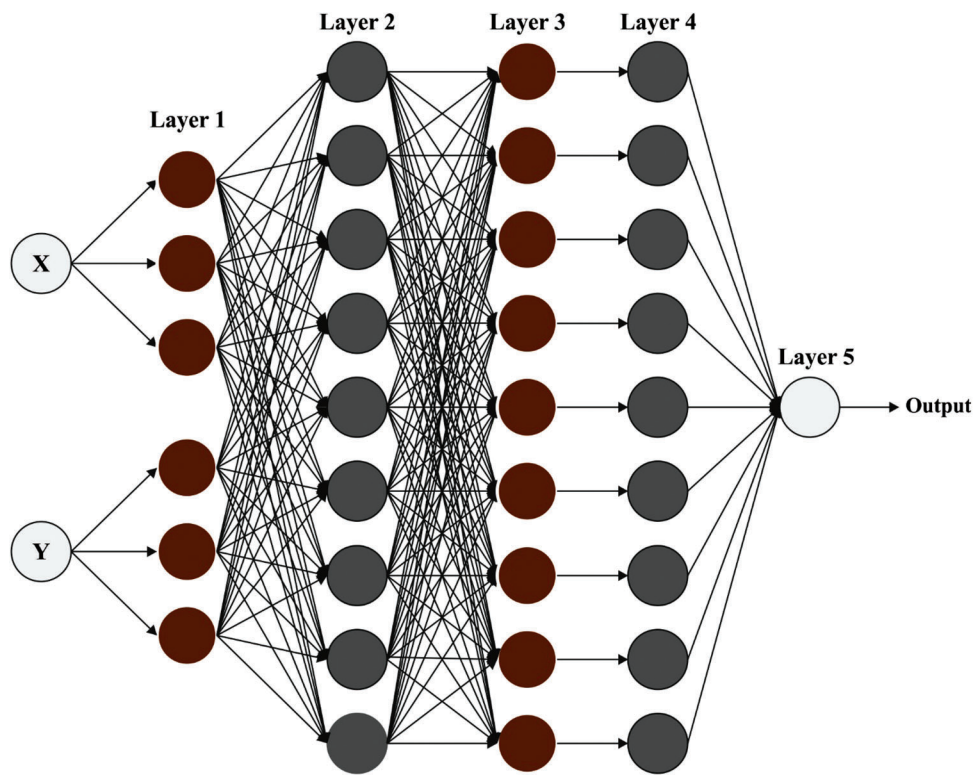


Figure 2: Structure of FNN

4 Results and Discussion

The face mask detection results of the DLFMD-RMG model are investigated using the dataset [22] comprising 1000 images described in Table 1. Fig. 3 illustrates some sample images with and without masks. The proposed model is simulated using Python 3.6.5 tool. The proposed model experiments on PC i5-8600k, GeForce 1050Ti 4 GB, 16 GB RAM, 250 GB SSD, and 1 TB HDD. The parameter settings are learning rate: 0.01, dropout: 0.5, batch size: 5, epoch count: 50, and activation: ReLU.

Table 1: Dataset details

Class	No. of sample images
With mask	500
Without mask	500
Total number of images	1000



Figure 3: Sample images

The confusion matrices accomplished by the DLFMD-RMG model on the identification of face masks on mass gathering in Fig. 4. The figures represented that the DLFMD-RMG model has categorized face mask images appropriately.

Table 2 offers face mask recognition outcomes of the DLFMD-RMG technique on 80% of training (TR) datasets and 20% testing (TS) data. Fig. 5 exhibits the quick face mask recognition results of the DLFMD-RMG system on 80% of the TR dataset. The DLFMD-RMG model has recognized images ‘with mask’ by $accu_y$ of 97.88%, $prec_n$ of 97.97%, $reca_l$ of 97.72%, F_{score} of 97.85%, and MCC of 95.75%. Moreover, the DLFMD-RMG technique has identified images ‘without mask’ by $accu_y$ of 97.88%, $prec_n$ of 97.78%, $reca_l$ of 98.02%, F_{score} of 97.90%, and MCC of 95.75%.

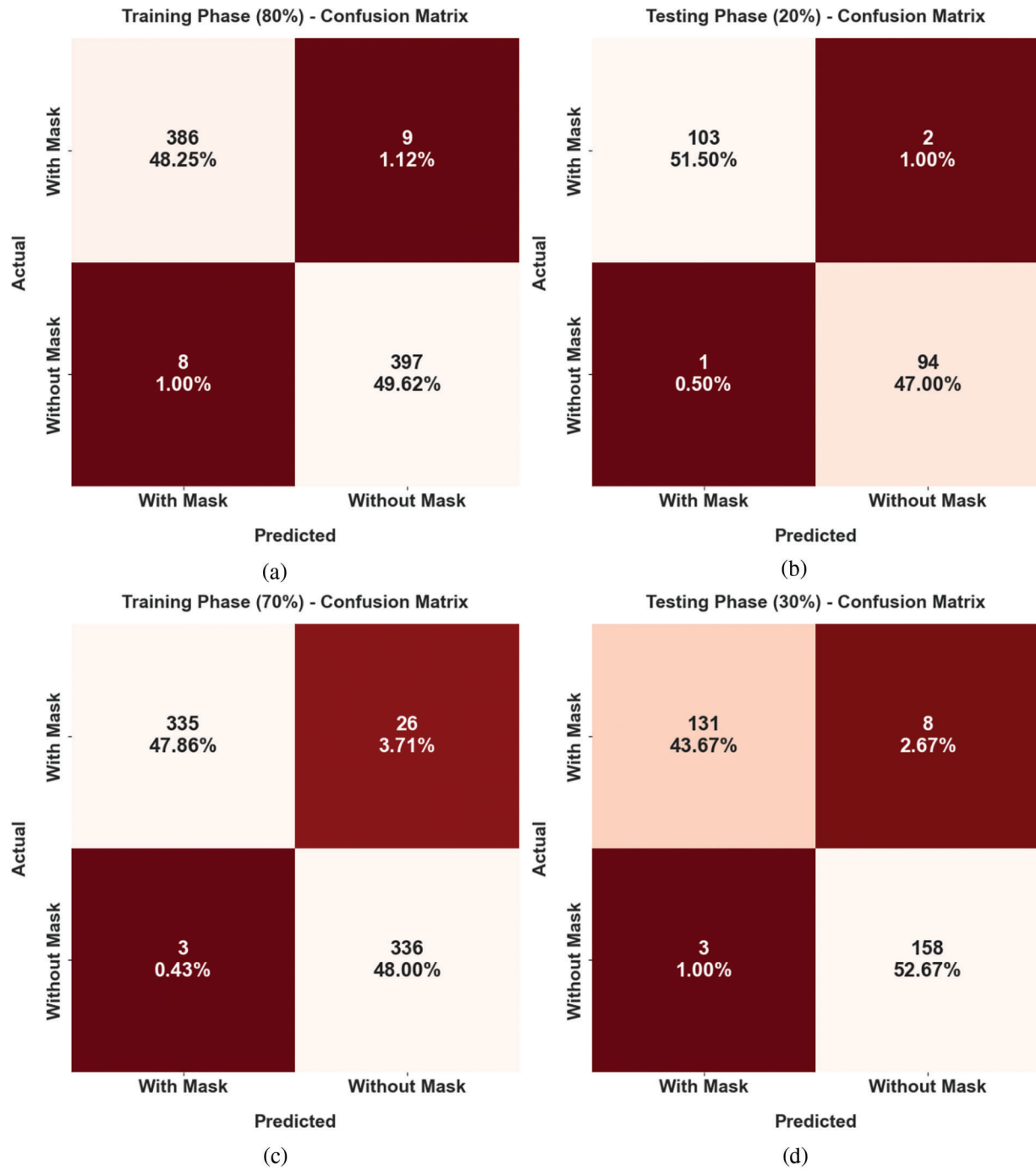


Figure 4: Confusion matrices of DLFMD-RMG approach (a) 80% of TR dataset, (b) 20% of TS data, (c) 70% of TR data, and (d) 30% of TS data

Table 2: Face mask recognition results of DLFMD-RMG methodology under 80:20 of TR and TS dataset

Training/Testing (80:20)					
Class	Accuracy	Precision	Recall	F-score	MCC
Training phase					
With mask	97.88	97.97	97.72	97.85	95.75
Without mask	97.88	97.78	98.02	97.90	95.75
Average	97.88	97.88	97.87	97.87	95.75
Testing phase					
With mask	98.50	99.04	98.10	98.56	97.00
Without mask	98.50	97.92	98.95	98.43	97.00
Average	98.50	98.48	98.52	98.50	97.00

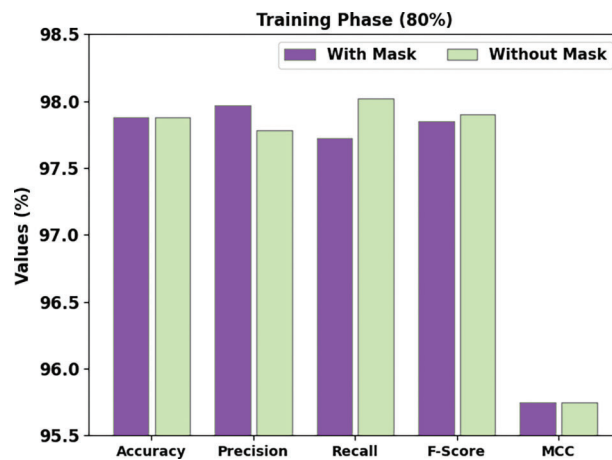
**Figure 5:** Face mask recognition result of DLFMD-RMG approach under 80% of TR dataset

Fig. 6 shows a brief face mask detection outcome of the DLFMD-RMG method on 20% of TS data. The DLFMD-RMG method has identified images ‘with mask’ by $accu_y$ of 98.50%, $prec_n$ of 99.04%, $reca_l$ of 98.10%, F_{score} of 98.56%, and MCC of 97%. Furthermore, the DLFMD-RMG method has identified images ‘without mask’ by $accu_y$ of 98.50%, $prec_n$ of 97.92%, $reca_l$ of 98.95%, F_{score} of 98.43%, and MCC of 97%.

Table 3 offers face mask recognition outcomes of the DLFMD-RMG system on 70% of TR and 30% of TS datasets. Fig. 7 shows a brief face mask detection outcome of the DLFMD-RMG system on 70% of TR data. The DLFMD-RMG system has identified images ‘with mask’ by $accu_y$ of 95.86%, $prec_n$ of 99.11%, $reca_l$ of 92.80%, F_{score} of 95.85%, and MCC of 91.92%. Furthermore, the DLFMD-RMG method has identified images ‘without a mask’ by $accu_y$ of 95.86%, $prec_n$ of 92.82%, $reca_l$ of 99.12%, F_{score} of 95.86%, and MCC of 91.92%.

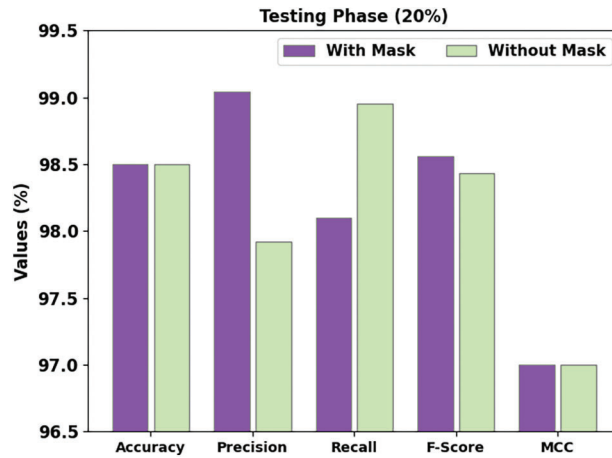


Figure 6: Face mask recognition result of DLFMD-RMG approach under 20% of TS dataset

Table 3: Face mask recognition results of DLFMD-RMG approach under 70:30 of TR and TS dataset

Training/Testing (70:30)					
Class	Accuracy	Precision	Recall	F-score	MCC
Training phase					
With mask	95.86	99.11	92.80	95.85	91.92
Without mask	95.86	92.82	99.12	95.86	91.92
Average	95.86	95.97	95.96	95.86	91.92
Testing phase					
With mask	96.33	97.76	94.24	95.97	92.66
Without mask	96.33	95.18	98.14	96.64	92.66
Average	96.33	96.47	96.19	96.30	92.66

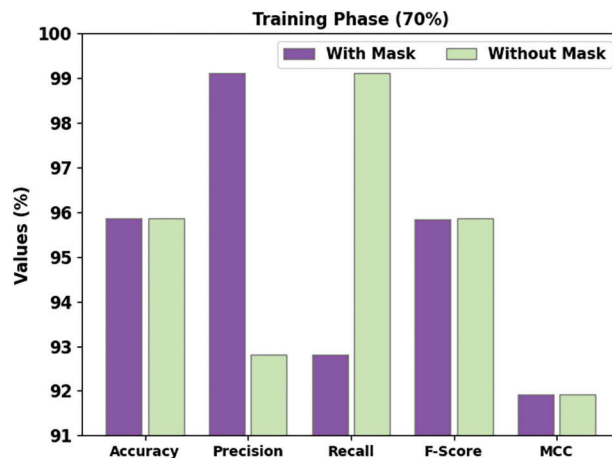


Figure 7: Face mask recognition result of DLFMD-RMG approach under 70% of TR dataset

Fig. 8 demonstrates a brief face mask detection outcome of the DLFMD-RMG system on 30% of TS data. The DLFMD-RMG method has identified images ‘with mask’ by $accu_y$ of 96.33%, $prec_n$ of 97.76%, $reca_l$ of 94.24%, F_{score} of 95.97%, and MCC of 92.66%. Furthermore, the DLFMD-RMG technique has identified images ‘without mask’ by $accu_y$ of 96.33%, $prec_n$ of 95.18%, $reca_l$ of 98.14%, F_{score} of 96.64%, and MCC of 92.66%.

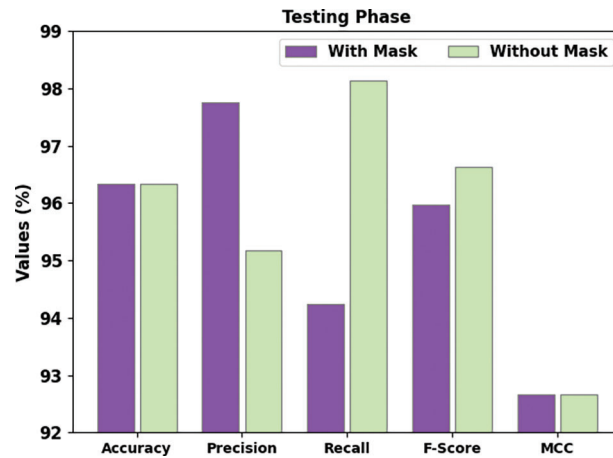


Figure 8: Face mask recognition result of DLFMD-RMG approach under 30% of TS dataset

A brief ROC investigation of the DLFMD-RMG method under the test database is illustrated in Fig. 9. The results indicated the DLFMD-RMG system had shown its capacity to classify different classes.

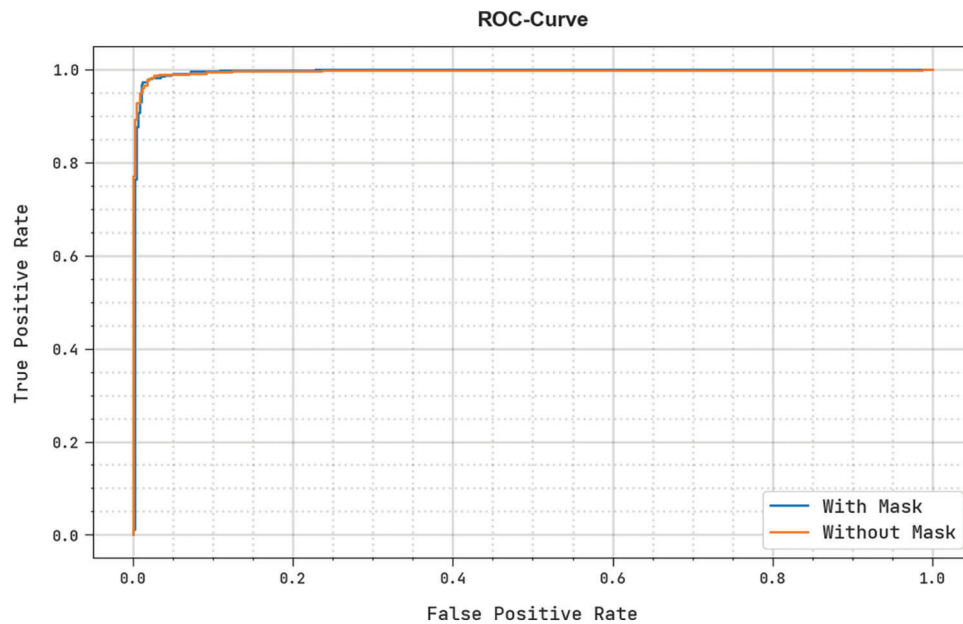


Figure 9: ROC curve analysis of the DLFMD-RMG model

Eventually, a comparative study of the DLFMD-RMG with recent approaches is given in Table 4 and Fig. 10 [12,16,23]. These results inferred that the DLFMD-RMG model had shown enhanced outcomes over other models. For instance, based on $accu_y$, the DLFMD-RMG model has attained an increased $accu_y$ of 98.50%. Besides, based on $prec_n$, the DLFMD-RMG model has achieved an improved $prec_n$ of 98.48%. Also, based on $reca_l$, the DLFMD-RMG model has accomplished an improved $reca_l$ of 98.52%. Therefore, the DLFMD-RMG model can be employed for face mask recognition in religious mass gatherings.

Table 4: Comparative analysis of the DLFMD-RMG system with recent methodologies

Methods	Accuracy	Precision	Recall	F-score
DLFMD-RMG	98.50	98.48	98.52	98.50
MobileNetV2-SVM	97.36	95.28	97.46	96.75
VGG19-K-NN	96.63	97.70	94.66	95.71
MobileNetV2-K-NN	95.06	94.07	95.40	95.21
Xception-SVM	94.51	96.99	97.35	96.92
DT3	96.93	97.92	94.54	97.91
LLE-CNN	78.06	95.84	96.61	95.79
YOLOv2-ResNet	81.21	95.70	97.27	94.42

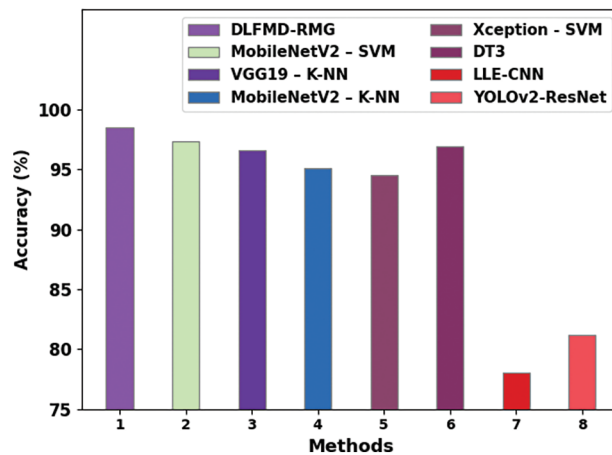


Figure 10: Comparative analysis of DLFMD-RMG system with recent methodologies

5 Conclusion

In this study, new DLFMD-RMG technology has been developed. The DLFMD-RMG technique focuses mainly on detecting face masks in religious mass gatherings. To accomplish this, the presented DLFMD-RMG technique undergoes two levels of pre-processing: BF technique and Contrast Enhancement. For face detection, the DLFMD-RMG technique uses YOLOv5 with a ResNet-50 detector. In addition, the face detection performance can be improved by the SOA for tuning the hyperparameter of the ResNet-50 model. At last, the faces with and without masks are classified using the FNN model. The experimental validation of the DLFMD-RMG algorithm is examined on a benchmark dataset, and

the results highlight the promising performance of the DLFMD-RMG technique over other recent approaches. Ensemble deep learning models can be employed for improved face mask recognition results. Besides, a hybrid metaheuristic algorithm can be designed to enhance classification performance.

Funding Statement: This work was funded by the Deanship of Scientific Research (DSR), King Abdulaziz University (KAU), Jeddah, Saudi Arabia, under grant no. (HO: 023-611-1443). The authors, therefore, gratefully acknowledge DSR technical and financial support.

Conflicts of Interest: The authors declare they have no conflicts of interest to report regarding the present study.

References

- [1] M. Razavi, H. Alikhani, V. faza, 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, pp. 1–8, 2022.
- [2] S. Hussain, Y. Yu, M. Ayoub, A. Khan, R. Rehman *et al.*, "IoT and deep learning based approach for rapid screening and face mask detection for infection spread control of COVID-19," *Applied Sciences*, vol. 11, no. 8, pp. 3495, 2021.
- [3] C. J. Worby and H. H. Chang, "Face mask use in the general population and optimal resource allocation during the COVID-19 pandemic," *Nature Communications*, vol. 11, no. 1, pp. 1–9, 2020.
- [4] H. Farman, T. Khan, Z. Khan, S. Habib, M. Islam *et al.*, "Real-time face mask detection to ensure covid-19 precautionary measures in the developing countries," *Applied Sciences*, vol. 12, no. 8, pp. 3879, 2022.
- [5] J. S. Talahua, J. Buele, P. Calvopiña and J. Varela-Aldás, "Facial recognition system for people with and without face mask in times of the covid-19 pandemic," *Sustainability*, vol. 13, no. 12, pp. 6900, 2021.
- [6] F. Crespo, A. Crespo, L. M. S. tinez, D. H. P. Ordóñez and M. E. M. Cayamcela, "A computer vision model to identify the incorrect use of face masks for covid-19 awareness," *Applied Sciences*, vol. 12, no. 14, pp. 6924, 2022.
- [7] 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, pp. 5236, 2020.
- [8] M. K. I. Rahmani, F. Taranum, R. Nikhat, M. R. Farooqi and M. A. Khan, "Automatic real-time medical mask detection using deep learning to fight covid-19," *Computer Systems Science & Engineering*, vol. 42, no. 3, pp. 1181–1198, 2022.
- [9] N. Dilshad, A. Ullah, J. Kim, and J. Seo, "LocateUAV: Unmanned aerial vehicle location estimation via contextual analysis in an IoT environment," *Internet of Things Journal*, pp. 1, 2022. <https://doi.org/10.1109/JIOT.2022.3162300>.
- [10] S. Alsubai, M. Hamdi, S. Abdel-Khalek, A. Alqahtani and A. Binbusayyis *et al.*, "Bald eagle search optimization with deep transfer learning enabled age-invariant face recognition model," *Image and Vision Computing*, vol. 126, pp. 104545, 2022.
- [11] P. Nagrath, R. Jain, A. Madan, R. Arora, P. Kataria *et al.*, "SSDMNV2: A real time DNN-based face mask detection system using single shot multibox detector and mobilenetv2," *Sustainable Cities and Society*, vol. 66, pp. 102692, 2021.
- [12] 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, pp. 108288, 2021.
- [13] S. Gupta, S. V. N. Sreenivasu, K. Chouhan, A. Shrivastava, B. Sahu *et al.*, "Novel face mask detection technique using machine learning to control COVID'19 pandemic," *Materials Today: Proceedings*, 2021. <https://doi.org/10.1016/j.matpr.2021.07.368>.

- [14] S. Asif, Y. Wenhui, Y. Tao, S. Jinhai and K. Amjad, "Real time face mask detection system using transfer learning with machine learning method in the era of COVID-19 pandemic," in *2021 4th Int. Conf. on Artificial Intelligence and Big Data (ICAIBD)*, Chengdu, China, pp. 70–75, 2021.
- [15] B. Varshini, H. R. Yogesh, S. D. Pasha, M. Suhail, V. Madhumitha *et al.*, "IoT-Enabled smart doors for monitoring body temperature and face mask detection," *Global Transitions Proceedings*, vol. 2, no. 2, pp. 246–254, 2021.
- [16] M. Loey, G. Manogaran, M. H. N. Taha and N. E. M. Khalifa, "Fighting against COVID-19: A novel deep learning model based on YOLO-v2 with ResNet-50 for medical face mask detection," *Sustainable Cities and Society*, vol. 65, pp. 102600, 2021.
- [17] 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, 2021.
- [18] W. Wu, H. Liu, L. Li, Y. Long, X. Wang *et al.*, "Application of local fully convolutional neural network combined with YOLO v5 algorithm in small target detection of remote sensing image," *PLoS One*, vol. 16, no. 10, pp. e0259283, 2021.
- [19] L. Wen, X. Li and L. Gao, "A transfer convolutional neural network for fault diagnosis based on ResNet-50," *Neural Computing and Applications*, vol. 32, no. 10, pp. 6111–6124, 2020.
- [20] K. B. Maji, B. P. De, R. Kar, D. Mandal and S. P. Ghoshal, "CMOS analog amplifier circuits design using seeker optimization algorithm," *IETE Journal of Research*, vol. 68, no. 2, pp. 1376–1385, 2022.
- [21] J. Fei and Z. Feng, "Fractional-order finite-time super-twisting sliding mode control of micro gyroscope based on double-loop fuzzy neural network," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 51, no. 12, pp. 7692–7706, 2020.
- [22] <https://github.com/prajnasb/observations/tree/master/experiments/>.
- [23] A. Oumina, N. El Makhfi and M. Hamdi, "Control the covid-19 pandemic: Face mask detection using transfer learning," in *2020 IEEE 2nd Int. Conf. on Electronics, Control, Optimization and Computer Science (ICECOCS)*, Kenitra, Morocco, pp. 1–5, 2020.