Learning Discriminant Spatial Features With Deep Graph-Based Convolutions for Occluded Face Detection

FIRAS ALBALAS[1], **AHMAD ALZU'BI**[1], (Member, IEEE), **ALANOUD ALGUZO**[1], **TAWFIK AL-HADHRAMI**[2], **AND ACHRAF OTHMAN**[3], (Senior Member, IEEE)

[1]Department of Computer Science, Jordan University of Science and Technology, Irbid 22110, Jordan
[2]School of Science and Technology, Nottingham Trent University, Nottingham NG11 8NS, U.K.
[3]Mada Center, Doha, Qatar

Corresponding author: Ahmad Alzu'bi (agalzubi@just.edu.jo)

ABSTRACT The use of face masks has become a widespread non-pharmaceutical practice to mitigate the transmission of COVID-19. However, achieving accurate facial detection while people wear masks or similar face occlusions is a major challenge. This paper introduces a model to detect occluded or masked faces based on fused convolutional graphs. This model includes a deep neural architecture with two spatial-based graphs that rely on a set of key facial features. First, a distance graph is used to identify geographical similarity between the facial nodes that represent certain key face parts. Second, a correlation graph is formulated to compute the correlations between every two nodes that represent two different augmented facial modalities. Transfer learning is then performed using a pretrained deep architecture as a baseline to map the abstract semantic information into multiple feature filters. Then, discriminant graph convolutions are constructed based on the fusion of distance and correlation graphs. This model evaluates two tasks of facial detection, which are the binary detection of masked or unmasked faces, and multi-category detection of masked, unmasked, or occluded face with no mask. The experimental results on two benchmarking real-world datasets show that the proposed deep learning model is highly effective with an accuracy of 98% achieved in binary detection. Even with high variance in image occlusions, our proposed model has great promise in detecting and distinguishing between types of facial occlusion with an accuracy of 86% reported in multi-category detection.

INDEX TERMS Correlation graphs, deep learning, distance graph, graph convolutional networks, face mask, occluded face detection, spatial features.

I. INTRODUCTION

Face masks, in combination with physical distancing and vaccination, are one of the most effective preventive measures for slowing down the rapid spread of SARS-CoV-2 (COVID-19) [1] and have become widely adopted as a standard precautionary measure [2]. Though face masks were already used in certain contexts such as healthcare and environments with excessive pollution, the commonplace use of masks by the public in daily life has increased significantly.

Owing to the rising popularity of masks, there is now a demand for accurate, real-time image recognition techniques that can effectively detect both masked and non-masked

The associate editor coordinating the review of this manuscript and approving it for publication was Michele Nappi.

faces for public health and security purposes. Masked faced detection (MFD) aims to address this by using facial recognition algorithms to detect whether a person is wearing a face mask or not. However, there are various shapes and styles of face masks that hide different amounts of a person's cheeks, nose, and mouth. In current literature, most MFD detection algorithms have been devised to detect the existence of masks over face [3], but few has attempted to simultaneously deal with MFD and occluded face detection (OFD) where the face is distracted by a non-mask object. Moreover, automation of the detection process is needed to make adoption of the technologies feasible in real-world security scenarios [4], such as the person authentication at access checkpoints [5].

Face detection technologies are mainly based on artificial intelligence (AI) and machine learning solutions, where many

F. Albalas *et al.*: Learning Discriminant Spatial Features With Deep Graph-Based Convolutions for Occluded Face Detection

IEEE *Access*

computer vision applications, including image classification and recognition [6]–[9], have reported high performance. From a facial recognition perspective, OFD and MFD have the added complication of fewer key facial parts visible to perform identification, such as nose, mouth, and cheeks. Even common effective face detectors, including YOLOv3 [10], RefineFace [11], LLE-CNNs [12], and Faster R-CNN [13], which largely rely on such important face parts, have shown a performance drop when dealing with face masks or occlusions [14]. Moreover, the available datasets for detecting face masks or occlusions are relatively insufficient to train the detection models [3], and it becomes essential to improve the discrimination quality of OFD systems.

In recent years, a large body of research has attempted to address such challenges in the domain of MFD and OFD based on deep learning approaches [12], [15]–[19]. Machine learning models based on deep neural networks, such as convolutional neural networks (CNNs), are superior in face detection and recognition, which utilize many popular pre-trained models [20]–[22].

Graph convolutional networks (GCNs) are among the more effective CNN-based models for identifying key features in various computer vision applications [23]–[28]. Conventional CNNs can be generalized by GCNs using spatial or spectral filters to deal with real data. However, learning graph representations are relatively complex with the restriction on the depth of architecture layers, as well as unfavorable redundant computations [29]. Therefore, there is a lack of research utilising GCN-based models for MFD or OFD. The current paper investigates the capability of GCN-based architectures in extracting discriminant key features to contribute towards detecting masked or occluded faces more accurately.

In this paper, a graph-based deep learning model that is efficiently able to detect faces with/without masks or faces with non-mask occlusions in images is introduced. The proposed model aims to generate a set of key facial features, utilizing two convolutional representations, which are distance and correlation graphs. Each individual spatial filter forms a separate graph notation of facial parts, which in turn provides a fused GCN-based architecture used to represent the final generic image descriptor. This model also benefits from the transfer learning conducted to use the learned kernels trained on general face detection images for the specific MFD/OFD task. Most importantly, the model performance is evaluated in two classification tasks, which are binary and multi-category detection. The main contributions of this work are summarized as follows:

1) A deep learning model is developed using two GCN-based distance and correlation representations to detect faces with masks or occlusions. On one hand, the distance graph indicates a distinct relationship between the facial features with the utilization of weighted distance functions. On the other, the correlation graph indicates the correlation between the spatial features based on the historical usages (i.e., inflow or outflow) in each time interval. These two graphs are then convolved (fused) to generate a multi-graph model that is then used to learn the general discriminative characterizations of occluded faces.

2) Two classification tasks are examined in this work. The first task determines whether the face is covered by face mask or not (i.e., binary detection). The second more challenging task categorizes the test face image into one of three classes: masked, unmasked, or occluded face with no masks.

3) Two large benchmarking datasets have been restructured and manipulated in order to evaluate the performance of the proposed model on a sufficient and diverse collection with a total of 38000 images.

The remaining sections of this paper are organized as follows. Section II presents the recent studies and related works in the domain of MFD and OFD using deep learning techniques. Section III introduces the work methodology and it illustrates the generic pipeline used to construct the graph-based image descriptors. Section IV discusses the experimental results obtained in the binary and multi-category face detection tasks. Finally, Section V concludes this paper with potential future directions.

II. RELATED WORK

Face detection has traditionally been an intricate challenge for computers due to the animated and dynamic nature of faces, but with the technological progress and the development of algorithms in the field of AI and deep learning, face detection has become effectively solvable [30]–[32].

Advances in the field of face detection are indebted to the work of Viola *et al.* [33] which takes a person's photo and compares it with images stored in a database. A shortcoming of this work is that in the case of any deviations from the facial expressions made in the database causes the matching process to be invalidated and subsequently the performance of the existing in-house face detection algorithms to drop noticeably. Research efforts have intensified in this domain to create models more robustly capable of detecting face masks under various circumstances.

Deep learning methods are among the most successful approaches that have been widely examined in the domain of non-mask face detection, including deep features [34], [35], face detection on social media [36], [37], and video-based face detection [38], [39]. Deep learning algorithms allow the artificial neural networks to learn the key facial points and their locations in a comprehensive way [40]. Many sophisticated approaches have adopted the unrestricted scenario for face detection techniques, due to the limitations caused by distortions, exaggerated expressions, or large obstructions, which also made the dataset used to process such scenarios so limited.

With this increasing research work in the field of MFD and OFD, a database containing a large number of images became an essential demand. There are many datasets of masked faces available to the public but few contain sufficient images to

feed and train the deep learning models on diverse images with challenging face occlusions.

However, many image collections have been recently proposed in literature for MFD and MFR. Vu *et al.* [41] proposed combined Local Binary Pattern (LBP) and deep learning features using RetinaFace, and they provided a new dataset called COMASK20. Wang *et al.* [4] provided three types of facemask datasets: Real-world Masked Face Recognition Dataset (RMFRD), Masked Face Detection Dataset (MFDD), and Simulated Masked Face Recognition Dataset (SMFRD). Prasad *et al.* [17] also introduced two datasets consisting of people's images in real natural scenes: MASK-face v1 and MASK-face v2. Liu *et al.* [19] constructed the HKBU-MARs V2+ dataset that consists of high-quality face masks taken under a wide variety of lighting settings, and a large number of related videos. Ge *et al.* [12] introduced the MAFA dataset which consists of 35 806 synthetic masked faces and real images collected from the internet with various orientations and occlusion degrees, in which at least one part of each face is occluded by mask. Also, they proposed LLE-CNNs to recover the missing facial cues by various masks.

Many contributions have responded to the need for effective approaches to detect face masks. However, the existing face detectors typically rely on the uncovered facial parts and have therefore shown a noticeable performance drop when applied to images of masked faces [14]. The following section discusses some recent approaches mainly based on the deep learning algorithms and architectures.

Lightweight face detection systems that run in real-time platforms such as mobile devices have also been developed. Nagrath *et al.* [18] proposed a deep learning model called SSDMNV2 based on MobilenetV2 for real-time mask detection. It has the advantage of using lightweight embedded devices such as Raspberry Pi and NVIDIA Jetson Nano. The model was evaluated on a medical face mask dataset. It outperformed the pre-existing models trained on the same dataset such as LeNet-5, ResNet-50, VGG-16, and AlexNet. Militante *et al.* [42] also built a real-time system using CNN-based VGG16 model to classify persons into masked or unmasked faces, and it triggers a warning alarm using Raspberry Pi to those not wearing facemasks.

GAN-based deep learning systems have also been utilised in many MFD applications. Din *et al.* [43] developed a GAN-based network to unmask the detected face masks and synthesize the missing facial parts with fine details and regions reconstruction. Two face discriminators were used: one for learning the global structure of the person's face, and one for learning the missing regions. A set of synthetic images is created using the CelebA dataset. Luo *et al.* [44] developed the EyesGAN model which is based on the composition of people's faces from their eyes. Two techniques were used to build this model, which are the perceptual loss and the self-attentional mechanism in GANs. Huang *et al.* [45] also used the Resnet50 classifier as a deep-learning framework for building a masked face recognition model that was evaluated on a new proposed Webface-OCC dataset.

Video-based detection approaches have also been introduced to detect or recognize face with masks or occlusions. Snyder *et al.* [46] developed a system to classify persons into masked or unmasked faces using a CNN-based model, which includes three steps: detecting human subjects in videos using a ResNet-50 model with a hierarchical feature network (FPN), extracting human faces from videos using multitasking convolutional neural networks (MT-CNN), and classifying faces into masked and unmasked by training a CNN classifier. A real data set was collected from educational settings. Joshi *et al.* [47] also presented a real-time video-based system to detect the face mask using MTCNN face detection model with MobileNetV2. Mandal *et al.* [48] used the Resnet50 classifier as a deep-learning framework for masked face recognition evaluated on the RMFRD dataset. Sethi *et al.* [49] examined two-stage deep-learning detectors using MobilenetV2, Resnet50, and AlexNet models for detecting face masks in real-time with different images. Draughon *et al.* [50] also presented a CNN-based system to track people's movement in public areas and detect any faces with occluded parts.

Other research efforts have adopted a hybrid approach that uses aspects of both traditional machine learning and newer deep learning models. Loey *et al.* [51] developed a hybrid system that detects the facemask by extracting image features using Resnet50 and classifies the facemask using SVM, decision trees, and ensemble algorithms. The model was evaluated on RMFRD and LFW datasets. Oumina *et al.* [52] proposed a hybrid system combining traditional machine learning algorithms with deep learning models to detect the face mask. It extracts image features using MobileNetV2, VGG19, and Xception then classifies the face mask using K-NN and SVM. The SVM classifier with MobileNet-V2 model reported the best detection accuracy.

Even though GCNs have shown superiority in various computer vision application, few studies have investigated the performance of GCNs in MFD and OFD tasks. Ren *et al.* [53] introduced a dynamic graph representation to adaptively remove the nodes representing the occluded parts in biometric including faces and iris images. The proposed model applies dynamic graphs matching with various distance measures and adjacent matrices. Ye *et al.* [54] proposed a deep learning model based on GCNs for detecting face masks. It extracts image features using DenseNet101 then recognizes and classifies the facemask using GCN on MAFA dataset. Recently, Alguzo *et al.* [55] also proposed a deep learning model based on multi-GCNs to determine and detect people wearing masks. This model produces a multi-graph structure using convolutional filters that utilizes a 4D facet tensor as an input function, and it includes a convergence layer to capture multiple facial expressions. The real-world masked face dataset (RWMFD) was used to evaluate the proposed model.

This paper introduces a new deep learning model based on the fusion of two spatial graphs with several filters constructed effectively to generate a discriminating generic
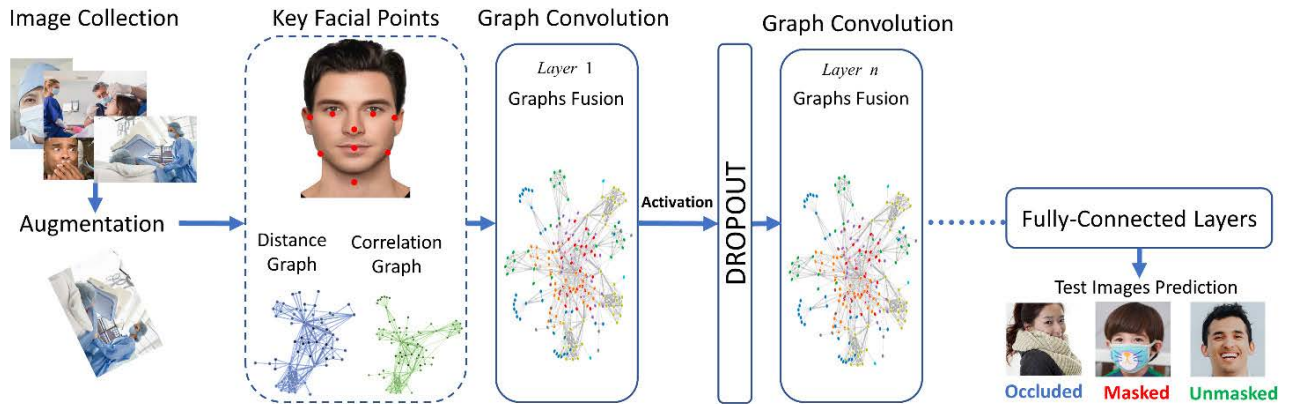
F. Albalas *et al.*: Learning Discriminant Spatial Features With Deep Graph-Based Convolutions for Occluded Face Detection

IEEE *Access*

FIGURE 1. The main phases of the proposed OFD pipeline.

image representation. The RWMFD dataset is also replaced with two alternative datasets, MAFA and Human Faces, which are more sufficient in terms of size and diversity. Additionally, the work in [55] only determines whether the detected face is covered by a mask or not, while in this paper we perform both the binary OFD and multi-category OFD (i.e., unmasked, masked and occluded).

III. METHODOLOGY

A. DEEP ARCHITECTURE PIPELINE

In this section, the deep learning model used to detect the face-mask based on two convolution-based graphs is introduced. As shown in Figure 1, this model pipeline involves several phases, as will be detailed in the following sub-sections. The dataset images are first prepared using two benchmarking collections preprocessed to fit the architecture requirements with data augmentation to increase the diversity of challenging and confusing images. The proposed architecture includes several graph convolutional layers that include the graph convolution fused using the correlation and distance graphs, which represent multiple filters with shared weights. The baseline architecture is initialized with the pretrained VGG16 model [21] to perform an efficient transfer learning procedure. Specifically, the weights learnt in the VGG16 model are used in the training process to obtain better generalization capability while learning the features of the new domain-specific images (i.e., masked and occluded faces) especially in the lower layers. This architecture also includes several top layers (e.g., dropout, dense, and fully connected layers), which end with a prediction layer to classify the input images.

B. DATASET PREPARATION

1) MAFA DATASET

The MAFA dataset [12] is one of the largest datasets contributing to the development of the MFD algorithms. It consists of 30 811 facial images collected from various websites such as Google, Bing and Flickr with different degrees of

FIGURE 2. Image samples from (a) MAFA data and (b) HFD data.

occlusion. Each image contains at least one face wearing a mask. MAFA is characterized by a variety of degrees of occlusion, types of face masks, textures, as well as a large proportion of hand-covered faces. MAFA can be grouped into two categories: faces covered with a typical face mask (i.e., medical or colored), and faces covered by hands, scarves, or any other objects impeding its identification. Figure 2(a) shows a set of image samples from the MAFA dataset.

2) HUMAN FACES DATASET

The Human Faces Dataset (HFD) [56] consists of about 7 200 images collected from the internet. It also contains few GAN generated 'fake faces' to challenge the functionality of dealing with real and generated faces. The dataset is characterized by its diversity of ethnicity, gender, and ages, with many images of elderly people. In our experiments, the HFD was used in the training phase to train the model on the fully-featured exposed face. Figure 2(b) shows a set of the HFD samples.

3) DATASET SPLIT

The dataset used to train, test, and validate the proposed model in the binary detection (i.e., unmasked or masked faces) consists of 29 209 images for training, 7 557 images for testing, and 1 245 images for validation (77%/20%/3%). The second task is to classify input images into one of three classes: masked, unmasked, and occluded faces. The MAFA

IEEE Access

F. Albalas *et al.*: Learning Discriminant Spatial Features With Deep Graph-Based Convolutions for Occluded Face Detection

TABLE 1. A description of images split used for training and testing.

Task	Class	Train split	Validate split	Test split	Total
Binary	Masked	25,159	714	4,938	30,811
	Unmasked	4,050	531	2,619	7,200
Multi-class	Masked	18,289	500	3,500	22,289
	Unmasked	5,223	500	1,477	7,200
	Occluded	5,895	500	2,127	8,522

and HFD were manually labelled, to train, test, and validate the proposed model. As result, it consists of 29 407 images for training, 7 104 images for testing, and 1 500 images for validation (77%/19%/4%). Table 1 summarizes the classes and number of images set for the binary and multi-category detection and classification tasks. The unmasked images are from the HFD dataset and the remaining images are all the MAFA masked and occluded face images.

4) IMAGE PREPROCESSING

Image data augmentation is a common and effective technique used in data preprocessing to artificially create modified versions of the images, such as flipping, scaling, rotation, and shearing. One of its most important benefits is to improve the generalization capability of the trained model. We have applied three data augmentation operations to the input images, which are rescaling image by a factor 1/255 to obtain a range of [0-1], generating a random brightness in the range [0.3-1.5] where the values above 1.0 brighten the images and values less than 1.0 darken the images, and rotating images randomly in the range [0-360] with a rotation range of 30.

C. GRAPH CONVOLUTIONAL REPRESENTATION

Given a graph $G(V, E)$ where V is a set of vertices and E is a set of edges, let $v_i \in V$ indicate a node and $e_{ij} = (v_i, v_j \in E)$ indicate an edge pointing from v_i to v_j. A neighborhood of a node v is defined as $N_{(v)} = u \in V | (u, v) \in E$. An adjacency matrix A is defined as $n \times n$ matrix with $A_{ij} = 0$ if $e_{ij} \notin E$ and $A_{ij} = 1$ if $e_{ij} \in E$. Each graph contains the node attributes X, where $X \in R^{n \times d}$ is a node feature matrix and $x_v \in R^d$ represents the features vector of node v. In addition, a graph has edge attributes X^e, where $X^e \in R^{m \times c}$ is a matrix of edge features with $x_{vu} \in R^c$ representing the features vector of edge (u, v).

For the task of face mask detection, the goal is to analyse the images through a graph capable of analysing the spatial details and extracting the best discriminant features from images. To this end, we introduce a graph-based convolutional network that contains three main steps: graph generation, graph fusion, and graph convolution.

1) GRAPH GENERATION

Graph construction is essential to develop a successful graph convolutional model. This is achieved based on the important relationships between the nodes. To capture the diverse relationships among nodes, we construct two subgraphs, which are the distance graph and correlation graph.

For the distance graph, each node is connected to its neighbors by edge, whereby close nodes indicate a distinct relationship which is based on the geographical law that states 'everything is related to everything else, but near things are more related than distant things'. This highlights the importance of determining the distance between the nodes in order to build the graph based on the distance between these nodes. To determine the weight between two nodes, we use the reciprocal of the distance so that closer pixels will be linked with higher weights. Two types of distance functions are utilized, which are a power-law function and a linear function [57]. The distance graph is defined as follows:

$$G_d(V, E)_{weight} = Distance^{-1} \quad (1)$$

For the correlation graph, the historical usages (outflow or inflow) of each node are calculated in each time interval and the correlations between every two nodes are then computed as the inter-node link weights in the graph. We use the Pearson coefficient to calculate the correlation [58], [59]. The co-relation graph is defined as follows:

$$A_c = \begin{pmatrix} 0 & d_{0,1} & \cdots & d_{0,N-1} \\ r_{1,0} & 0 & \cdots & r_{1,N-1} \\ \vdots & \vdots & \ddots & \vdots \\ d_{N-1,0} & \cdots & \cdots & 0 \end{pmatrix} \quad (2)$$

where $r_{i,j}$ denotes the Pearson correlation between I and J.

2) GRAPHS FUSION

A two-graph convolutional layer is used to fully exploit different graphs that contain heterogenous useful spatial information. Our model conducts graph fusion then graph convolution using the formulated distance and correlation graphs.

Firstly, both graphs are merged into one graph by calculating the weighted summation of their adjacency matrices extracted from each image. Because adjacency matrices of graphs are located in different ranges, the adjacency matrix A for each graph is normalized as follows:

$$\hat{A} = D^{-1}A + I \quad (3)$$

where I is an identity matrix and D is a diagonal matrix calculated as follows:

$$D = \begin{pmatrix} \sum_{J=0}^{N-1} A_{0,j} & 0 & \cdots & 0 \\ 0 & \sum_{J=0}^{N-1} A_{i,j} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sum_{J=0}^{N-1} A_{N-1,j} \end{pmatrix} \quad (4)$$

The results from Equation (4) are a normalized adjacency matrix for each image with a self-loop in which the self-loop

F. Albalas *et al.*: Learning Discriminant Spatial Features With Deep Graph-Based Convolutions for Occluded Face Detection

IEEE *Access*

maintains the information of the node in the convolution part, a key design strategy in graph neural networks [57]. Then, the fusion result is further normalized by adding a softmax operation to the weight matrix. Specifically, if we have N graphs to blend together, the fused graph can be defined as follows:

$$F = \sum_{i=1}^{N} W_i^l \circ A_i^l \tag{5}$$

where W_i^l is weight matrix.

3) GRAPHS CONVOLUTION
Now, we perform the convolution operation based on the fused graph F denoted as $\left[H_l^1, H_l^2, \ldots, H_l^k\right]$, where H is the hidden state. The convolution layer is then applied to each segment to produce a sequence of feature matrices denoted as $\left[Z_l^1, Z_l^2, \ldots, Z_l^K\right]$. These feature matrices are fed into the convolutional layer chronologically. Finally, we take the output feature matrix of the last hidden state H as the output of the convolutional layer.

Accordingly, any input images are fed to the input layer then to the first multigraph layer. This layer generates multiple distinct graphs for image matrix and selects the graph with the optimal features that represent the facial image parts. Then, the architecture includes a dropout layer to drop the graph features with no key discriminating values (i.e., values approaching zero). The output of the first multigraph is passed to the next multigraph layer then again to a dropout layer, and so forth.

In the top layers, a dense layer is added to reduce the dimension of node features to a smaller size which reduces the training and testing run time and complexity. Finally, the last softmax layer uses this final descriptor of graph features to predict the test images and labels them with masked or unmasked faces in the binary task and masked, unmasked or occluded faces in the multi-category detection. It is worth mentioning that we use the ELU activation function in the low graph convolutional layers and the ReLU activation function in the top fully connected layers. This model is trained end-to-end to generate the optimal learning weights through a set of training epochs and image batches.

IV. RESULTS AND DISCUSSION
A. EVALUATION SETUPS
All experiments were carried out using Python as a development language and a set of rich libraries including TensorFlow and Keras as a development platform with a powerful GPU processor. Table 2 lists the hyperparameters used in the model training and testing. The best parameter listed in this table were empirically chosen, where the number of nodes was set to 100 but finally reduced to keep the most informative nodes and relations. Also, the number of epochs were initially set to 50 but an early stopping approach was used to stop training the model when the performance becomes stabilized.

TABLE 2. The list of hyperparameters used in our experiments.

Hyperparameter	Value
Dropout	0.2
Loss function	Cross-entropy
Optimizer	Adam
Batch size	64
Initial graph nodes	100
Learning rate	1e-4
Verbose	1
Epochs	50

Given that the model prediction is finally determined by binary and multi-classification, we have utilized four different measures to evaluate the performance of multi-graph deep model as follows: *True Positives* (TP) indicates that the model predicted the class of image as (Mask) and it is actually (Mask); *False Positives* (FP) indicates that the model predicted the class of image as (unmasked) and it is actually (Mask); *True Negatives* (TN) indicates that the model predicted the class of image as (unmasked) and it is actually (unmasked); and *False Negatives* (FN) indicates that the model predicted the class of image as (Mask) and it is actually (unmasked). Additionally, the model performance is also measured by a set of standard metrics, including *recall*, *precision*, *accuracy*, and *F1-score*.

B. EXPERIMENTAL RESULTS
In this section, we discuss the experimental results of the proposed multi-graph deep model for the binary and multi-class detection tasks.

1) BINARY DETECTION
Figure 3 demonstrates the training and validation results of the model for the binary detection task. As can be observed in Figure 3(b), the accuracy increases smoothly to reach about 99% on the training phase, while it decreases slightly on the validation phase to reach approximately 98%. As a result, the model maintains high training and validation accuracy, which confirms that there is no overfitting or underfitting that could affect the model generalization ability. Moreover, the model can quickly learn and converge in a small amount of training epochs. Similarly, in Figure 3(a), the misclassification rates denoted by the loss function are very low in the training phase which is around 0.016, while in the validation phase reaches around 0.057.

The proposed model generated over 2 million nodes to represent and select the optimal discriminating descriptor of facial key features, which in turn gained an overall accuracy of 98%. Additionally, the detector accuracy is high in terms of precision, recall and *F1*-score, which shows the capability of our model in detecting the masked or unmasked faces. The results confirm the high capability of our proposed model in predicting masked faces in the test dataset which achieved 0.98, 1.00, and 0.99 for precision, recall, and *F1*-score, respectively. Similarly, it achieves a high accuracy in predicting the unmasked faces in the test dataset which achieved 0.99, 0.95, and 0.97 precision, recall, and *F1*-score,
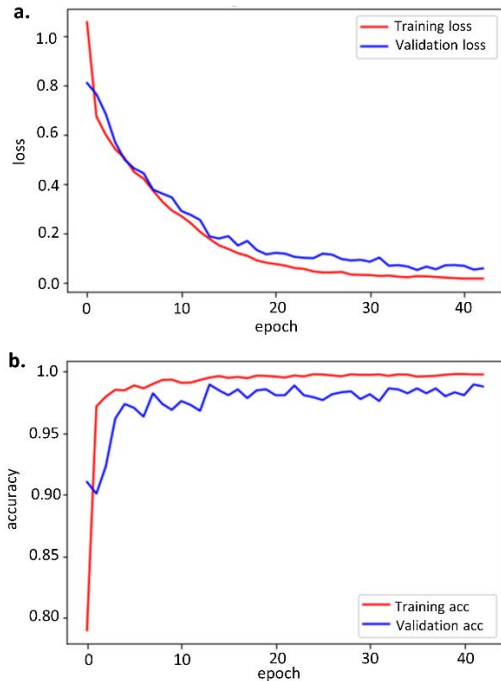
IEEE *Access*

F. Albalas *et al.*: Learning Discriminant Spatial Features With Deep Graph-Based Convolutions for Occluded Face Detection

FIGURE 3. The results of (a) loss and (b) accuracy rates for the binary detection task during in training phase.

TABLE 3. The performance measures of binary classifier.

	Precision	Recall	F1-score
Masked Faces	0.98	1.00	0.99
Unmasked Faces	0.99	0.95	0.97
Weighted Avg.	0.98	0.98	0.98

respectively with a total of 7,557 supported images. Therefore, the model is able to accurately detect masks even with different colors, styles, shapes and coverage areas.

We conclude that the model can accurately detect masks with an accuracy of 99% achieved in the training phase and an accuracy of 98% achieved in the testing phase.

A summary of the performance results calculated from the confusion matrix and achieved by the binary classifier are shown in Table 3.

2) MULTI-CATEGORY DETECTION

Figure 4 demonstrates the training and validation results of the model for multi-category detection (i.e., masked, unmasked, and occluded faces). As can be observed in Figure 4(b), the accuracy is high in the training phase, but drops down in the validation phase to approximately 85.19%. This accuracy variance indicates a slight overfitting even though a cross-validation was applied in the model training, as well as an early stopping rule to guide the learning procedure before it begins to over-fit. However, this is justified in the model because the number of occluded images, selected in the validation phase of the binary and multi-category detectors, was relatively small compared to the number of images available for training the model on the masked faces.

FIGURE 4. The results of (a) loss and (b) accuracy rates for the multi-category detection task in the training phase.

As a result, the model tends to consume additional facial attributes (represented by graph nodes) while learning the features of occluded faces effectively. As we focus on investigating the effectiveness of GCN-based deep features in detecting and classifying occluded or masks faces, adding sufficient images with occluded faces to the collection generated by the data augmentation applied in our model may help in providing better performance.

Nevertheless, the model can quickly learn and converge in a small amount of training epochs. Similarly, in Figure 4(a), the misclassification rates indicated by the loss function are also low in the training phase, estimated at 0.19, while in the validation phase it reached about 0.4. In the testing phase, the overall accuracy achieved is 0.86, and the model reported 0.90, 0.87, and 0.89 for precision, recall, and *F1*-score, respectively. Similarly, it achieves high accuracy in predicting the unmasked faces for the test dataset which achieved 0.98, 0.91, and 0.95 precision, recall, and *F1*-score, respectively. Also, the results of predicting the occluded faces are 0.65, 0.75, and 0.66 for precision, recall, and *F1*-score, respectively. A summary of the performance results calculated from the confusion matrix and achieved by the multi-category classifier is shown in Table 4.
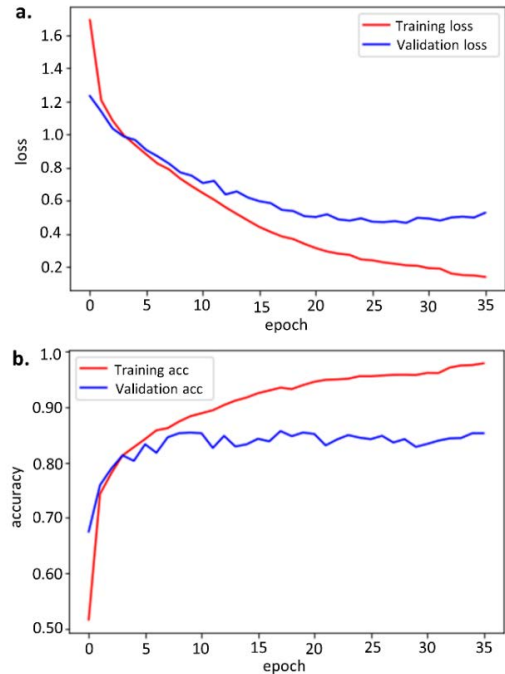
We conclude that the model is able to accurately distinguish between the face mask and any other face occlusions with different coverage areas. Figure 5 shows a set of sample images predicted by the model and classified into masked, occluded or unmasked faces. However, the nature of images with occluded faces in each category have influenced the learning and generalization ability of multi-category classifier. Therefore, predicting the category of the

F. Albalas *et al.*: Learning Discriminant Spatial Features With Deep Graph-Based Convolutions for Occluded Face Detection

IEEE*Access*

TABLE 4. The performance measures of multi-category classifier.

	Precision	Recall	F1-score
Masked Faces	0.90	0.87	0.89
Unmasked Faces	0.98	0.91	0.95
Occluded Faces	0.60	0.75	0.65
Weighted Avg.	0.89	0.86	0.87

FIGURE 5. Sample of test images predicted in the testing phase.

TABLE 5. Comparison with the recent state-of-the-art methods.

Ref.	Detector (dataset)	Accuracy
Prasad *et al.* [17]	MaskedFaceNet (P-MAFA)	95.5
Avanzato *et al.* [60]	YOLOv3 (MAFA+WIDER)	84.0
Vinh *et al.* [61]	YOLOv3 (P-MAFA)	90.1
Zhang *et al.* [62]	R-CNN (P-MAFA)	84.1
Ryumina *et al.* [63]	CNN (MAFA+RMFD)	96.0
Ye *et al.* [52]	GCN (MAFA+MASK)	88.3
Ge *et al.* [12]	LLE-CNNs (MAFA)	76.4
Wang *et al.* [64]	YOLO (MAFA+WIDER)	89.0
This paper	Binary Detector (MAFA+HFD)	98.0
This paper	Multi-class (MAFA+HFD)	86.0

detected face remains challenging due to the high similarities between the face masks and other occlusion objects hiding the facial parts. Providing benchmarking datasets with sufficient diverse types of masks and occlusions, real-world or synthesis images, remains one of the main challenges in the field of MFD and OFD.

3) ACCURACY COMPARISON
Most of the previous related work is based on typical deep learning models such as CNN, R-CNN, YOLOv3, SSD, GCN, DenseNet and LLE-CNNs. However, there are no benchmarking datasets commonly used for evaluating the performance of MFD. Therefore, the use of facial images with masks generated synthetically is one of favourable alternatives. Moreover, the main aim of research works that use MAFA or partial MAFA dataset is to detect masks, where our proposed model is also able to classify the detected face either occluded by a face mask or any other objects. Therefore, we present here only the recent related works evaluated on MAFA, partial MAFA, or merged with other datasets under different setups but with the aim of detecting the mask rather than classifying the face images.

Table 5 summarizes the main characteristics and performance of our GCN-based model and several recent works devoted to the MFD detection task. The performance is shown in terms of achieved accuracy on MAFA dataset or partial MAFA (P-MAFA) where the training and testing images are selected by different approaches in these related works. Our proposed model is distinguished by representing the masked facial features using a multi-graph convolutional network and dealing with two classification tasks, which are binary and multi-category detection. The proposed model shows a comparable accuracy to the existing works with an accuracy of 98% for the binary detection. It also achieves 86% in multi-category detection obtained by a simple yet efficient GCN-based deep architecture.

V. CONCLUSION
This paper introduced a new deep learning model using GCNs to detect faces with masks or occlusions. This architecture was developed with the aim of verifying its ability to extract key features of faces obscured by masks or other elements and to learn generic discriminant descriptors. This model is mainly based on fusing the correlation and distance graphs in a convolutional layer that is then followed by several layers, such as dropout, dense and softmax layers. The training and validation procedures were efficiently conducted without overfitting problems. The model proved its ability in dealing with the binary MFD detection task with high performance rates of precision, recall, and *F1*-score, by which the overall accuracy reached 98%. The multi-category OFD task aims at classifying images into one of three classes: mask, no mask, and occluded face with non-mask object. The model also proved its ability to deal with the OFD task with an overall accuracy of 86% achieved. The proposed model outperforms many recent state-of-the-art approaches. As a result, the proposed two-graph representations of key facial features proved its superiority in providing a generic

discriminating descriptor for the task of detecting occluded faces. In future, other graph-based features could be integrated into GCN architectures, such as using an interaction graph to indicate whether two nodes representing two facial parts interact with each other frequently. Integrating more graph nodes representing various informative relationships between the face parts and any mask or occlusion objects may offer improved performance. The proposed model can also be utilised as a monitoring system to track and warn people not wearing masks, especially in public areas. Moreover, this proposed model can serve as a masked or occluded face detector for face identification to recognize the identity of persons obscured with face coverings.

REFERENCES

[1] T. Kumar, "An evidence review of face masks against COVID-19," *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 9, no. 12, pp. 919–923, Dec. 2021.

[2] Centers for Disease Control and Prevention (CDC). *Types of Masks and Respirators*. Accessed: Jan. 22, 2021. [Online]. Available: https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/types-of-masks.html

[3] A. Alzu'bi, F. Albalas, T. AL-Hadhrami, L. B. Younis, and A. Bashayreh, "Masked face recognition using deep learning: A review," *Electronics*, vol. 10, no. 21, p. 2666, Oct. 2021, doi: 10.3390/electronics10212666.

[4] Z. Wang, G. Wang, B. Huang, Z. Xiong, Q. Hong, H. Wu, P. Yi, K. Jiang, N. Wang, Y. Pei, H. Chen, Y. Miao, Z. Huang, and J. Liang, "Masked face recognition dataset and application," Mar. 2020, *arXiv:2003.09093*.

[5] A. Anwar and A. Raychowdhury, "Masked face recognition for secure authentication," Aug. 2020, *arXiv:2008.11104*.

[6] X. Fan, M. Jiang, and H. Yan, "A deep learning based light-weight face mask detector with residual context attention and Gaussian heatmap to fight against COVID-19," *IEEE Access*, vol. 9, pp. 96964–96974, 2021.

[7] J. Ker, L. Wang, J. Rao, and T. Lim, "Deep learning applications in medical image analysis," *IEEE Access*, vol. 6, pp. 9375–9389, 2017.

[8] W. Wang, Q. Lai, H. Fu, J. Shen, H. Ling, and R. Yang, "Salient object detection in the deep learning era: An in-depth survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, early access, Jan. 12, 2021, doi: 10.1109/TPAMI.2021.3051099.

[9] A. Alzu'bi, A. Amira, N. Ramzan, and T. Jaber, "Improving content-based image retrieval with compact global and local multi-features," *Int. J. Multimedia Inf. Retr.*, vol. 5, no. 4, pp. 237–253, Nov. 2016.

[10] C. Li, R. Wang, J. Li, and L. Fei, "Face detection based on YOLOv3," in *Recent Trends in Intelligent Computing, Communication and Devices* (Advances in Intelligent Systems and Computing), vol. 1006, V. Jain, S. Patnaik, F. P. Vlădicescu, and I. Sethi, Eds. Singapore: Springer, 2020, pp. 277–284, doi: 10.1007/978-981-13-9406-5_34.

[11] S. Zhang, C. Chi, Z. Lei, and S. Z. Li, "RefineFace: Refinement neural network for high performance face detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 11, pp. 4008–4020, Nov. 2021.

[12] S. Ge, J. Li, Q. Ye, and Z. Luo, "Detecting masked faces in the wild with LLE-CNNs," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Honolulu, HI, USA, Jul. 2017, pp. 426–434.

[13] H. Jiang and E. Learned-Miller, "Face detection with the faster R-CNN," in *Proc. 12th IEEE Int. Conf. Autom. Face Gesture Recognit. (FG)*, Washington, DC, USA, May/Jun. 2017, pp. 650–657.

[14] M. Ngan, P. Grother, and K. Hanaoka, "Ongoing face recognition vendor test (FRVT) part 6B: Face recognition accuracy with facemasks using post-COVID-19 algorithms," NIST Interagency/Internal Rep., Nat. Inst. Standards Technol., Gaithersburg, MD, USA, Tech. Rep. 8331, Nov. 2020, pp. 1–83, doi: 10.6028/NIST.IR.8331.

[15] Y. Li, K. Guo, Y. Lu, and L. Liu, "Cropping and attention based approach for masked face recognition," *Int. J. Speech Technol.*, vol. 51, no. 5, pp. 3012–3025, May 2021.

[16] R. Zhu, K. Yin, H. Xiong, H. Tang, and G. Yin, "Masked face detection algorithm in the dense crowd based on federated learning," *Wireless Commun. Mobile Comput.*, vol. 2021, Oct. 2021, Art. no. 8586016, doi: 10.1155/2021/8586016.

[17] S. Prasad, Y. Li, D. Lin, and D. Sheng, "MaskedFaceNet: A progressive semi-supervised masked face detector," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Waikoloa, HI, USA, Jan. 2021, pp. 3389–3398.

[18] P. Nagrath, R. Jain, A. Madan, R. Arora, P. Kataria, and J. Hemanth, "SSDMNV2: A real time DNN-based face mask detection system using single shot multibox detector and MobileNetV2," *Sustain. Cities Soc.*, vol. 66, Mar. 2021, Art. no. 102692.

[19] S.-Q. Liu, X. Lan, and P. C. Yuen, "Multi-channel remote photoplethysmography correspondence feature for 3D mask face presentation attack detection," *IEEE Trans. Inf. Forensics Security*, vol. 16, pp. 2683–2696, 2021.

[20] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 770–778.

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

[22] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017.

[23] D. Xu, Y. Zhu, C. B. Choy, and L. Fei-Fei, "Scene graph generation by iterative message passing," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Honolulu, HI, USA, Jul. 2017, pp. 3097–3106.

[24] J. Yang, J. Lu, S. Lee, D. Batra, and D. Parikh, "Graph R-CNN for scene graph generation," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. (ECCV)*, Munich, Jermany, Sep. 2018, pp. 690–706.

[25] S. Zhang, H. Tong, J. Xu, and R. Maciejewski, "Graph convolutional networks: A comprehensive review," *Comput. Social Netw.*, vol. 6, no. 1, pp. 1–23, Dec. 2019.

[26] A. Jain, A. R. Zamir, S. Savarese, and A. Saxena, "Structural-RNN: Deep learning on spatio-temporal graphs," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 5308–5317.

[27] S. Yan, Y. Xiong, and D. Lin, "Spatial temporal graph convolutional networks for skeleton based action recognition," in *Proc. 32nd AAAI Conf. Artif. Intell.*, New Orleans, LA, USA, Feb. 2018, pp. 7444–7452.

[28] M. Narasimhan, S. Lazebnik, and A. G. Schwing, "Out of the box: Reasoning with graph convolution nets for factual visual question answering," in *Proc. 32nd Adv. Neural Inf. Process. Syst.*, Montreal, QC, Canada, Dec. 2018, pp. 2654–2665.

[29] F. Wu, T. Zhang, A. Souza, C. Fifty, T. Yu, and K. Q. Weinberger, "Simplifying graph convolutional networks," in *Proc. 36th Int. Conf. Mach. Learn.*, Long Beach, CA, USA, Jun. 2019, pp. 6861–6871.

[30] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, "Joint face detection and alignment using multitask cascaded convolutional networks," *IEEE Signal Process. Lett.*, vol. 23, no. 10, pp. 1499–1503, Oct. 2016.

[31] H. Li, Z. Lin, X. Shen, J. Brandt, and G. Hua, "A convolutional neural network cascade for face detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Boston, MA, USA, Jun. 2015, pp. 5325–5334.

[32] M. Mathias, R. Benenson, M. Pedersoli, and L. Gool, "Face detection without bells and whistles," in *Proc. 13th IEEE Comput. Soc. Conf. Comput. Vis. (ECCV)*, Zurich, Switzerland. Cham, Switzerland: Springer, Sep. 2014, pp. 720–735.

[33] P. Viola and M. J. Jones, "Robust real-time face detection," *Int. J. Comput. Vis.*, vol. 57, no. 2, pp. 137–154, 2004.

[34] G. Guo, H. Wang, Y. Yan, J. Zheng, and B. Li, "A fast face detection method via convolutional neural network," *Neurocomputing*, vol. 395, pp. 128–137, Jun. 2020.

[35] H. Filali, J. Riffi, A. M. Mahraz, and H. Tairi, "Multiple face detection based on machine learning," in *Proc. Int. Conf. Intell. Syst. Comput. Vis. (ISCV)*, Fez, Morocco, Apr. 2018, pp. 1–8.

[36] J. Mehta, E. Ramnani, and S. Singh, "Face detection and tagging using deep learning," in *Proc. Int. Conf. Comput., Commun., Signal Process. (ICCCSP)*, Chennai, India, Feb. 2018, pp. 1–6.

[37] A. Priadana and M. Habibi, "Face detection using Haar cascades to filter selfie face image on Instagram," in *Proc. Int. Conf. Artif. Intell. Inf. Technol. (ICAIIT)*, Yogyakarta, Indonesia, Mar. 2019, pp. 6–9.

[38] D. Garg, P. Goel, S. Pandya, A. Ganatra, and K. Kotecha, "A deep learning approach for face detection using Yolo," in *Proc. IEEE Punecon*, Pune, India, Nov./Dec. 2018, pp. 1–4, doi: 10.1109/PUNECON.2018.8745376.

[39] R. J. Cárdenas, C. A. Beltrán, and J. C. Gutiérrez, "Small face detection using deep learning on surveillance videos," *Int. J. Mach. Learn. Comput.*, vol. 9, no. 2, pp. 189–194, Apr. 2019.

[40] X. Fan and M. Jiang, "RetinaFaceMask: A single stage face mask detector for assisting control of the COVID-19 pandemic," May 2020, *arXiv:2005.03950*.

F. Albalas *et al.*: Learning Discriminant Spatial Features With Deep Graph-Based Convolutions for Occluded Face Detection

IEEE *Access*

[41] H. N. Vu, M. H. Nguyen, and C. Pham, "Masked face recognition with convolutional neural networks and local binary patterns," *Appl. Intell.*, vol. 52, pp. 5497–5512, Mar. 2022.

[42] S. V. Militante and N. V. Dionisio, "Real-time facemask recognition with alarm system using deep learning," in *Proc. 11th IEEE Control Syst. Graduate Res. Colloq. (ICSGRC)*, Shah Alam, Malaysia, Aug. 2020, pp. 106–110.

[43] N. U. Din, K. Javed, S. Bae, and J. Yi, "A novel GAN-based network for unmasking of masked face," *IEEE Access*, vol. 8, pp. 44276–44287, 2020.

[44] X. Luo, X. He, L. Qing, X. Chen, L. Liu, and Y. Xu, "EyesGAN: Synthesize human face from human eyes," *Neurocomputing*, vol. 404, pp. 213–226, Sep. 2020.

[45] B. Huang, Z. Wang, G. Wang, K. Jiang, K. Zeng, Z. Han, X. Tian, and Y. Yang, "When face recognition meets occlusion: A new benchmark," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Toronto, ON, Canada, Jun. 2021, pp. 4240–4244.

[46] S. E. Snyder and G. Husari, "Thor: A deep learning approach for face mask detection to prevent the COVID-19 pandemic," in *Proc. SoutheastCon*, Mar. 2021, pp. 1–8.

[47] A. S. Joshi, S. S. Joshi, G. Kanahasabai, R. Kapil, and S. Gupta, "Deep learning framework to detect face masks from video footage," in *Proc. 12th Int. Conf. Comput. Intell. Commun. Netw. (CICN)*, Bhimtal, India, Sep. 2020, pp. 435–440.

[48] B. Mandal, A. Okeukwu, and Y. Theis, "Masked face recognition using ResNet-50," Apr. 2021, *arXiv:2104.08997*.

[49] S. Sethi, M. Kathuria, and T. Kaushik, "Face mask detection using deep learning: An approach to reduce risk of Coronavirus spread," *J. Biomed. Inform.*, vol. 120, p. 103848, 2021.

[50] G. T. S. Draughon, P. Sun, and J. P. Lynch, "Implementation of a computer vision framework for tracking and visualizing face mask usage in urban environments," in *Proc. IEEE Int. Smart Cities Conf. (ISC)*, Sep./Oct. 2020, pp. 1–8.

[51] 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, Art. no. 108288.

[52] A. Oumina, N. El Makhfi, and M. Hamdi, "Control the COVID-19 pandemic: Face mask detection using transfer learning," in *Proc. IEEE 2nd Int. Conf. Electron., Control, Optim. Comput. Sci. (ICECOCS)*, Kenitra, Morocco, Dec. 2020, pp. 1–5.

[53] M. Ren, Y. Wang, Z. Sun, and T. Tan, "Dynamic graph representation for occlusion handling in biometrics," in *Proc. 34th AAAI Conf. Artif. Intell.*, New York, NY, USA, Feb. 2020, vol. 34, no. 7, pp. 11940–11947.

[54] Q. Ye and R. Li, "Mask recognition method based on graph convolutional network," *J. Phys. Conf. Ser.*, vol. 1920, no. 1, May 2021, Art. no. 012117.

[55] A. Alguzo, A. Alzu'bi, and F. Albalas, "Masked face detection using multi-graph convolutional networks," in *Proc. 12th Int. Conf. Inf. Commun. Syst. (ICICS)*, Valencia, Spain, May 2021, pp. 385–391.

[56] *Human Faces Dataset.* Accessed: Dec. 4, 2021. [Online]. Available: https://www.kaggle.com/ashwingupta3012/human-faces

[57] D. Chai, L. Wang, and Q. Yang, "Bike flow prediction with multi-graph convolutional networks," in *Proc. 26th ACM SIGSPATIAL Int. Conf. Adv. Geograph. Inf. Syst.*, Washington, DC, USA, Nov. 2018, pp. 397–400.

[58] M. Lv, Z. Hong, L. Chen, T. Chen, T. Zhu, and S. Ji, "Temporal multi-graph convolutional network for traffic flow prediction," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 6, pp. 3337–3348, Jun. 2021.

[59] W. Wang, J. Chen, Y. Zhang, Z. Gong, N. Kumar, and W. Wei, "A multi-graph convolutional network framework for tourist flow prediction," *ACM Trans. Internet Technol.*, vol. 21, no. 4, pp. 1–13, Jul. 2021.

[60] R. Avanzato, F. Beritelli, M. Russo, S. Russo, and M. Vaccaro, "YOLOv3-based mask and face recognition algorithm for individual protection applications," in *Proc. CEUR Workshop*, 2020, pp. 41–45.

[61] T. Q. Vinh and N. T. N. Anh, "Real-time face mask detector using YOLOv3 algorithm and Haar cascade classifier," in *Proc. Int. Conf. Adv. Comput. Appl. (ACOMP)*, Quy Nhon, Viet Nam, Nov. 2020, pp. 146–149.

[62] J. Zhang, F. Han, Y. Chun, and W. Chen, "A novel detection framework about conditions of wearing face mask for helping control the spread of COVID-19," *IEEE Access*, vol. 9, pp. 42975–42984, 2021.

[63] E. Ryumina, D. Ryumin, D. Ivanko, and A. Karpov, "A novel method for protective face mask detection using convolutional neural networks and image histograms," in *Proc. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. (ISPRS Arch.)*, Moscow, Russia, Apr. 2021, pp. 177–182.

[64] Z. Wang, P. Wang, P. C. Louis, L. E. Wheless, and Y. Huo, "WearMask: Fast in-browser face mask detection with serverless edge computing for COVID-19," Jan. 2021, *arXiv:2101.00784*.

FIRAS ALBALAS received the master's degree from The University of Jordan, in 2000, and the Ph.D. degree in computer science from Glamorgan (South Wales) University, Cardiff, U.K., in 2009. He is currently an Associate Professor with the Department of Computer Science, Jordan University of Science and Technology. His major research interests include mobile computing, the Internet of Things (IoT), cloud computing, and deep learning.

AHMAD ALZU'BI (Member, IEEE) received the M.Sc. degree in computer science from the Jordan University of Science and Technology, Jordan, in 2009, and the Ph.D. degree in computer science from the University of the West of Scotland (UWS), U.K., in 2016. He is currently an Assistant Professor with the Department of Computer Science, Jordan University of Science and Technology. He has many publications in reputed journals and conferences. His research interests include deep learning, image retrieval, medical imaging, and computer vision.

ALANOUD ALGUZO received the bachelor's degree in computer science from Al-Balqa Applied University, Jordan, in 2018, and the master's degree in computer science from the Jordan University of Science and Technology, Jordan, in 2021.

Her major research interests include data meaning, machine learning, image recognition, graph theory, and regression analysis.

TAWFIK AL-HADHRAMI received the M.Sc. degree in IT/applied system engineering from Heriot-Watt University, Edinburgh, U.K., and the Ph.D. degree in mesh communication and IoT from the University of the West of Scotland, Glasgow, U.K., in 2015. He is currently working as a Senior Lecturer at Nottingham Trent University (NTU), Nottingham, U.K. He is also a member of the Network Infrastructure and Cyber Security Group (NICS), NTU. He was involved in research at the Networking Group, University of the West of Scotland. He is involved in different projects with industries. He has established collaboration with the different international institutions over the world. His research interests include the Internet of Things (IoT) and applications, data science, network infrastructures and emerging technologies, artificial intelligence, computational intelligence, and 5/6G wireless communications. He is an Associate Editor of IEEE Access, *PeerJ computer Science*, and *International Journal of Cyber Forensics and Advanced Threat Investigations* journals.

ACHRAF OTHMAN (Senior Member, IEEE) is currently working as the Head of the ICT Accessibility Innovation and Research Section, Mada Center, Doha, Qatar. He is also working on research projects to enable persons with disabilities using innovative technologies and artificial intelligence (AI) solutions.

• • •