International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. 10, No.1 (Apr-2021)

http://dx.doi.org/10.12785/ijcds/100144

Feature Extraction and Fusion for Face Recognition Systems using Pre-Trained Convolutional Neural Networks

Soad Almabdy[1] and Lamiaa Elrefaei[1,2]

[1] *Computer Science Department, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia*
[2] *Electrical Engineering Department, Faculty of Engineering at Shoubra, Benha University, Cairo, Egypt*

Received 05 May. 2020, Revised 04 Jun. 2020, Accepted 10 Jun. 2020, Published 01 Apr. 2021

Abstract: Recently, face recognition applications achieved promising results by using Convolutional Neural Network (CNN). CNN has the capability to extract features automatically from images and does not need to extract hand-crafted features as traditional algorithms. Feature fusion aims to provide improvements of data validity for both traditional algorithms and deep learning algorithms. In this paper we propose a feature fusion approach for face recognition, the approach performs fusion at the feature level by applying two pre-trained CNNs AlexNet and ResNet-50. Firstly, extracting the feature from both pre-trained CNN AlexNet and ResNet-50 separately. Secondly, fuse the feature maps learned from AlexNet and ResNet-50. Finally, a Support Vector Machine (SVM) classier is used for the classification task. Experiments are conducted on the following datasets: FEI face, GTAV face, ORL, F_LFW, Georgia Tec Face, LFW, DB_Collection, demonstrate the effectiveness of the proposed approach. In addition, the fusion of the two CNN based models AlexNet and ResNet-50 lead to significant performance improvement. In particular, the fusion approach achieves accuracy in range (96.21%-100%) on all datasets.

Keywords: Convolutional Neural Network (CNN), Face Recognition, Feature extraction, Feature Level Fusion, Deep learning.

1. INTRODUCTION

Face recognition is an active area of research of computer vision. It has been used in different applications such as: security application, law enforcements, commercial use. The face recognition applications contain four stages including face detection, face alignment, feature extraction, and classification. The Feature extraction stage is an important step for recognition system.

In constrained environments, hand designed features that extracted with traditional algorithms such as: Local Phase Quantization (LPQ) [1], and Local Binary Patterns (LBP) [2], have achieved respectable performance for face recognition. Nevertheless, in unconstrained environments the performance degrades dramatically because, face images cover complex and several challenges such as pose, expression, occlusion, and illumination. CNNs is the typical feature extraction approach in unconstrained environments that achieved surpassing face recognition accuracy [3].

Convolutional Neural Networks CNN's [4] has been successfully applied for research problems in computer vision area, such as recognizing and classifying given objects and images. It is a specialized form of neural network which comprises of a grid topology. CNN networks mainly contain of filters, which is applied at different locations of an organized input data in order to produce an output map. Convolution networks have played a vital role in deep learning evolution and are an example of how information and insights from studying the brain can be applied to machine learning applications.

The fusion of the data at the feature level consists of the combination of feature map sets of the feature vector. The sets of the feature vector had a rich information about the raw biometric data. Combination at this level is achieved better recognition performance [5], [6]. However, fusion at this level is an important process, which provides the most discriminatory information from original various feature sets. Feature fusion techniques has two categories are Parallel feature fusion and Serial feature fusion [7] .

E-mail: salmabdy@kau.edu.sa, laelrefaei@kau.edu.sa, lamia.alrefaai@feng.bu.edu.eg

The aim of this paper to improve the face classification accuracy by using different CNN architectures. The main contributions are evaluating the performance when apply feature fusion for both AlexNet model and ResNet-50 model, and for classification task using Support Vector machine SVM [8]. And evaluate the proposed face recognition approach using different datasets include FEI face dataset [9], GTAV face dataset [10], ORL dataset [11], Georgia Tec Face [12], Frontalized Labeled Faces in The Wild (F_LFW) [13], Labeled Faces in The Wild (LFW) [14], and DB_Collection.

The rest of the paper is organized as follows: Section 2 reviews the related works. In Section 3, present the methodology with a detail of the proposed approach. Section 4 analyzed the experimental results. Finally, Conclude the paper in section 5.

2. Related works

This section reviews some feature fusion methods that use Convolutional Neural Networks CNNs and traditional feature fusion and applied for face detection and face recognition systems. N. Bodla et al. [15] proposed a feature fusion approach named deep heterogeneous for face recognition using deep CNNs (DCNNs). The approach efficiently fused the feature from different deep networks. They evaluated the approach on the IARPA Janus Challenge Set 3 (Janus CS3). The results of the experiment showed the feature fusion network method can efficiently enhance the recognition accuracy. In addition, the results demonstrate efficiency of the method and the improvements over the state-of-the-art. M. Simón et al. [16], proposed a face recognition method based on deep CNN model. They proposed a multimodal facial recognition system. Also, applied a fusion of CNN features with various hand- craft features. The system evaluated on RGB-D-T database. Results of the experiments demonstrate that the combination between the classical feature and CNN feature enhances the performance of the system. W. Zhang et al. [17], they applied feature fusion technique for both face verification and recognition. they considered the confidence level and face representation level. For both level the investigation as following: first, in the representation level, they considered two algorithms for feature fusion PCA and LDA. Second, in the confidence level they investigated two rules sum and product. The experiment conducted on FERET dataset and their own face database. The experiments results showed that the performance of both face verification and recognition can be improving when using appropriate information fusion. And the fusion results of representation level are more robust than the confidence level. F. Simanjuntak and G. Azzopardi [18], proposed a feature fusion method based on Combination of Shifted Filter Responses (COSFIRE) [19], and the pre-trained CNN VGG [20]. the method used face images for gender recognition. The proposed method evaluated on GENDER-FERET dataset. The experiment results achieved 97.45% of recognition rate. Also, the experiments results demonstrate that COSFIRE method provides corresponding features to CNNs. J. Feng et al. [21], proposed fusion scheme for 3D face recognition using CNNs model. The method fused the 2D feature and 3D feature, the take the output of the feature fusion as input for softmax classifier. The proposed method evaluated on CASIA-3D. The results of the experiments showed that the face recognition method improves the recognition performance. Z. Lu et al. [22], proposed face recognition technique called covariance matrix regularization (CMR). The fusion method reduced the overfitting at the fusion process. In the method they are using the training data matrix for assigning the weights to cross-feature covariance. And applied four schemes for feature fusion: first, feature fusion of three-color channel values. Second, feature fusion of Local binary pattern (LBP) features. Third, feature fusion of LBP features. Finally, feature fusion of two CNN models features. The proposed technique is conducted on Georgia Tech dataset, MultiPIE dataset, LFW datasets, and AR dataset. Results showed that CMR outperforms the technique of single feature. W. Tian et al. [23], proposed detection with Feature Fusion and Segmentation Supervision technique (DF2S2) for face detection. In the method the structure of Feature fusion pyramids applies semantic information from feature maps of higher-level to augment feature maps of low-level. They used pre-trained on ResNet-50 models. The experiments of proposed technique are conducted on WIDER FACE benchmark. And the and technique achieved state-of-art results. Y. Ye et al. [24], proposed facial expression recognition method named Region-based Convolutional Fusion Network (CMR). The models built based on VGG model [44]. CMR consist of six convolutional layers and three pooling layers. CMR networks used to extract features then fuse them for facial expression recognition. The experiment conducted on KDEF and CK+, and Oulu-CASIA datasets. Experimental results showed that CMR achieved comparable performance facial expression recognition. Y. Rehman et al. [25], evaluated the feature fusion method for face liveness detection using CNN models. The method adaptively balances the feature fusion for face images in real-world and images generated by CNN. the experiment conducted for face liveness detection on databases including: Replay-Attack face anti-spoofing, OULU, and CASIA. the proposed method is effective for face detection. A. A. Moustafa et al. [26], proposed a system for age-invariant face recognition based on VGG model. The system consists of four main stages include: pre-processing, feature extraction, feature fusion, and classification. In the model feature extraction stage is done using VGG model. Then, extracted features are fused using the real-time feature-level multi-discriminant correlation analysis in order to reduces feature dimensions and improve the accuracy for the system. The experiments are accomplished on FGNET and MORPH datasets. The result of the system showed that the age-invariant face recognition system outperforms the state-of-the-art system on same data.

3. METHODOLOGY

In this study we are investigate the face recognition performance through Pre-trained CNNs models. We are applying feature fusion from the pre-trained CNN ResNet-50 and AlexNet models. We extract suitable image features using two models AlexNet and ResNet-50. Then utilize the fused features in the classification phase. Finally, for classification task using Support Vector Machine SVM. The approach overview shown in “Fig. 1”.

The study follows the following stages:

- Pre-processing stage: using face detection algorithm then cropping face image, resizing all images to appropriate size of pre-trained CNN models, and converting all input images to RGB images.
- Features Extraction stage: using two pre-trained CNNs AlexNet and ResNet-50 to extract suitable image features.
- Features Fusion stage: applying fusion at feature level by fuse feature maps learned from AlexNet and ResNet-50.
- Classification stage: for the classification task using SVM.
- Finally, the experiments are conducted on several datasets. Looking at different results and analyzing the performance for the single net and fusion approach.

A. *pre-processing*

In the implementation we start with pre-processing step by detected face in images using Viola-Jones algorithm then cropping the face images region and convert any grayscale images to RGB. For AlexNet model resize images to (227-by-227). ResNet-50 network resize to 224-by-224.

B. *Feature extraction*

The features learned from each network AlexNet and ResNet-50 are extracted. Feature level fusion is done by applying concatenate technique for the feature vector from each network obtained to form a powerful face representation. In the study as process for feature level fusion first, extract feature vector-1 from AlexNet model as A, from layer ‘fc7’ that before classification layer. The output of feature extraction is (4,096-dimensional) feature vector. Second, ResNet-50 extract feature vector-2 as B, from layer ‘fc1000’ with (2,048 –dimensional) feature vector. The goal is to combine these feature vectors in order to produce a new feature sets by using serial feature fusion. The serial combined feature C is (A+B)-dimensional. The new feature vector as result from serial combination is (6144-dimension).

1) AlexNet

Krizhevsky et al. [27] introduced AlexNet network as shown in “Fig. 2”. It was winning the ImageNet challenge in 2012 with a top5 error of 16.4%. AlexNet consists of 5 Convolutional layers (Conv1-5), 3 Max-Pool layers, and 3 fully-connected layers (FC). It accepts images as input with size [227 x 227 x 3] image as an input. The first convolution layer has 96 kernels, each with size of [11 x 11 x 3] each, with stride of 4 and pad of 0. This is followed by max a pool layer with filter size of [3 x 3], stride of 2 and a normalization layer. A second convolutional layer has 256 filters with size of [5 x 5 x 96], stride of 1, pad of 2, followed by a max pool layer with filter size of [3 x 3], stride of 2 and a normalization layer. There are then three convolutional layers in sequence with 384|3 x 3 x 256 - Stride 1 - Pad 1, 384|3 x 3 x 384 - Stride 1 - Pad 1, and Conv5, 256|3 x 3 x 384 - Stride 1 - Pad 1 respectively. After these three convolutions there is a max pool layer followed by three fully connected layers of size [4096 x 4096 x 1000] with ReLU activations and dropout probability of 0.5. Finally, the output layer cost function used is softmax. The learning rate is manually decayed by factor of 10 when validation accuracy is plateaued. An ensemble of seven CNN’s was used in the 2012 ImageNet competition.

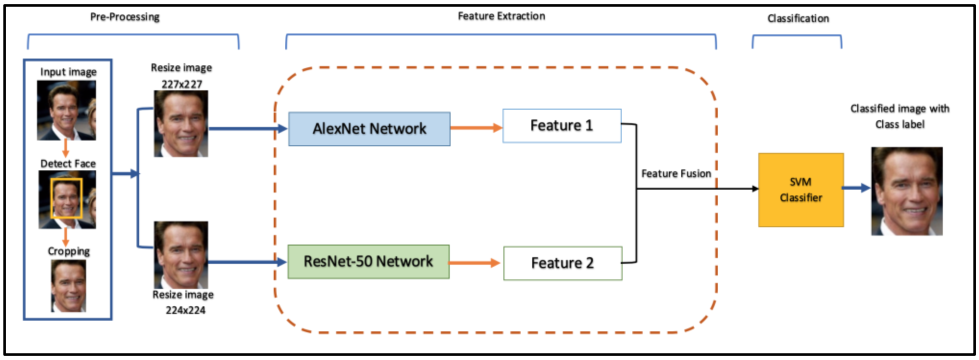

Figure 1. Feature fusion from AlexNet Net and ResNet-50 Net

2) ResNet-50

ResNet Network proposed by Kaiming He et al. [28]. It is wins image classification challenge in 2015 with a top5 error of 3.57%. ResNet-50 architecture shown in "Fig. 3". It is built by combing 5 convolutional layers. Architecture of ResNet-50 uses a [224 x 224 x 3] image as an input. ResNet-50 is composed of 5 convolution layers. Each layer includes Residual Blocks that contain three convolution layers [1x1], [3x3], [1x1]. The first convolution layer has 64 kernels, each with size of [7 x 7 x 3], with stride of 2 and pad of 3. This is followed by max-pool layer with filter size of [3 x 3], stride of 2. A second convolutional layer has three Residual Blocks as stacked layer: the first is a [1x1] convolution layer, then [3x3] convolution layer, and finally a 1x1 convolution layer with depth of 256. The second convolutional layer contains four Residual Blocks with 512 dimensions for depth. The third layer contains 6 Residual Blocks with depth of 1024. and finally, 3 Residual Blocks with depth dimension 2048. After these five convolutions there is a max pool layer followed by 1 fully connected layer with size [1000]. Finally, the output layer cost function used is softmax. ResNet-50 is currently state-of-the-art and often the default choice when it comes to using convolutional neural networks in practice. In ResNet-50, layer 'fc1000' is used to extract features and then used these features as input for SVM classifiers.

C. Classification

In this process the proposed method used the SVM classifier. SVM [8] is supervised binary classification method. It used hyperplane between two input classes that maximizes the margin for them. Suppose we have the input data X and the output data Y. The function y = f (x, α) find the hyperplane in two-dimensional space, where α are the parameters of the function.

4. Experimental Results and Analysis

The experiment in the study evaluated the performance when applying feature fusion between feature vectors from AlexNet model and ResNet-50 model and using SVM classifier. In the experiment the data were split into 80% training and 20% testing randomly. "Fig. 1" presented the process for the fusion approach.

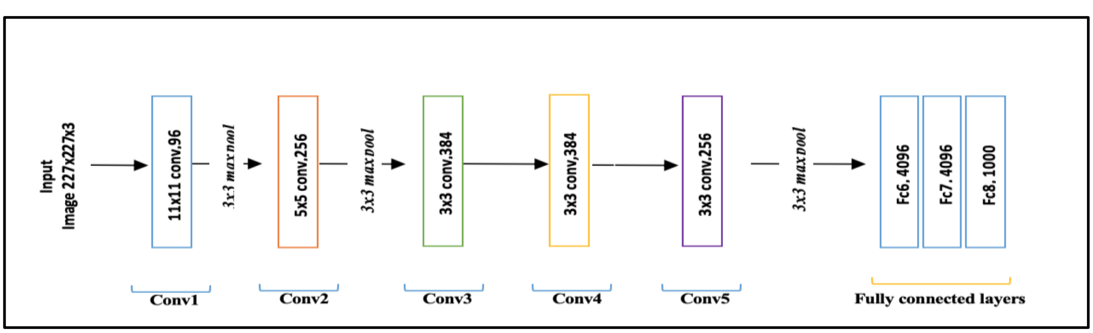

Figure 2. The Architecture of AlexNet

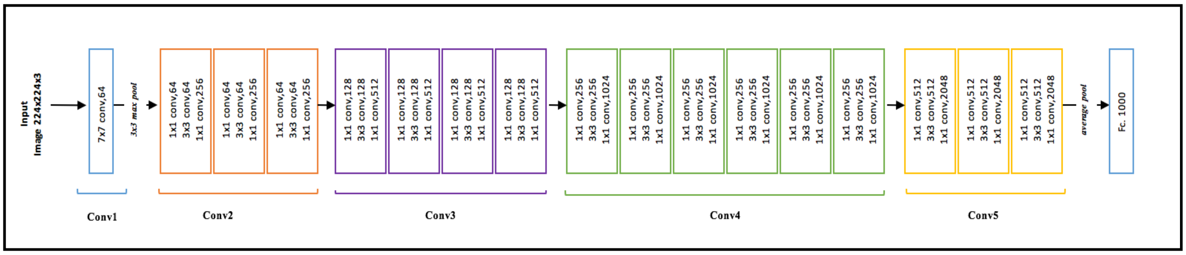

Figure 3. The Architecture of ResNet-50 [3]

A. *The Experimental Setup*

All experiments in this study were conducted on Windows platform. The platform properties are Intel® Core i7- CPU, 16 GB, and NVIDIA GEFORCE GTX 1050TI. The implementation accomplishes using MATLAB version 2018a. In the study, the performance evaluation based on recognition accuracy. The accuracy is defining as the percentage of correct label prediction.

B. *Dataset Description*

In this study all experiments are conducted on the following datasets: FEI face, GTAV face, Georgia Tec Face, ORL, F_LFW, LFW, and DB_Collection. TABLE I described all datasets and their specifications that used in the study. ORL dataset [11], which include 400 images for 40 individuals with different facial expressions. GTAV Face Database [10] include 704 images for 44 individuals, the dataset contains images with different pose angles and illuminations. In our study, we chosen 34 images per each person. Georgia Tech Face Database [12], which contains images for 50 individuals, the dataset includes images with different illumination appearance, and facial expression. FEI face [9] contains 2800 images for 200 individuals. In our study, we choose 700 images for 50 persons with frontal images only. LFW [14], the dataset includes more than 13,000 images for more than 1680 individual. LFW dataset proposed for unconstrained face recognition research. In our study we choose 700 images for 50 individuals. F_LFW [13], contains the frontalized images of LFW dataset. In the study selected 700 images for 50 persons. F_LFW created in research [29]. Finally, DB_Collection that generated in this study. the dataset includes combination of images from all datasets that used in the study. The total pictures in the dataset are 2880 images for 30 persons from each dataset.

TABLE I. DESCRIPTION OF FACE DATASETS

Datasets	Images	Identities	Size
ORL	400	40	92x112
GTAV Face	704	44	240x320
Georgia Tech Face	700	50	131 × 206
FEI face	700	50	640x480
LFW	700	50	250x250
F_LFW	700	50	272x323

C. *Result analysis*

The experiment is performed on the following datasets: FEI face, GTAV face, ORL, Georgia Tec Face, F_LFW, LFW, and DB_Collection. The results of this method when applying fusion between two models AlexNet and ResNet-50 illustrated in "Fig. 4". As comparison of the results when using single net or fusion between AlexNet and ResNet-50 presented in TABLE II. In case of dataset simple and not include a lot of limitation such as orientation, illumination, complex background, etc., the single net performs better than fusion as result with GTAV face, ORL datasets. The fusion network achieved 99%, 98.31% ,96.21%, and 99.07%, and 99.89% on FEI, Georgia Tec Face, LFW, F_LFW, and DB_Collection respectively. The result better than a single net because the fusion performs good result with a complex dataset.

D. *Comparison with the state-of-the-art*

This section provides a comparison for the proposed method with the state-of- the-art in terms of feature fusion level method with CNN models. TABLE III presents a comparison for face recognition accuracy with the state-of-the-art methods include: DCNNs [15], VGG-Face CNN [18], 3D Face Recognition Method [21], CMR [22], DF2S2 [23], and MDCA [26]. As presented in TABLE III, all methods applied feature fusion at feature level using different CNN models and datasets with different number of samples, also different experimental setting. Our method achieved high accuracy as comparable with the methods of state-of- the-art in term of fused feature from two different models. The accuracy results were in range of (98%-100) with most of datasets. The results were higher than DF2S2 [23] and MDCA [26], which fused feature from one model.

5. CONCLUSION

In this paper, convolutional neural networks AlexNet, RasNet-50 and fusion between AlexNet and ResNet-50 are investigated on different datasets. In the study we conducted an experiment to evaluate the performance for a fusion at feature level from Convolutional Neural Networks ResNet-50 and AlexNet. First extracting learned features from both networks AlexNet and ResNet-50. Second, combine these two feature sets in order to produce a new feature vector by using serial feature fusion. Finally, using SVM classifier. The investigation study is conducted FEI face, GTAV face, LFW, ORL, F_LFW, Georgia Tec Face, and DB_Collection. The results indicated that the single net performs very well with a simple dataset, but the fusion net performs good result with a complex dataset. In the future, we propose to additional improvement for recognition accuracy. To do so, different data fusion will be performing such as decision level fusion and score level fusion.

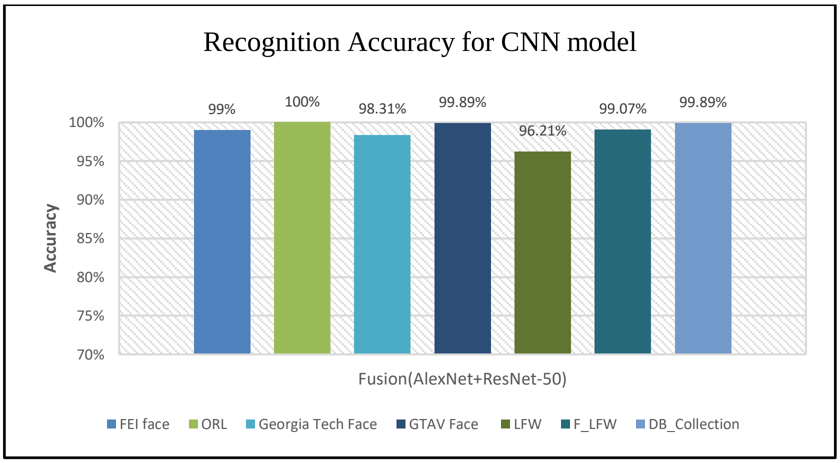

Figure 4. Recognition Accuracy for Fusion (AlexNet and ResNet-50)

TABLE II. COMPARISON BETWEEN THE SINGLE NET AND FEATURE FUSION OF (ALEXNET AND RESNET-50) NETWORKS

	FEI	ORL	Georgia Tech	GTAV	LFW	F_LFW	DB_Collection
AlexNet	97.50%	99.17%	96%	99.55%	94%	98%	97%
ResNet-50	98.50%	100%	96%	100.0%	94%	96%	97.50%
Fusion (AlexNet+ResNet-50)	**99%**	**100%**	**98.31%**	**99.89%**	**96.21%**	**99.07%**	**99.89%**

TABLE III. COMPARISON FOR FACE RECOGNITION ACCURACY WITH THE STATE-OF-THE-ARTS METHODS

References	Model	Feature extraction method	Classifier	Dataset	Accuracy
N. Bodla et al.[15]	DCNNs	Two CNN	The joint Bayesian metric	Janus CS3	98.42%
F. Simanjuntak and G. Azzopardi [18]	VGGFace CNN	VGG + COSFIRE filter	SVM	GENDER-FERET	97.45%
J. Feng et al. [21]	3D Face Recognition Method	Two CNN	Softmax	CASIA-3D FaceV1	98.44%
Z. Lu et al. [22]	covariance matrix regularization (CMR) technique	VGG-Face + ResNetShort	Softmax	Multi PIE, Georgia Tech, AR and LFW	98.92%
W. Tian et al .[23]	DF^2S^2 (Detection with Feature Fusion and Segmentation Supervision)	ResNet-50	Loss function	WIDER FACE	95.6%
A. A. Moustafa et al. [26]	Multi-discriminant correlation analysis (MDCA)	VGG	SVM	FGNET and MORPH	80.4% (FGNET) 96.5% (MORPH)
Our Proposed (2020)	Fusion (AlexNet+ResNet-50)	AlexNet+ResNet-50	SVM	FEI	99%
				GTAV	99.89%
				ORL	100%
				Georgia Tec	98.31%
				LFW	96.21%
				F_LFW	99.07%
				DB_Collection	99.89%

6. REFERENCES

[1] C. Turan and K. M. Lam, "Histogram-based local descriptors for facial expression recognition (FER): A comprehensive study," J. Vis. Commun. Image Represent., vol. 55, no. May, pp. 331–341, 2018.

[2] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," IEEE Trans. Pattern Anal. Mach. Intell., vol. 24, no. 7, pp. 971–987, 2002.

[3] S. Almabdy and L. Elrefaei, "Deep Convolutional Neural Network-Based Approaches for Face Recognition," Appl. Sci., vol. 9, no. 20, p. 4397, 2019.

[4] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf, "Closing the gap to human-level performance in face verification. deepface," IEEE Comput. Vis. Pattern Recognit., 2014.

[5] A. Ross and R. Govindarajan, "Feature level fusion in biometric systems," Proc. Biometric Consort. Conf., vol. 2, pp. 1–2, 2004.

[6] V. A. Sindagi and V. M. Patel, "DAFE-FD: Density Aware Feature Enrichment for Face Detection," pp. 2185–2195, 2019.

[7] L. Yang, X. Ban, Y. Li, and G. Yang, "Multiple features fusion for facial expression recognition based on ELM," Int. J. Embed. Syst., vol. 10, no. 3, pp. 181–187, 2018.

[8] Boser, B. E., I. M. Guyon, and V. N. Vapnik., "A training algorithm for optimal margin classifiers.," Proc. Fifth Annu. Work. Comput. Learn. Theory, pp. 144–152, 2003.

[9] C. E. Thomaz, "FEI face database," 2012. [Online]. Available: https://fei.edu.br/~cet/facedatabase.html. [Accessed: 06-Apr-2019].

[10] F. Tarres and A. Rama, "GTAV face database," 2011. [Online]. Available: https://gtav.upc.edu/en/research-areas/face-database. [Accessed: 06-Apr-2019].

[11] "ORL face database." [Online]. Available: http://www.uk.research.att.com/facedatabase. html. [Accessed: 06-Apr-2019].

[12] A. V. Nefian, "Georgia Tech Face Database." [Online]. Available: http://www.anefian.com/research/face_reco.htm. [Accessed: 06-Apr-2019].

[13] "Frontalized faces in the wild," 2016. [Online]. Available: https://www.micc.unifi.it/resources/datasets/frontalized-faces-in-the-wild/. [Accessed: 06-Apr-2019].

[14] G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller, "Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments," 2007.

[15] N. Bodla, J. Zheng, H. Xu, J. C. Chen, C. Castillo, and R. Chellappa, "Deep heterogeneous feature fusion for template-based face recognition," Proc. - 2017 IEEE Winter Conf. Appl. Comput. Vision, WACV 2017, pp. 586–595, 2017.

[16] M. Simón et al., "Improved RGB-D-T based face recognition," IET Biometrics, pp. 297–304, 2016.

[17] W. Zhang, S. Shan, W. Gao, Y. Chang, B. Cao, and P. Yang, "Information fusion in face identification," Proc. - Int. Conf. Pattern Recognit., vol. 3, pp. 950–953, 2004.

[18] F. Simanjuntak and G. Azzopardi, Fusion of CNN- and COSFIRE-Based Features with Application to Gender Recognition from Face Images, vol. 943. Springer International Publishing, 2020.

[19] G. Azzopardi, A. Greco, and M. Vento, "Gender recognition from face images with trainable COSFIRE filters," 2016 13th IEEE Int. Conf. Adv. Video Signal Based Surveillance, AVSS 2016, no. August, pp. 235–241, 2016.

[20] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," pp. 1–14, 2014.

[21] J. Feng, Q. Guo, Y. Guan, M. Wu, X. Zhang, and C. Ti, 3D face recognition method based on deep convolutional neural network, vol. 670. Springer Singapore, 2019.

[22] Z. Lu, X. Jiang, and A. Kot, "Feature fusion with covariance matrix regularization in face recognition," Signal Processing, vol. 144, pp. 296–305, 2018.

[23] W. Tian et al., "Learning Better Features for Face Detection with Feature Fusion and Segmentation Supervision," 2018.

[24] Y. Ye, X. Zhang, Y. Lin, and H. Wang, "Facial expression recognition via region-based convolutional fusion network," J. Vis. Commun. Image Represent., vol. 62, no. April, pp. 1–11, 2019.

[25] Y. Rehman, L. M. Po, M. Liu, Z. Zou, W. Ou, and Y. Zhao, "Face liveness detection using convolutional-features fusion of real and deep network generated face images," J. Vis. Commun. Image Represent., vol. 59, pp. 574–582, 2019.

[26] A. A. Moustafa, A. Elnakib, and N. F. F. Areed, "Age-invariant face recognition based on deep features analysis," Signal, Image Video Process., 2020.

[27] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," Adv. Neural Inf. Process. Syst., pp. 1–9, 2012.

[28] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," 2016 IEEE Conf. Comput. Vis. Pattern Recognit., pp. 770–778, 2016.

[29] C. Ferrari, G. Lisanti, S. Berretti, and A. Del Bimbo, "Effective 3D based frontalization for unconstrained face recognition," 2016 23rd Int. Conf. Pattern Recognit., pp. 1047–1052, 2016.

Soad Almabdy is currently a Teacher Assistant in the Department of Computer Science at FCIT, King Abdulaziz University, and Jeddah. She obtained his master's degree from King Abdulaziz University, Saudi Arabia. Her research interests are in the fields of Computer Vision and Machine learning.

LAMIAA A. ELREFAEI, received the B.Sc. degree (Hons.) in electrical engineering (electronics and telecommunications), and the M.Sc. and Ph.D. degrees in electrical engineering (electronics) from the Faculty of Engineering at Shoubra, Benha University, Egypt, in 1997, 2003, and 2008, respectively. She held a number of faculty positions at Benha University, as a Teaching Assistant, from 1998 to 2003, as an Assistant Lecturer, from 2003 to 2008, and has been a Lecturer, since 2008. She is currently an Associate Professor with the Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia. Her research interests include computational intelligence, biometrics, multimedia security, wireless networks, and nano networks.