

Person identification based on facial biometrics in different lighting conditions

Marem H. Abdulabas, Noor D. Al-Shakarchy

Department of Computer Science, Computer Science and Information Technology, University of Kerbala, Kerbala, Iraq

Article Info

Article history:

Received May 28, 2022

Revised Sep 17, 2022

Accepted Oct 13, 2022

Keywords:

Age gap

Convolution neural networks

Facial recognition

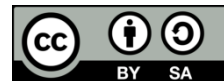
Haar cascade

MobileNet

ABSTRACT

Technological development is an inherent feature of this time, that reliance on electronic applications in all daily transactions (business management, banking, financial transfers, health, and other important aspects of life). Identifying and confirming identity is one of the complex challenges. Therefore, relying on biological properties gives reliable results. People can be identified in pictures, films, or real-time using facial recognition technology. A face individual is a unique identifying biological characteristic to authenticate them and prevents permits another person to assume that individual's identity without their knowledge or consent. This article proposes the identification model by facial individual characteristics, based on the deep neural network (DNN). The proposed method extracts the spatial information available in an image, analysis this information to extract the salient features, and makes the identifying decision based on these features. This model presents successful and promising results, the accuracy achieves by the proposed system reaches 99.5% (+/- 0.16%) and the values of the loss function reach 0.0308 over the Pins Face Recognition dataset to identify 105 subjects.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Marem H. Abdulabas

Department of computer science, Computer Science and Information Technology, University of Kerbala
Kerbala, Iraq

Email: marem.h@s.uokerbala.edu.iq

1. INTRODUCTION

The human face is one of the most distinctive features of the human body. A person's face can help them recognize others. Face recognition is included in biometric identification alongside voice recognition, retinal scanning, and fingerprint recognition [1]. Face recognition has emerged as one of the most intriguing study issues in computer vision. Various studies have been conducted for a computer to recognize a person's face [2]. Furthermore, requests for the implementation of the technology have sparked interest in facial recognition research.

Recently, with the development of technology, biological properties have become a necessity for person identification and verification systems. One important biometric is facial. Facial identification is the process of recognizing or validating an individual's identification by looking at their face [3]. The facial recognition software may be used to identify persons in photographs, videos, or in real-time.

System identification is the process of implying models from observations and research into the behaviors and attributes of systems. The task of developing mathematical models of dynamical systems based on observed data is known as system identification, in other words; these systems attempt to find a model relationship and find the system ordering and the unknown function's approximation [4]. Depending on the available information that characterizes the system's behavior, there are two techniques for system

identification. The state-space approach (internal description) is the first approach, which describes the system's internal state and is employed whenever the system dynamical equation is accessible [5]. The second method is the black-box approach (input-output description), which is utilized when there is no other knowledge about the system except its inputs and outputs. One of the active research subjects is automatic person identification, which aims to teach an independent system the ability to recognize a person using biometrics [6]. The unit may use the person's identity as input for a wide range of real-world applications by detecting their identity. Security identification and verification, for example, might be used to follow people's movements in the real world, similar to how automatic license plate readers track vehicles by plate numbers. In other nations, such as the United States, China, and Malaysia, real-time facial recognition is already in use, even during athletic events. Many elements influence the trustworthiness of any identity and authentication system, such as something you already know (such as a password), have something (such as a smart card), something you're doing (such as a fingerprint or other biometric method). When a user declares their identity (for example, with a username), identification occurs, and authentication occurs when users prove their identity. Users are authenticated, for example, when they supply both their username and the proper password. Users are then awarded permissions, rights, and privileges depending on their verified identities. Establishing a reliable face recognition necessitates the development of a set of dependable procedures that play a key role in successful person identification, such as effective feature extraction methods for capturing salient features in various lighting circumstances.

Luminance information is a problem for traditional methods of person identification based on face biometric [7]. When images were obtained with varied resolutions and elimination settings, these systems performed poorly. Most computer vision applications work with images in a variety of lighting situations and resolutions, and real-world applications must consider accuracy and execution speed. Deep neural networks (DNNs) introduced an efficient way for extracting robust features like facial features, which analyzes the facial mesh to make a judgment. In this work facial biometric framework for person identification in different lighting conditions has been proposed based on using convolutional neural networks (CNN) pre-trained model called MobileNet.

Several studies on the problem of face recognition have been conducted. Al-Shakarchy *et al.* [8] based on the new DNN architecture, The proposed approach for categorization of the eyes status in this research, which includes those that are open or closed in images of eyes with various elimination conditions, is described in this paper DNNs. The deep neural network classification (DNNC) method was suggested, and it worked effectively with short training and modeling times. With a 96 percent training set, the proposed system can attain accuracy. The highest value of the loss function is 0.01. With adequate hardware resources, this system can be implemented.

Adnan *et al.* [9] the main idea behind this work is to feature extracted from the face that used techniques such as (histogram of gradients (HOG), local binary patterns (LBP), principal component analysis (PCA), speeder up robust feature (SURF), and Harris). Next, the k-nearest neighbors (KNN) algorithm is used to find the similarity ratio between the training image preserved within the system and the testing image. The data contains approximately 100 images (60 training images and 40 testing images) of different persons. The hog algorithm is the most efficient when compared to other approaches, with an 85 percent hit rate.

Wang *et al.* [10] For face recognition, the two-dimensional quaternion principal component analysis (F-2D-QPCA) approach was utilized face recognition (FR). To optimize picture variance, they used the F-norm and a greedy iterative strategy to ensure the method's convergence and robustness. When compared to other current models, the experimental findings of this model on a variety of color face picture databases confirmed its efficacy and accuracy. Wang *et al.* [11] For FR, a combination of LBP and PCA with beetle antennae search algorithm (BASA) was utilized. The SoftMax was used to reduce the time it took to complete the FR process using multi-face classification. Tang *et al.* [12] The ORL's FR rate was boosted to 100 percent, and Yale-score B's rose to 97.51 percent. The LBP operator and 10 CNNs with five different brain architectures were used to connect the layers. They were employed for greater feature extraction and tweaking of network parameters. to improve the accuracy and speed of 3-dimensional FR, Shi *et al.* [13] LBP was used to extract information from the 3-D facial depth picture. The information was then classified using a support vector machine (SVM) and light detection and classification LBP, SVM and LBP were offered as a combined approach for computer vision research. Rameswari *et al.* [14] a team of researchers has developed a computer system that uses facial recognition and radio frequency identification (RFID) technology to enhance system security. When compared to other face identification algorithms such as local binary patterns histogram (LBPH), FaceNet obtained a higher accuracy of 97 percent in this system. Omara *et al.* [15] Kernel SVM is employed as a classifier, and it outperformed standard classifiers. The model was evaluated on an AR dataset and obtained 99.85 percent accuracy. It surpassed numerous state-of-the-art multi-modal approaches to create multimodal biometric systems.

For high and low-resolution face photos, Zangeneh *et al.* [16] For each type of resolution to be translated into a common space, a connected mapping approach architecture with two branches of deep

convolutional neural networks was used. The branch for high-resolution face image transformation has a fourteen-layer network, whereas the branch for low-resolution face image transformation has an additional five-layer network coupled to the fourteen-layer network. The facial recognition technology (FERET), labeled faces in the wild (LFW), and multiple biometric grand challenge (MBGC) datasets were used to test it [17], and the architecture proposed outperformed the other methods, achieving a 5 percent higher accuracy of 97.2 percent when compared to the traditional methods used previously. When applied to very low-resolution photos of 6*6 pixels, it outperformed the other approaches. Zhang and Wang [18] The approach labels the face photos as fatigued if they have more than 80% of their eyes closed. The researchers used an algorithm based on the sequential forward floating selection (SFFS) algorithm. Gao *et al.* [19] The authors offer a CNN for extracting facial characteristics, and the face alignment technique is utilized to further localize the critical spots on the face. A joint Bayesian framework (JBF) was designed to score feature vector similarity, and PCA was used to reduce the dimensionality of the deep features. On the CAS-PEAL dataset, an accuracy of 98.52 percent was attained [20]. To solve the problem of age estimation in real-time applications, Khan *et al.* [21] The mean absolute error (MAE) for the FG NET (limited) dataset was 3,446, while the MAE for the UTKFace (unconstrained) dataset was 4,867. The model was fine-tuned for the age group classification job using the Adience dataset, and its overall accuracy was 61.4 percent. Khan *et al.* [22] A model has an 82.3 percent identification rate after 10575 trials. The LFW face database was used for face verification, with 6,000 pairs of face comparison trials producing an average recognition rate of 84.5 percent. Peng *et al.* [23] used a modified 3D CNN architecture with a ResNet backbone to collect dynamic actions of the video data. For mouth motion, the structure received an F1 score of 82 percent, and for facial palsy grading, it received an F1 score of 88 percent. Storey *et al.* [24] The 'deep coupled ResNet model' has been proposed by researchers as a model for low-resolution FR. Trunk and branch networks were used to extract discriminative characteristics shared by face photos of various resolutions. With varying probe sizes and datasets, the suggested model obtained 93.6 percent -98.7 percent accuracy in face verification. The remaining sections are organized as follows: section 2 the method. While section 3 results and discussion. Finally, section 4 conclusion.

2. METHOD

The proposed model is built on a pre-trained MobileNet architecture that is training and testing on a dataset called Pins face recognition, which contains photos obtained from Pinterest, cropped and it is a publicly available dataset the total dataset consists of 105 celebrities and 17,534 faces. This dataset consists of 105 classes representing celebrity's persons male and female and each class has a different number of faces this work proposes a person identification system via facial biometrics based on a deep neural network. The proposed model employ pre-trained CNN models to find a model relationship and using a set of input and output data, find the system ordering and approximation of the unknown function. The main aim of the proposed model is to predict the person's identity by his/her face. The general block diagram of the proposed Person Identification model can be illustrated in Figure 1. The proposed system performs with three main stages respectively, pre-processing stage, feature extraction stage to extract the salient features of the input image then person identification stage based on these features. Each stage performed a specific function and employed some layers which are used to perform this specific purpose.

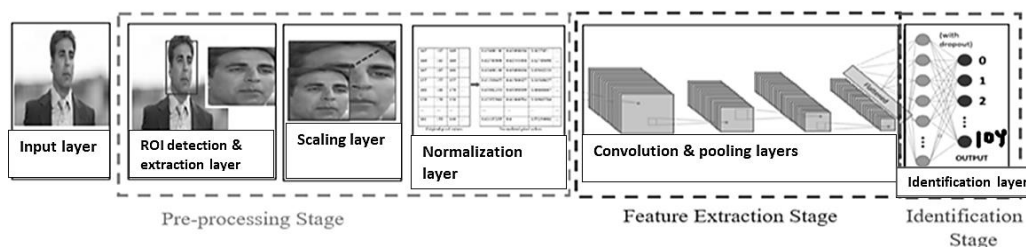


Figure 1. The block diagram for proposed system

2.1. Pre-processing stage

The pre-processing stage completes all necessary preparations on the input data to make it suitable for use in the suggested classification prediction stage. This stage consists of a region of interest (ROI) detection and extraction, resizing (scale), and normalization steps. The ROI can be described as an image part

representing the boundaries of the object being considered. According to the features employed by the proposed system (facial features), the interesting region implemented with the proposed system is the face region. Viola-Jones object detection framework based on Haar features is used in this step to detect the face region and then extract this region to be used later in other stages.

Scaling the dimensionality of cropped images in samples is a more crucial step in providing generality to samples and making them appropriate for deep learning prediction models. Image scaling translates to image resizing, which entails reconstructing an image from a one-pixel grid to another by changing the total number of pixels in images of samples. For better results, image resizing uses one of the image scaling algorithms that employ the interpolation of known data (the values at surrounding pixels) in an image to estimate missing values at missing points.

Neural network models deal with small weight values during inputs are processed. The learning process can be disrupted or slowed down due to large integer value inputs. The normalization process changes the intensity range of input values with normally viewed so that each input value has a value range between 0 and 1. By dividing all values by the biggest value, the input value is normalized (which is 255).

2.2. Feature extraction and identification stages

Images of various illuminations may be found in standard datasets, as well as, some of these photos may be taken in low lighting or with low-quality cameras, and as a result of these occurrences, the photos are dark, unclear, or have a lot of noise. Deep neural networks presented an efficient and promising method to extract robust features in different lighting conditions. Therefore, these stages are achieved based on the proposed pre-trained MobileNetV2 Network Structure. To obtain the benefits of the pre-trained CNN Models, there are many applications and systems based on the modification of these models to gain accurate results by only modifying some layers in the original architecture. This research is using MobileNetV2 [25] as a backbone architecture to identify the states of a person's face. The distinguishes MobileNetV2 is that it requires very little computation power to run or apply transfer learning. The architecture of the Modified MobileNetV2 model is revealed in Figure 2.

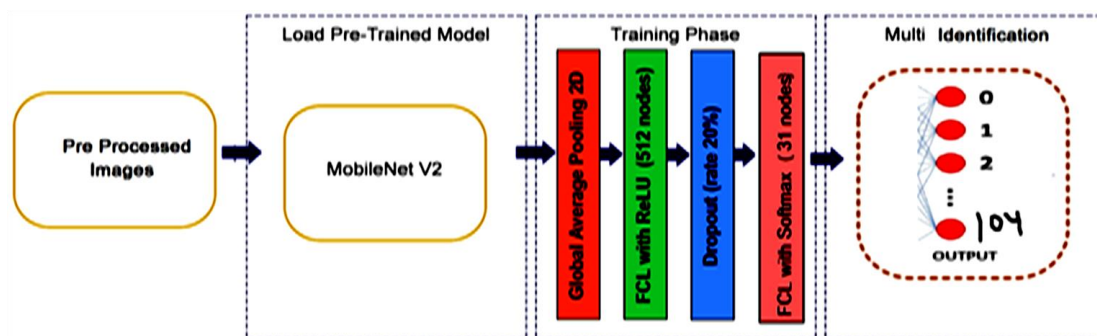


Figure 2. Pre-trained MobileNetV2 CNN architecture

Figure 2 shows that the base model is a MobileNetV2 network, ensuring the head FC layer sets are left off. In the training stage, these steps have been followed: firstly, generate class names. The next step is to freeze all layers except the final ones and train for a few epochs until you reach Plateau (no improvement stage). Finally, unfreeze all of the layers and train all of the weights while gradually lowering the learning rate until you reach a Plateau.

The identification stage is concerned with finding a model relationship and using fully connected layers, also known as Dense, and applying the "ReLU" activation function to these layers, except the final fully connected layer, called the output layer or decision-making layer, which implements a SoftMax function, determining the system orders and approximation of the unknown function. The Dropout layer additionally adds roughly 20% of the neurons picked at random in this step.

3. RESULTS AND DISCUSSION

The performance of the training and evaluation modes is evaluated with two key points (metrics), which are accuracy and loss functions. A quick way to understand the behavior for the learning of the proposed model on a specific dataset is by evaluating the training and a validation dataset for each epoch and

plotting the results as can be shown in Figure 3. In the Figure 3 part (a) show model accuracy and part (b) show the model loss for the proposed system.

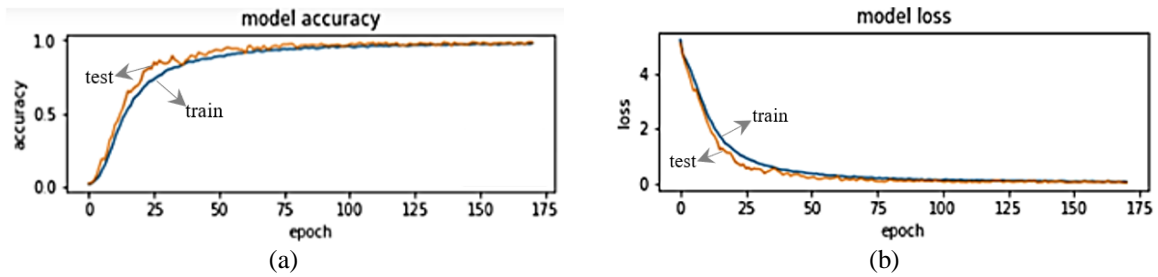


Figure 3. Accuracy and loss function of proposed model (a) model accuracy and (b) model loss

3.1. Evaluation model

Deep learning models are stochastic, which means that each time the same model is fitted to the same data, it will provide different predictions and, as a result, have varied overall abilities. The model is evaluated using the k-fold cross-validation approach for estimating model skill (controlling for model variance), which yields different results when the same model is trained on different data. The second procedure is estimating a stochastic model's skill (model stability control); different results when the same model is trained on the same data; by repeating the experiment of evaluating a non-stochastic model multiple times and then calculating the mean of the estimated mean model skill, the so-called mean. Results of precision, recall, and F1-measure of the model have been in Table 1.

Table 1. Metrics for evaluation of the pre-trained CNN model

Accuracy	Classification Error	Recall	Precision	F1-measure
0.9905	0.0095	0.9906	0.9911	0.9906

3.2. K-folds cross-validation

By testing machine learning models on a specific resampling data set based on a single parameter called k, the cross-validation technique evaluates the skill of a machine learning model on unseen data. In other words, during the training of the model, the limited sample is utilized to estimate how the model is typically predicted to predict unknown data. The procedure of 10-fold cross-validation with training and testing is shown in Figure 4.

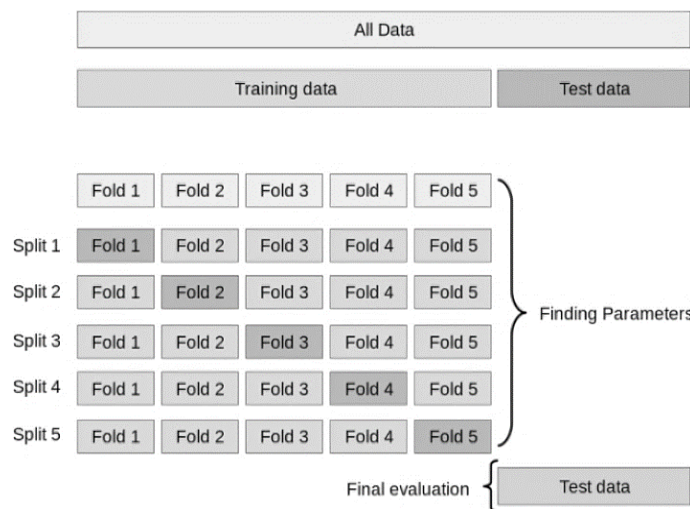


Figure 4. Cross-validation procedure with 5-folds

The proposed models are given $k = 10$, for each value of K , it will split dataset into Tr (80%) + Va (10%) = 90% and $Te = 10\%$, recording the testing performance according to the metric used (adopted accuracy). Finally, the average of the performance is computed to represent the final result. The 10-Fold CV implementation and Table 2 summarize the findings of the trial.

Table 2. 10-Fold CV for the proposed model

No of fold	Accuracy of each fold	Model accuracy
1	99.05%	
2	99.05%	
3	99.05%	
4	99.05%	
5	99.05%	
6	99.05%	99.5% (+/- 0.16%)
7	99.05%	
8	99.05%	
9	99.05%	
10	99.05%	

4. CONCLUSION

The identification system using deep neural networks can be considered the best approach to achieve high accuracy and give better results than other traditional approaches in terms of accuracy and loss functions. Period of this study, the main purpose of this study is to show the proposed system successfully identifies the person using his or her image. The ability to work with raw data provided by CNN saved the time and effort required for the vexing pre-processing required improving image quality. The ReLU activation function integrated with the convolution deep neural network layer (CNN) is used to extract salient features and ignore weak features, resulting in dealing with sample noise. The ReLU accomplishes this by removing all noise elements from the sequence and keeping only those with positive values. Going to add the Batch Normalization layers after the convolutional layers ensures that the convolutional assets are obeyed, reducing the pose variation problem in the sample. The batch normalization layer (BN) normalizes various elements of the same feature map at different places in the same way, regardless of their direct spatial surroundings.




REFERENCES

- [1] Z. Mortezaie and H. Hassanpour, "A survey on age-invariant face recognition methods," *Jordanian Journal of Computers and Information Technology*, vol. 5, no. 2, pp. 87–96, 2019, doi: 10.5455/jjcit.71-1554841475.
- [2] D. T. Long, "A lightweight face recognition model using convolutional neural network for monitoring students in E-learning," *International Journal of Modern Education and Computer Science*, vol. 12, no. 6, pp. 16–28, Dec. 2020, doi: 10.5815/ijmecs.2020.06.02.
- [3] S. F. Kak, F. Mahmood Mustafa, and P. Valente, "A review of person recognition based on face model," in *IEC2018 Proceedings Book*, 2018, pp. 1–19. doi: 10.23918/iec2018.01.
- [4] S. Rachad, B. Nsiri, and B. Bensassi, "System identification of inventory system using ARX and ARMAX models," *International Journal of Control and Automation*, vol. 8, no. 12, pp. 283–294, Dec. 2015, doi: 10.14257/ijca.2015.8.12.26.
- [5] C. M. Pappalardo and D. Guida, "System identification algorithm for computing the modal parameters of linear mechanical systems," *Machines*, vol. 6, no. 2, 2018, doi: 10.3390/machines6020012.
- [6] H. Mandalapu *et al.*, "Audio-visual biometric recognition and presentation attack detection: A comprehensive survey," *IEEE Access*, vol. 9, pp. 37431–37455, 2021, doi: 10.1109/ACCESS.2021.3063031.
- [7] M. Zulfiqar, F. Syed, M. J. Khan, and K. Khurshid, "Deep face recognition for biometric authentication," in *2019 International Conference on Electrical, Communication, and Computer Engineering (ICECCE)*, Jul. 2019, no. August, pp. 1–6. doi: 10.1109/ICECCE47252.2019.8940725.
- [8] N. D. Al-Shakarchy and I. H. Ali, "Open and closed eyes classification in different lighting conditions using new convolution neural networks architecture," *Journal of Theoretical and Applied Information Technology*, vol. 97, no. 7, pp. 1970–1979, 2019.
- [9] S. Adnan, F. Ali, and A. A. Abdulmunem, "Facial feature extraction for face recognition," *Journal of Physics: Conference Series*, vol. 1664, no. 1, Nov. 2020, doi: 10.1088/1742-6596/1664/1/012050.
- [10] M. Wang, L. Song, K. Sun, and Z. Jia, "F-2D-QPCA: A quaternion principal component analysis method for color face recognition," *IEEE Access*, vol. 8, pp. 217437–217446, 2020, doi: 10.1109/ACCESS.2020.3041847.
- [11] F. Wang, F. Xie, S. Shen, L. Huang, R. Sun, and J. Le Yang, "A novel multiface recognition method with short training time and lightweight based on ABASNet and H-Softmax," *IEEE Access*, vol. 8, pp. 175370–175384, 2020, doi: 10.1109/ACCESS.2020.3026421.
- [12] J. Tang, Q. Su, B. Su, S. Fong, W. Cao, and X. Gong, "Parallel ensemble learning of convolutional neural networks and local binary patterns for face recognition," *Computer Methods and Programs in Biomedicine*, vol. 197, 2020, doi: 10.1016/j.cmpb.2020.105622.
- [13] L. Shi, X. Wang, and Y. Shen, "Research on 3D face recognition method based on LBP and SVM," *Optik*, vol. 220, Oct. 2020, doi: 10.1016/j.ijleo.2020.165157.
- [14] R. Rameswari, S. Naveen Kumar, M. Abishek Ananth, and C. Deepak, "Automated access control system using face recognition," *Materials Today: Proceedings*, vol. 45, pp. 1251–1256, 2021, doi: 10.1016/j.matpr.2020.04.664.




- [15] I. Omara, A. Hagag, S. Chaib, G. Ma, F. E. Abd El-Samie, and E. Song, "A hybrid model combining learning distance metric and DAG support vector machine for multimodal biometric recognition," *IEEE Access*, vol. 9, pp. 4784–4796, 2021, doi: 10.1109/ACCESS.2020.3035110.
- [16] E. Zangeneh, M. Rahmati, and Y. Mohsenzadeh, "Low resolution face recognition using a two-branch deep convolutional neural network architecture," *Expert Systems with Applications*, vol. 139, Jan. 2020, doi: 10.1016/j.eswa.2019.112854.
- [17] N. Singhal, V. Ganganwar, M. Yadav, A. Chauhan, M. Jakhar, and K. Sharma, "Comparative study of machine learning and deep learning algorithm for face recognition," *Jordanian Journal of Computers and Information Technology*, vol. 7, no. 3, pp. 313–325, 2021, doi: 10.5455/JJCIT.71-1624859356.
- [18] F. Zhang and F. Wang, "Exercise fatigue detection algorithm based on video image information extraction," *IEEE Access*, vol. 8, pp. 199696–199709, 2020, doi: 10.1109/ACCESS.2020.3023648.
- [19] B. Cao, S. Shan, X. Zhang, and W. Gao, "Baseline evaluations on the CAS-PEAL-R1 face database," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 3338, no. May 2014, pp. 370–378, 2004, doi: 10.1007/978-3-540-30548-4_42.
- [20] G. Kahl, "Family-specific domain (FSD)," in *The Dictionary of Genomics, Transcriptomics and Proteomics*, vol. 06, no. 02, Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA, 2015, p. 1. doi: 10.1002/9783527678679.dg04283.
- [21] S. Khan, M. H. Javed, E. Ahmed, S. A. A. Shah, and S. U. Ali, "Facial recognition using convolutional neural networks and implementation on smart glasses," in *2019 International Conference on Information Science and Communication Technology (ICISCT)*, Mar. 2019, no. May 2020, pp. 1–6. doi: 10.1109/CISCT.2019.8777442.
- [22] S. Khan, E. Ahmed, M. H. Javed, S. A. A. Shah, and S. U. Ali, "Transfer learning of a neural network using deep learning to perform face recognition," in *2019 International Conference on Electrical, Communication, and Computer Engineering (ICECCE)*, Jul. 2019, pp. 1–5. doi: 10.1109/ICECCE47252.2019.8940754.
- [23] S. Peng, H. Huang, W. Chen, L. Zhang, and W. Fang, "More trainable inception-ResNet for face recognition," *Neurocomputing*, vol. 411, pp. 9–19, Oct. 2020, doi: 10.1016/j.neucom.2020.05.022.
- [24] G. Storey, R. Jiang, S. Keogh, A. Bouridane, and C.-T. Li, "3DPalsyNet: A facial palsy grading and motion recognition framework using fully 3D convolutional neural networks," *IEEE Access*, vol. 7, pp. 121655–121664, 2019, doi: 10.1109/ACCESS.2019.2937285.
- [25] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Jun. 2018, pp. 4510–4520. doi: 10.1109/CVPR.2018.00474.

BIOGRAPHIES OF AUTHORS



Marem H. Abdulabas    Holds a bachelor's degree in Computer Science from Karbala University, in 2012. Currently, she studies a master's degree in Computer Science at the Faculty of Computer Science and Information Technology at Karbala University. She can be contacted at email: marem.h@s.uokerbala.edu.iq.



Noor D. Al-Shakarchy    Faculty member of Computer Science and Information Technology College, Kerbala University, Karbala, Iraq. She did get my Ph.D. From The University of Babylon. B.Sc. and M.Sc at the University of Technology, Computer Science and Information Systems/Information Systems Department, Bagdad, Iraq in 2000 and 2003 respectively. Her research interests include computer vision, pattern recognition, Information Security, Artificial intelligence, and deep neural networks. She can be contacted at email: noor.d@uokerbala.edu.iq.