International Journal on
"Technical and Physical Problems of Engineering"
(IJTPE)
Published by International Organization of IOTPE

ISSN 2077-3528
IJTPE Journal
www.iotpe.com
ijtpe@iotpe.com

December 2022 | Issue 53 | Volume 14 | Number 4 | Pages 219-224

AGE ESTIMATION UTILIZING DEEP LEARNING CONVOLUTIONAL NEURAL NETWORK

F.K. Al Jibory [1] O.A. Mohammed [2] M.S.H. Al Tamimi [3]

1. Ministry of Education, Baghdad, Iraq, csd1b0037@gmail.com
2. Diyala Governorate Council, Baqubah, Iraq, omaramerpc@gmail.com
3. Department of Computer Science, College of Science, University of Baghdad, Baghdad, Iraq, m_altamimi75@yahoo.com

Abstract- Estimating an individual's age from a photograph of their face is critical in many applications, including intelligence and defense, border security and human-machine interaction, as well as soft biometric recognition. There has been recent progress in this discipline that focuses on the idea of deep learning. These solutions need the creation and training of deep neural networks for the sole purpose of resolving this issue. In addition, pre-trained deep neural networks are utilized in the research process for the purpose of facial recognition and fine-tuning for accurate outcomes. The purpose of this study was to offer a method for estimating human ages from the frontal view of the face in a manner that is as accurate as possible and takes into account the majority of the challenges faced by existing methods of age estimate. Making use of the data set that serves as the foundation for the face estimation system in this region (IMDB-WIKI). By performing preparatory processing activities to setup and train the data in order to collect cases, and by using the CNN deep learning method, which yielded results with an accuracy of 0.960 percent, we were able to reach our objective.

Keywords: Estimation, Age, Deep Learning, IMDB, CNN.

1. INTRODUCTION

Real-world applications that require age data include social comprehension, biometric data, proof of identity, surveillance, human-computer interaction, digital consumer and crowd behavioral research, internet ads, and item recommendation. Automatically determining a person's age from a facial photograph is a challenging challenge to tackle, despite its many uses. This is largely due to the fact that there are several causes of intra-class variances in the face pictures of humans, which limits the applicability of these models in real-world applications [1]. Over the course of many decades, face analysis tasks have become an important topic of study. Age recognition activities might be characterized as those that convey some basic yet vital information about a face [2]. These data have the potential to be useful in a wide variety of applications, including self-monitoring [3].

Recent years have seen an explosion in interest in the process of automatically estimating a person's age from frontal-type facial photographs, as it has numerous applications in a variety of face analysis problems [1].

Over the course of several decades, facial analysis tasks have emerged as a prominent area of research interest. The calculation of unrestricted age by software can be complicated by a number of factors, including inadequate resolution, bright illumination, diverse human races, particularly dark-skinned humans like those of African ancestry, or closely related forms. Age recognition is one of these occupations. Simple observations that can shed light on important aspects of a person's appearance. These data have the potential to be useful in a wide variety of applications, including self-monitoring [4]. Standard machine learning (ML) approaches and deep learning (DL) methods have been accepted by researchers as having significant potential in age recognition applications, according to the methodology currently in use [5]. Approaches based on deep learning have also met with a great deal of success in a variety of different contexts. CNN is the most well-known use of deep learning, and sufficient quantities of training data are required. When there are more training data points, it is easier to generalize the network more thoroughly.

On the other hand, appropriately disaggregated data for age and gender recognition based solely on facial features may be limited and difficult to get. When training photos are lacking, there is a greater chance that overshooting may occur. There is a wide variety of approaches that can be taken to reduce the impact of the issue of over-fitting [6]. Human frontal faces account for the majority of face identification methods in the field of image processing, which has made them a hot topic for recent research. In the field of computer vision, face recognition has received a lot of attention. Identity verification, intelligent visual surveillance and automated immigration processing systems are all examples of practical applications where vision is an issue [7]. It is really a major issue in many real applications, such as identity verification, intelligent monitoring, and automated immigration clearance systems [8].

It is still difficult to use face systems to identify people in real-world circumstances, as several case studies have shown [9]. The face is a flexible body that can take on a variety of different looks depending on factors such as age, angle, facial expression, and, most significantly, the amount of light that is falling on it. This is the fundamental reason why this occurs. In addition, there are a great deal of factors that continue to influence one's capacity to recognize faces [10], such as blockages and postures [18].

2. AGE ESTIMATION SYSTEM

Classifications of ages or age estimates Researchers have been working on age groups ever since the first Face Recognition Vendor Test (FRVT) was published in 2002. The findings of that test focused on the influence of demographic variables on facial recognition performance. Since then, researchers have found some inconsistent results when looking at age groups. It has been proven in almost all previous studies that it is simpler to identify senior adults as opposed to younger people [11]. Facial recognition technology, also known as FERET, has been shown to have a harder time recognizing younger faces than older ones. The reasoning for this is based on the assumption that younger faces have fewer distinguishing characteristics. Previous scholars have conducted many different kinds of studies to investigate the ways in which sub-branch factors could influence a person's capacity for facial recognition. They examined three well-known algorithms, including Principal Component Analysis, and came to the conclusion that younger individuals, particularly those of childhood, are more challenging to recognize than older individuals, and this is true regardless of the method that is used to recognize people. There is now a clearly distinguishable viewpoint. demographics utilizing [12][19].

3. RELATED WORK

The estimation of human age has been intensively researched over the last many years, and this thesis will focus on a few of these studies. Using a facial recognition database, the system trains a deep CNN model. Regarding the Adience database, [13] describes a possible approach for estimating ages. On the database, this model was trained. This article provides three important additions to the past study on the issue. (1) Using a CNN-trained face recognition model may help increase age estimate accuracy, according to this research; (2) The issue of overfitting may be avoided by using a CNN that has been pretrained on a large database for face recognition; and (3) Pre-training of the VGG-CNN used to establish data modalities affects the age estimate model's performance in addition to the number of training pictures and participants in the training database. As a consequence, the system has an accuracy rating of 59.90.

In [14], the authors give a preliminary examination into how postprocessing techniques might be used to enhance the performance of pre-trained deep learning models. In order to extract characteristics from the facial picture input, a variety of convolutional neural networks (CNNs) that have been previously trained are used. Finally, a feed-forward neural network is used to estimate the individual's age by fusing together various features of the data (FFNN). Audience Benchmark of Unfiltered Faces for Age Classification was used to assess the method's performance using a dataset that was gathered by the designers themselves and had non-ideal samples rotated in a controlled manner. Age estimation using this approach was superior to or on par with that of cutting-edge methods, and it functioned well even when circumstances weren't ideal. The results also illustrate that CNNs trained with large datasets may achieve adequate accuracy on a range of validation pictures without the need for extra fine-tuning. Thus, it was concluded that the MAE had dropped to 0.46.

To circumvent the difficulties of estimating age, [15] offers the Fusion Network design for CNN (Fusion-Net). In addition to the whole face picture, FusionNet captures numerous age-specific facial patches as input in order to emphasize the qualities that are unique to each age. Experiments indicate that the FusionNet model beats other cutting-edge models with respect to the MORPH II benchmark. This entails offering FusionNet as a solution to the problem of age estimate based on facial features. Inputs for the model include not just the user's actual face, but also a variety of aging patches. There may be a way to think of the face patches as network shortcuts that boost the learning efficacy of certain age traits. As previously mentioned, this is due to the input face patches including age-specific data. Experiments have shown that the system performed better on the MORPH II benchmark than current CNN-based state-of-the-art techniques. Approximately 80% of the total dataset is randomly assigned for training, with the remaining 20% for testing. There is no repeat of material in the training or testing sets. It was possible to undertake statistical analysis because of the use of twenty distinct divisions (with the same ratio but different distribution). An astounding 96.37 percent of the time, the treatment was a complete success.

Support vector regression and convolutional neural networks are both used by the authors of [16] to determine an individual's age from their face. After CNNs have been trained for representation learning, metric learning, and SVR, SVR is conducted on the newly obtained features. As a consequence, the lack of large datasets with age annotations may be circumvented by first training the CNN for face recognition. The suggested method outperforms current state-of-the-art techniques on the MORPH-II and FG-Net datasets. Due to the obvious limited size of the dataset, it's imperative that retraining SVR layers rather than CNN layers be used to demonstrate domain adaptability for datasets like FG-Net.

The system is up and running around 84.18 percent of the time. The author of [17] recommends preprocessing photographs and implementing data augmentation techniques, as well as modifying the output layer, including age estimate methods like classification and regression, and emphasizing important attributes. Noise vectors like ambient information that is unrelated to the faces in the image may have a similar influence on age estimation accuracy as existing approaches. The conclusion may be drawn using an MAE of 3.81 and an accuracy of 70.19.

4. ESTIMATION AGE PROPOSED SYSTEM

The method that has been proposed offers a human age estimation system that is based on a large data collection that may contain as many as 150,000 photos. The availability of a wide variety of human circumstances is one of the defining characteristics of IMDB-WIKI. Utilizing a multitasking convolutional neural network (CNN) method, it should be noted that the face image of a human being is the input to the suggested system. This face image is required by the system so that it can estimate the person's age. Figure 1 is a diagrammatic representation of the technique that has been presented for determining the ages of human beings.

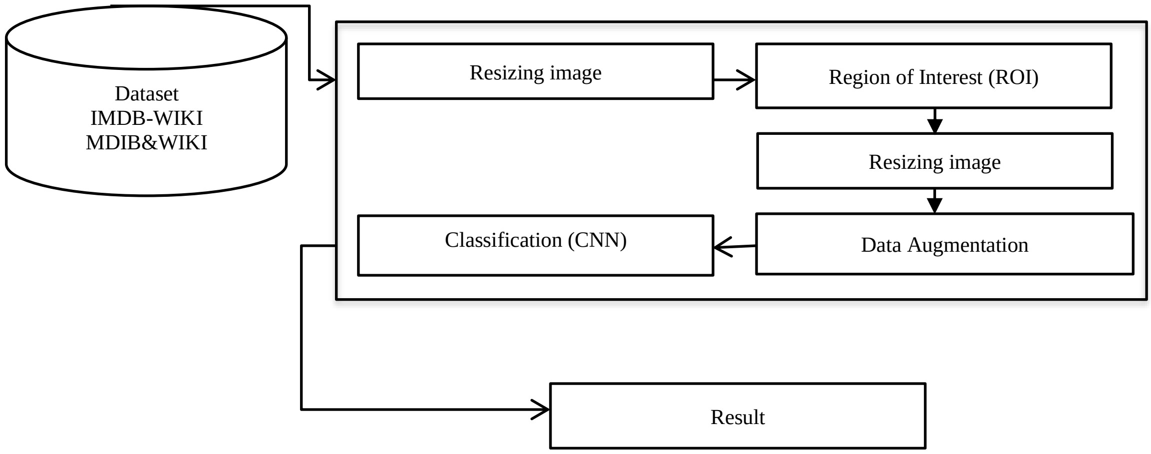

Figure 1. Systematized Block Diagram of proposed system

4.1. Dataset in the Proposed System

Since publicly accessible face picture collections are often small to medium in size and lack age information, we compiled a massive collection of celebrities. Designers compiled a list of the top one hundred thousand actors on IMDb and (automatically) browsed their profiles for birthdates, names, genders, and photos. We crawled all Wikipedia profile photographs that included identical meta data. Without a timestamp, designers erased pictures (the date the photo was taken). Each picture might be assigned an actual biological age if we believe that a single-faced image adequately depicts the actor and that the time stamp and date of birth are true. Despite our best efforts, we cannot guarantee the accuracy of the age information. These photographs were obtained from long-running films with erroneous time stamps. Photos (Figure 2) from 20,284 IMDb profiles and 62,328 from Wikipedia total 523,051 face images [1]. This is the largest publicly accessible dataset of face photos with gender and age labels that we are aware of. The analysis (Figure 3) highlights the significance of researching and using deep learning for the estimate of apparent age in still facial photos.

Figure 2. Samples in the IMDB-WIKI dataset [13]

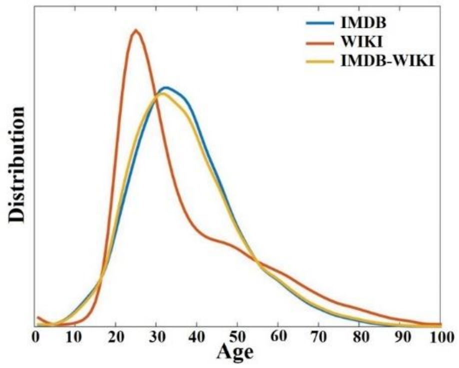

Figure 3. Analysis of IMDB-WIKI dataset [13]

4.2. Region of Interest (ROI)

The process of extracting an important area within the group, which consisted of (68) areas or points that were identified based on practical experience, where the benefit of this process or technology is to limit the so-called deduction area called ROI within the center of the face. Extracting an area within the face based on the use of (landmark) techniques. The dlib library, one of the most widely used facial recognition packages, is used in this approach. It is possible to accurately and reliably recognize people's faces in a variety of lighting settings thanks to the Max-Margin (MMOD) CNN face detector.

This library is the foundation of the ROI procedure, since it uses the 68 points stated in the proposed method to find the center of the face. Using shape prediction algorithms and landmarks that localize and depict the major portions of the face (eyes, eyebrows, nose, mouth, and jawline) in the context of facial characteristics, the objective is to identify significant facial structures on the face.

4.3. Resizing Image

The resizing procedure is vital since it strives to bring the image's dimensions together. The dataset in the proposed system is vast, and most importantly, it varies widely in size, with sizes ranging from (the most petite 71×70) to (300×300), and the image size cannot be adjusted. The image size was changed to 250×250, and this number was approved based on the experiment on different sizes and their impact on accuracy. Before the ROI process, to implement the process correctly, and as a result, the ROI output will be images of very different sizes, the image size was changed to 224×224.

4.4. Data Augmentation

By growing the data set in an artificial manner, the software model will be able to receive a great number of extra states and image modes. It is only training or teaching the model, but it is done. The system took advantage of data analytics (DA) to train the model on a huge number of cases that may not have been accessible in the data set. Epoch is the sole stage in which it is not included in the evidence collection; this is the single exception. Any extension idea is essentially dummy training for many examples that do not exist, and the only reason for it is to increase the system's accuracy in predicting age and gender in a variety of forms using the suggested system. The operations that are being performed are called Rotation, Construct, Brightness, and Flip.

4.5. Classification by Using CNN

This promising approach for data analysis employs complicated algorithms and artificial neural networks to enable robots and computers to learn from experience and to categorize and identify data similarly to the human brain. CNNs, which stands for "Convolutional Neural Networks", are a type of artificial neural network that play a significant role in Deep Learning and are used for the detection and classification of objects. The fact that numerical data are input into the algorithm is the most important component of the system that has been proposed. Therefore, the work that the algorithm does will be completed more quickly since it will not be necessary to transform the inputs to numbers and then utilize an array to process them as if they were images. As a direct consequence of this, the proposed system data number, which makes the algorithm more efficient in terms of the amount of time it takes.

Table 1. layer of proposed system description

No.	Layer name	Description
1	Input layer	CNN's input layer needs image data. Three-dimensional image data must be molded into a single column. 224×224 = 50,176, so multiply it by 1 before passing it in. The mean matrix of pixel values in the shape of (width, height, channels).
2	Rescaling	This layer rescales input values. This layer rescales each image by multiplying by a scale and adding an offset.
3	Normalization	Its key job is to avoid overfitting, which causes implementation modifications and unstable outcomes.
4	Zero Padding	The technique allows us to preserve the original input size. This is something that specifies on a per-convolutional layer. It is simply a process of adding layers of zeros to our input
5	Conv2D	a layer that creates a convolution kernel that is wind with layers input which helps produce a tensor of outputs that mean Applies a 2D convolution over an input signal composed of several input planes.
6	Batch Normalization	Applies a modification that keeps output mean around 0 and standard deviation near 1. Training and inference use batch normalization differently.
7	Rulle	The activation function in a neural network defines how the weighted sum of the input is transformed into an output from a node or nodes in a layer of the network.
8	Soft maxs	Softmax turns a vector of integers into a vector of probabilities, with probabilities proportional to scale.
9	Dense (pred_age)	It works on estimating the age of the images, i.e., people, and determining the age range between (1 to 101).

The CNN algorithm was utilized in the system that was proposed, and it had a 9-layer structure. The stretcher of CNN in the proposed system is depicted in Table 1. The classification step is the first step in the estimating system. This is because the classification step is the step that gives the findings based on the pre-processes. The age can be estimated to fall anywhere between one and one hundred years with the use of the system that has been proposed.

5. EVALUATION METRICS AND RESULT

The methodology that was suggested for the evaluation is going to be used. A confusion matrix [2, 3]. Which will be carried out in error rate and provides a summary of the number of events that were either correctly or incorrectly predicted by a classification model. It is common practice to refer to the results of counts calculated in a confusion matrix using the following terminology:

- Precision: It is measured using the data collection's standard deviation, whereas bias is measured using the disparity between the data collection's mean and the known value of the object that is being quantified. According to the following *precision* equation, precision is defined as the percentage of records in a group that the classifier has identified as belonging to a positive class that are, in fact, real positives.

$$Precision\,(p) = \frac{TP}{TP + FP} \tag{1}$$

By applying the Equation (1) for all images, the equation result is:

- Recall: *Recall* quantifies the proportion of positive examples predicted adequately by the classification, and its value is identical to the actual positive rate.

$$Recall(r) = \frac{TP}{TP + FN} \tag{2}$$

- F_1: The harmonic mean of accuracy and recall, as well as the flowing equation, denotes F_1, and the equation is.

$$F_1 = \frac{2rp}{r+p} = \frac{Two \times TP}{2 \times TP + FP + FN} \tag{3}$$

By applying the Equation (3) for all images, the equation result is:

• Accuracy: The classification *accuracy* will be tested on the search data collection. It is assumed that each category of membership is described.

$$Accuracy = \frac{\text{number of correctly classified image}}{\text{total nunber of image}} \times 100\% \quad (4)$$

By applying the Equation (4) for all images calculated by using the mean average, the equation is:

$$Accuracy = 0.98 \quad (5)$$

The Figure 4 shows the steps of the system work that show the measurement results. Table 2 also shows the result of the system. Figure 5 shows a summary of previous studies on the results and a comparison of the proposed system with these studies.

Table 2. Result of the system (Age)

Age	Result
Accuracy	0.960
precision	0.954
Recall	0.960
F_1	0.959

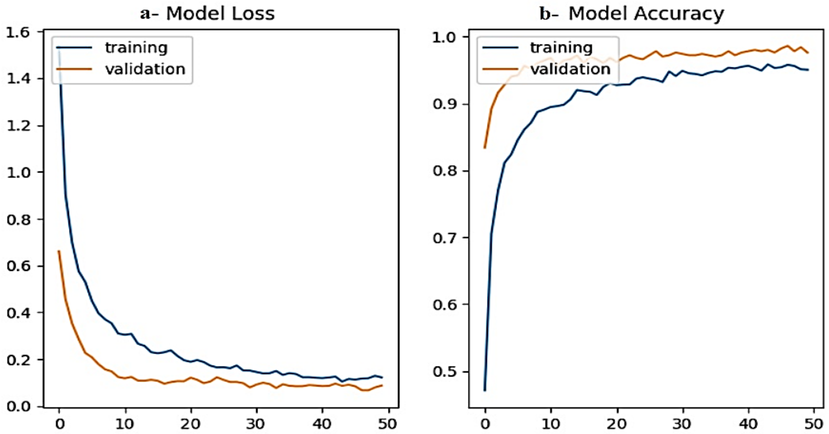

Figure 4. Training loss and accuracy model, (a) Model loss, (b) Model accuracy

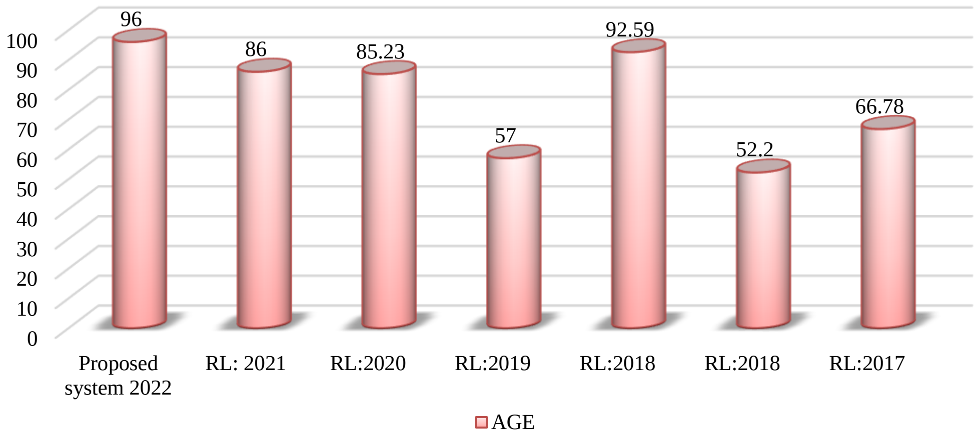

Figure 5. Histogram of summarized results for previous studies

6. CONCLUSIONS

To improve the accuracy of recognition, verification, or authentication, soft biometrics like age estimation using faces can be used as a supplement to traditional biometric technologies like facial recognition, fingerprint recognition, or iris recognition. Estimating someone's age is one way to spot a face that hasn't changed much throughout the years. Increasing the accuracy of the (central) solid biometric system can be accomplished through the use of iris recognition, hand geometry recognition, and fingerprint recognition. Using the CNN algorithm of deep learning, the proposed system attempted to achieve its goal of more accurately estimating the human age by basing its calculations on the images of the human's foreground. The (DA) procedure is one of the most crucial fundamental components in the area of face systems since it produces fictional situations that were not included in the data set in order to boost the system's accuracy. Because of this, the main idea is how to create and deal with massive data to obtain a high-quality classification process, and the reason for this is that it provides high accuracy in the field of face systems. It is recommended that, for future work, the YOLO v5s algorithm be utilized as an alternative to the algorithm that was previously used.

REFERENCES

[1] A. Abdolrashidi, M. Minaei, E. Azimi, S. Minaee, "Age and Gender Prediction from Face Images Using Attentional Convolutional Network", arXiv, Vol. 3, pp. 1-6, 2010.

[2] J.M. Al Tuwaijari, S.A. Shaker, "Face Detection System Based Viola-Jones Algorithm", The 6th International Engineering Conference Sustainable Technology and Development (IEC), pp. 211-215, 2020.

[3] A.K. Jain, A. Ross, S. Prabhakar, "An Introduction to Biometric Recognition", IEEE Trans. Circuits Syst. Video Technol., Vol. 14, No. 1, pp. 4-20, 2004.

[4] R. Ranjan, S. Sankaranarayanan, C.D. Castillo, R. Chellappa, "An All-In-One Convolutional Neural Network for Face Analysis", The 12th IEEE International Conference on Automatic Face and Gesture Recognition, pp. 17-24, 2017.

[5] C.Y. Hsu, L.E. Lin, C.H. Lin, "Age and Gender Recognition with Random Occluded Data Augmentation on Facial Images", Multimed. Tools Appl., Vol. 80, No. 8, pp. 11631-11653, 2021.

[6] W. Wang, J. Shen, S. Member, H. Ling, "A Deep Network Solution for Attention and Aesthetics Aware Photo Cropping", Vol. 8828, pp. 1-12, 2018.

[7] G. Guo, N. Zhang, "A Survey on Deep Learning-Based Face Recognition", Comput. Vis. Image Underst., Vol. 189, No. August, p. 102805, 2019.

[8] J.M. AL Tuwaijari, S.I. Mohammed, "Face Image Recognition Based on Linear Discernment Analysis and Cuckoo Search Optimization with SVM", International Journal of Computer Science and Information Security, Vol. 15, No. 11, IJCSIS, 2017.

[9] M. Lal, K. Kumar, R.H. Arain, A. Maitlo, S.A. Ruk, H. Shaikh, "Study of Face Recognition Techniques: A Survey", Int. J. Adv. Comput. Sci. Appl., Vol. 9, No. 6, pp. 42-49, 2018.

[10] S.H. Lee, S. Hosseini, H.J. Kwon, J. Moon, H.I. Koo, N.I. Cho, "Age and Gender Estimation using Deep Residual Learning Network", Int. Work. Adv. Image Technol. IWAIT 2018, pp. 1-3, 2018.

[11] S. Singh, S.V.A.V. Prasad, "Techniques and Challenges of Face Recognition: A Critical Review", Procedia Comput. Sci., Vol. 143, pp. 536-543, 2018.

[12] R. Ranjan, et al., "A Fast and Accurate System for Face Detection, Identification, and Verification", IEEE Trans. Biometrics, Behav. Identity Sci., Vol. 1, No. 2, pp. 82-96, 2019.

[13] S.H. Abdurrahim, S.A. Samad, A.B. Huddin, "Review on the Effects of Age, Gender, and Race Demographics on Automatic Face Recognition", Vis. Comput., Vol. 34, No. 11, pp. 1617-1630, 2018.

[14] K. Solanki, P. Pittalia, "Review of Face Recognition Techniques", Int. J. Comput. Appl., Vol. 133, No. 12, pp. 20-24, 2016.

[15] Z. Qawaqneh, A.A. Mallouh, B.D. Barkana, "Deep Convolutional Neural Network for Age Estimation based on VGG-Face Model", arXiv, No. 1, pp. 1-8, 2017. http://arxiv.org/abs/1709.01664.

[16] K. Zhang, et al., "Age Group and Gender Estimation in the Wild with Deep RoR Architecture", IEEE Access, Vol. 5, No. Cccv, pp. 22492-22503, 2017.

[17] H. Wang, X. Wei, V. Sanchez, C.T. Li, "Fusion Network for Face-Based Age Estimation", Proc. - Int. Conf. Image Process, ICIP, pp. 2675-2679, 2018.

[18] V. Greu, et al., "A New Algorithm for Optimizing IEEE 802. 11 Network Based on Interference Activity Evaluation Criteria", Int. J. "Technical Phys. Probl. Eng., Vol. 7, No. 4, pp. 55-61, 2015.

[19] R. Esmaeilzadeh, A.R. Milani, "Efficient Electric Price Forecasting Using Neural Networks", International Journal on Technical and Physical Problems of Engineering, Issue 32, Vol. 9, No. 3, pp. 30-34, September 2017.

BIOGRAPHIES

Farah Khiled Al Jibory was born in Iraq on November 3, 1992. He received the Bachelor degree in Computer Science from College of Economic Sciences University, Baghdad, Iraq in 2014. She received the Master degree in Computer Science from Baghdad University, Baghdad, Iraq in 2021. Currently, she is an assistant teacher in Iraqi Ministry of Education, Baghdad, Iraq. Her interests are deep learning technologies and branches of artificial intelligence.

Omar Amer Mohammed was born in Diyala, Baquba, Iraq on February 12, 1986. He received the Bachelor and Master degrees from Computer Department, College of Science, Diyala University, Iraq in 2008 and 2022, respectively. He is currently an employee in Diyala Governorate Council Office, Baqubah, Iraq. His interests are deep learning technologies and branches of artificial intelligence.

Mohammed Sabbih Hamoud Al Tamimi was born in Iraq on June 16, 1975. He received the B.Sc. degree in Computer Science from University of Al-Fateh, Tripoly, Libya in 1998, the M.Sc. degree in Computer Science from University of Science and Technology, Sana'a, Yemen in 2000, and Ph.D. degree in Computer Science from Faculty of Computing, University of Technology Malaysia, Johor Bahru, Malaysia in 2015. Currently, he is an Assist. Prof. in Department of Computer Science, College of Science, University of Baghdad, Baghdad, Iraq.