Predicting Age and Gender Using AlexNet

Qaswaa Khaled Abood [1], Farah khiled AL-Jibory [2]

[1]University of Baghdad computer science department, Iraq
[2]Ministry of Education, Karkh First Directorate of Education, Baghdad, Iraq

***Abstract* – Due to the availability of technology stemming from in-depth research in this sector and the drawbacks of other identifying methods, biometrics has drawn maximum attention and established itself as the most reliable alternative for recognition in recent years. Efforts are still being made to develop a user-friendly system that is up to par with security-system requirements and yields more reliable outcomes while safeguarding assets and ensuring privacy. Human age estimation and Gender identification are both challenging endeavours. Biomarkers and methods for determining biological age and gender have been extensively researched, and each has advantages and disadvantages. Facial-image-based positioning is crucial for many applications, including safety and security systems, border control, human engagement in sophisticated ambient analytics, and biometric identification. Determining a person's age and gender is a complex study method. With the advent of deep learning, the study of face systems has been completely transformed, and estimation accuracy is a crucial parameter for evaluating algorithms and their efficacy in predicting absolute ages. The UTKFace dataset, which serves as the backbone of the face estimating system, was used to assess the method. The eyes, cheeks, nose, lips, and forehead provide the foundation of this function. AlexNet achieves a 98% accuracy rate across its lifespan of system results.**

***Keywords* –Biometrics, Estimation, Age, Deep learning, IMDB, CNN.**

DOI: 10.18421/TEM121-61
https://doi.org/10.18421/TEM121-61

Corresponding author: Farah khiled AL-Jibory,
Ministry of Education, Karkh First Directorate of Education, Baghdad, Iraq
Email: csd1b0037@gmail.com

Received: 29 October 2022.
Revised: 10 February 2023.
Accepted: 24 February 2023.
Published: 27 February 2023.

1. Introduction

Real-world applications rely heavily on accurate age estimates, including human-computer interaction, biometrics, proof of identification, surveillance cameras, digital consumers, crowd analysis, online advertising, and product recommendation. Although there are numerous potential benefits, age estimation from facial images is difficult. This is because numerous factors may lead to intra-layer differences in a person's face image, generalizing from a training set challenges [1]. Over the years, facial analysis tasks have been thoroughly investigated. Recognizing a person's age is a simple, yet crucial approach to gaining additional insight into their appearance [2]. Self-monitoring is but one of the many applications of this information [3]. In recent years, automatic age estimation from frontal facial photographs has attracted much attention [1].due to the diverse challenges it may answer in face analysis. Facial analysis research has gained prominence over time. Software estimation of unrestricted age is challenging due to the identification that results from operations such as age recognition (poor precision, bright light, diverse human races, especially dark-skinned humans, such as the African race, or closely related forms) and immediately relevant details of a person's appearance [3]. There are numerous applications for this information, including introspection. Researchers have utilized existing technology to integrate machine learning (ML) and deep learning (DL) techniques in age recognition applications due to their enormous potential. There has been a tremendous deal of success with deep learning methods in numerous fields. AlexNet is the most well-known use of deep learning, and training data is vital to its success [4]. An increase in the number of learning algorithms facilitates the creation of a complete network generalization. Face and gender-based age-disaggregated data may be hard to obtain [5]. If inadequate training images are utilized, the likelihood of passing increases. Overfitting is a problem that has multiple solutions Recognizing human faces is a pressing problem in image processing, with most existing algorithms focusing on the frontal lobes of individuals' faces [6].

However, computer facial recognition is an established topic of study. Visibility is an issue for various real-world applications, including biometrics-based identity verification, intelligent video monitoring, and computer-assisted migration processing [7]. There are several real-world applications where this is a crucial vision challenge, including automated immigration clearance systems, intelligent visual monitoring, and identity verification [8]. Facial recognition is cumbersome to implement in various real-world contexts [9]. Different facial expressions, ages, viewpoints, and illumination levels contribute to how a person's face might be viewed. This is the primary reason why various factors, such as impediments and situations, hinder our ability to distinguish between faces [10]. This study examines a specific type of biometric technology whose primary purpose is to identify a person's age and gender.

2. Biometric System

A biometric system is a pattern recognition approach that takes in biometric data, generates a feature set from that data, and then compares that feature set to a template set stored in a database. It is possible to employ biometrics for authentication and identification purposes [11]. When it comes to biometrics, the most widely utilized feature for identifying people is a photograph of their faces, which is why this technique is so popular. Facial recognition has several potential uses, from the strictly controlled verification of "mug-shots" to the free-flowing identification of faces in a chaotic environment. There are two main ways to analyze facial features, which are commonly used [12].

I. Eyes, brows, nose, lips, chin, and their respective spatial connections and proportions to one another.

II. International face image investigation shows a weighted composite of numerous canonical faces.

Commercial face systems have good verification performance but require a stable background or special illumination to capture facial images. These systems also have trouble recognizing faces in photos taken from two distinct angles and under varying illumination. Without context, a person's face alone may not be enough to identify them among several identities securely. Soft biometrics are non-unique physical and behavioral traits that can be used to identify, verify, and describe people. Height, weight, and gender [12]. As an alternative to conventional biometrics, it is gaining favor as a non-invasive authentication method.

There are currently separate experimental systems that utilize soft biometrics. In the 18th century, the Bertillon method was used to detect criminals based on their appearance.

The lengths of a person's head, forehead, index finger, left foot, and cubit is anthropometric measurements. Body and facial geometry were utilized to classify these traits. Soft biometrics include factors such as age and gender. Gender information can be utilized in healthcare, intelligent environments, and biometric access control systems [13]. Intelligent space operations can be modified to create a more gratifying experience for each user by using gender data. A biometric system's efficiency can be enhanced by incorporating gender as a soft biometric. There are only two viable options for solving the problem of gender classification [14]:

I. Male (M).

II. Female (F).

It is possible to classify gender recognition systems as either appearance-based or non-appearance-based, depending on the features used as inputs to the classifier. Scientists use a specific feature of the human face to determine a person's age. At the same time, morphology or biology is the study of patterns in language. In this sense, the face is a group of flexible muscles and skin responsible for a wide range of facial textures. In this way, craniofacial morphology is defined as the study of the face and skull structure. Inevitably, the aging process will change a person's physical appearance. Specific craniofacial morphology, including facial texture, changes with age. Changes in skin texture are a common side effect of puberty [15].

3. Age and gender Estimation

The development of media and social networks has made automatic age and gender classification useful for various applications. While there have been significant improvements in the related tasks of face recognition and image classification, the performance of current approaches on real-world photographs is still significantly inadequate. Facial recognition technology (FERET) has shown difficulty recognizing younger faces more than older ones, as specific experiments on different databases have found [16]. This is due to the widespread belief that younger people's faces are less defined. Previous scholars have studied the influence of sub-branch characteristics on the accuracy of facial recognition. They looked at three popular algorithms (Principal Component Analysis) and concluded that younger people, especially those in childhood, are more difficult to distinguish than older people. Insight into the demography of the situation has allowed for a new perspective to emerge [17].

4. Related Work

This thesis summarizes recent articles on human age estimation.

In [18] A suggested system based on deep CNN trained on a facial recognition database is used to forecast audience database age data. This study makes three significant contributions. This study shows that a CNN trained on an extensive database for facial recognition can increase age estimation performance. The effectiveness of the VGG pre-training CNN depends on the number of training pictures and individuals in the training database, which affects the accuracy of the age estimate model. Predictability is 59.90% accurate.

In [19] Suggested system to investigate how post-processing can improve pre-trained deep neural networks. Several trained Convolutional Neural Networks (CNNs) extract facial features. Similar methods aggregate information, decrease feature dimensions, and forecast ages using Feed-Forward Neural Networks (FFNN). The Audience Benchmark of Unfiltered Faces for the Recognition System and a privately obtained dataset of non-ideal samples with controlled rotations were used to evaluate the approach. The age estimation method outperformed state-of-the-art techniques in less-than-ideal settings. The results show that CNNs trained on large datasets can achieve acceptable accuracy without fine-tuning.

In [20] The suggested System with AE is equal to 0.46. For age range estimation, CNN's Fusion Network comes highly recommended (Fusion-Net). FusionNet excludes the whole face in favor of age-targeted patches to emphasize the characteristics of a given age group. Based on experimental results, FusionNet is the most effective state-of-the-art network for MORPH II. FusionNet is used to estimate a face's age. The model also uses age-specific facial patches. Input face patches are network shortcuts that facilitate age-specific learning. The network outperforms CNN's cutting-edge MORPH II algorithms. 80% of the dataset is separated for testing and training. Training and testing use the same data—statistical analysis requires dividing data into 20 sections (with the same ratio but different distribution). The System was 96.37% efficient.

In [21] The suggested system based on the face age prediction article uses CNN and SVR (SVR). CNNs learn features through representation learning, metrics learning, and SVR. Facial recognition technology has the potential to close the generational data gap in enormous databases. The proposed strategy provided better results for both MORPH-II and FG-Net than the state-of-the-art methods. For limited data sets like FG-Net, retraining the SVR layer rather than the CNN could improve accuracy by 84.28%.

In [22] Using Classification and regression are merged. Essential traits are highlighted by preprocessing pictures and data augmentation. The output layer is changed. Reduce non-face-related noise vectors, including ambient information, to increase age estimation accuracy.

The result of the system is Accuracy = 7.19 with a Mean Absolute Error of 3.81.

In [23] Develop a better neural net loss function. A smooth loss function is employed (SGD) to optimize a system using stochastic gradient descent. If you want SGD to be broader and get closer to its optimal point, you should use the recommended loss function, which has a lower gradient than existing loss functions. Studies show that the suggested method outperforms the best existing algorithms, both specifically and in general. The system achieved 81.57 percent of its potential.

5. Facial system for indicating age and gender

The suggested system presents a method for predicting human age and gender based on a large and vital dataset, the UTKFace dataset. The availability of several human situations distinguishes it. The suggested method uses the Convolutional Neural Network (AlexNet) model to determine the age and gender of a person based on a face image. Figure 1 is a diagrammatic representation of the proposed approach for determining human lifespan:

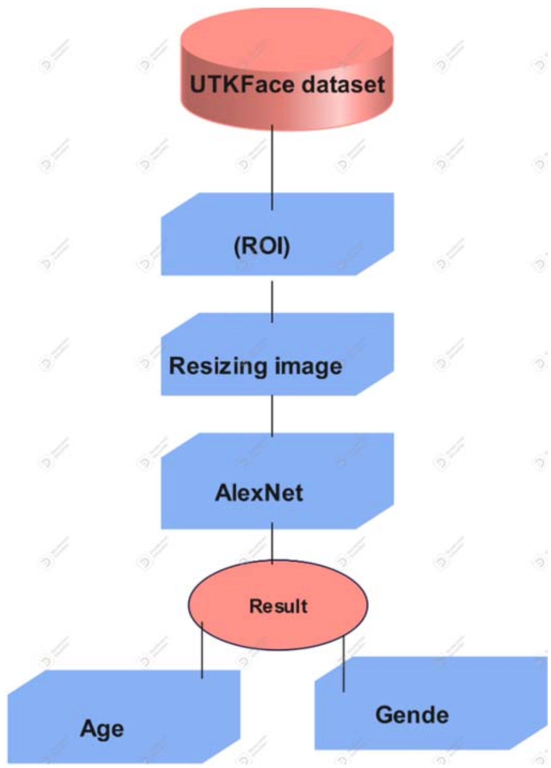


Figure 1 System Diagram.

5.1. Dataset in proposed system.

UTKFace is a long-term dataset (ranging from 0 to 116 years old). There are 20,000 photos in total, and they're all tagged with data like age, gender, and race.

Position, expression, lighting, visibility, and clarity are all variables in these images. This dataset makes face recognition, age estimation, age progression, and landmark localization more possible.

In the proposed system, use 8.1.1, where that means (8 equals 80% for training, one equals 10% for testing, and one equals 10% for validation). Obtaining these residencies has been applied through the use of a cross-validation technique. The following Figure (2) shows example image of dataset UTKFace.

Figure 2 Sample of Dataset in proposed system.

5.2. Region of Interest (ROI)

This methodology or technology can limit the so-called ROI deduction zone to the middle of the face. Dlib, a popular facial recognition library, does this. Face recognition based on CNN's max-margin object detection (MMOD) is reliable and accurate, recognizing faces in various lighting and camera angles. According to CNN and this library, the ROI method is based on 68 face points (eyes, eyebrows; nose; mouth; jawline). The purpose is to identify essential facial components using shape prediction to find the face's center. Figure (3) shows the steps of ROI.

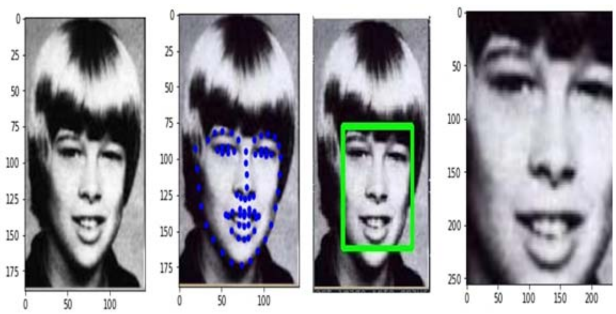


Figure 3 Steps of ROI.

5.3. Image resizing

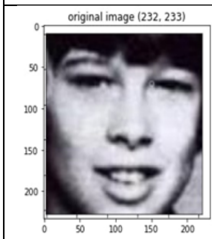
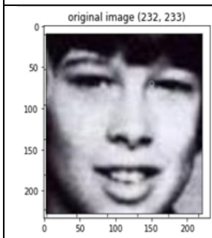
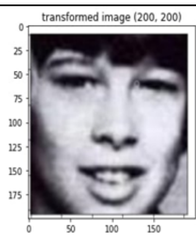
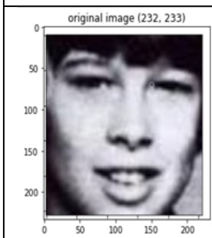
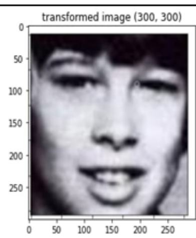
Vital as it is, the resizing procedure seeks to normalize the image's dimensions.

The suggested system's data set is extensive and variable in size, from the most miniature 70 *70 arrays to the most extensive 300 *300 array. Because the images produced by the ROI method will be of wildly varying dimensions, the image size was set at 200 * 200 and accepted after an experiment was conducted to determine the optimal resolution. Figure (4) shows the image in ROI stage and resizing stage.

Figure 4 Various sizes and resizing image.

The following Table (1) shows whay choosing resizing 200*200 as best in result of accuracy.

Table 1 Different Resizing Image

Original images	Resizing image
	
	
	

5.4. Classification Using AlexNet

The categorization of face photos or videos into preset age groups is a challenging process. A Convolutional Neural Network (AlexNet) model to analyze pictures and videos of human faces have recently demonstrated impressive results. On the (UTKFace) dataset, AlexNet was tested for age and gender classification. AlexNet -based deep learning models were employed to create a 200-dimensional representation that quantifies the age group's faces.

Due to better classification performance, AlexNet has tremendous success classifying people by age and gender. The majority of earlier studies enhanced classification accuracy by modifying network architecture. However, the classification of age groups and gender with an uneven data distribution has not yet been fully addressed. This work showed an AlexNet model capable of categorizing age groups with unequal data distribution. The following Figure (5) shows the Structure of the proposed System with trained model AlexNet.

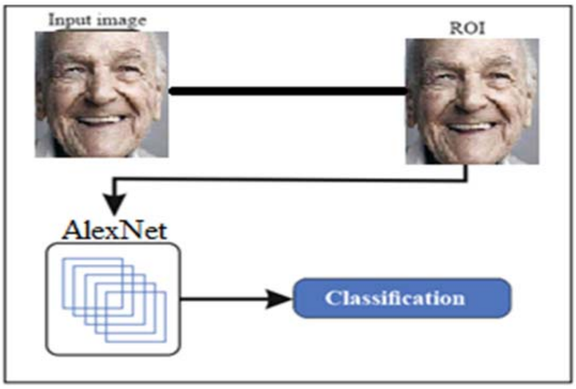


Figure 5 Structure of the proposed System with AlexNet.

5.5. Testing and training initialization

Each layer's initial weights are uniformly distributed to random variables with a zero mean and a 0.003 standard deviation, Gaussian distribution. Only the pictures and labels used to train the network are included in the training dataset.

The ground truth classes are represented as sparse binary vectors that serve as goal values for training. The target label vector for each training image has 91 classes, with 1 being the ground truth and 0 representing all other classes.

Overfitting can be avoided by using network architecture and two other methods. There are two 0.5-dropout layers in the network. There's a 50% probability that no neurons will fire. Training is performed with 64 images per batch using stochastic gradient descent at an initial learning rate of E3. An Nvidia graphics processing unit (GPU) with 4GB of video memory was used for training.

6. Evaluation Metrics and Result

The recommended evaluation approach will utilize the confusion matrix [24], [25], [8]. Classification models can either correctly estimate or wrongly forecast the number of occurrences, as shown in Figure (6). Confidence matrix counts are commonly referred to as:

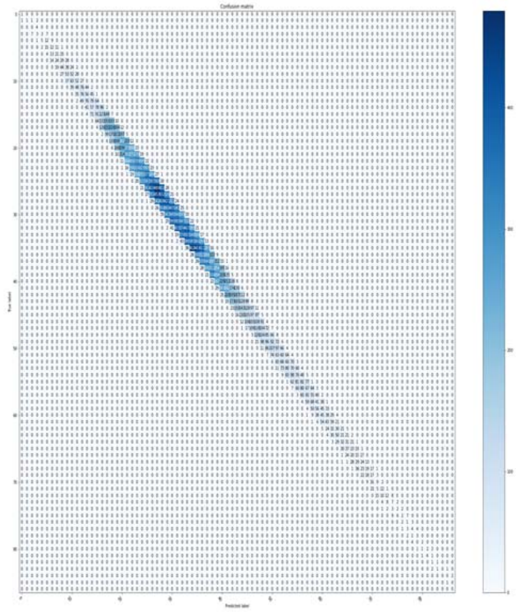


Figure 6 confusion matrices.

I. **Precision**: Mean, standard deviation, and item value are used by researchers to measure bias, whereas standard deviation is used to calculate dispersion in data collection. Accuracy measures how many true positives exist within the set of records that the classifier decided were true positives.

$$\text{Precision(p)} = \frac{\text{TP}}{\text{TP} + \text{FP}} \ ...(1)$$

II. **Recall**: The recall value is the actual positive rate, calculated by dividing the number of anticipated positive cases by the number of positive cases.

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

III. **F1**: Precision and recall harmonic mean is indicated by F1 with a matching equation.

$$\text{F1} = \frac{2\text{rp}}{\text{r} + \text{p}} = \frac{\text{Two} \times \text{TP}}{2 \times \text{TP} + \text{FP} + \text{FN}} \ .....(3)$$

IV. **Accuracy:** Search data will test categorization accuracy. Each membership category is described.

$$\textbf{Accuracy} = \frac{\textbf{namber of correctly classified image}}{\textbf{total namber of image}} * \%100 \ ..(4)$$

The following Table (2) displays the measurement results**:**

Table 2 Final results of deep learning in gender.

Type	Precision	Recall	F1-Score
0=female	0.96	0.98	0.96
1=male	0.99	0.98	0.98
Total accuracy		0.98	

The final deep-learning results are displayed in a Table (3) below.

Table 3 Final results of deep learning in age.

Class name	Precision	Recall	F1-Score
Adult	0.98	0.98	0.99
Child	0.84	1.00	0.92
Senior	0.97	0.98	0.99
Teenage	0.98	0.97	0.97
Young	0.98	0.99	0.99
Total accuracy		0.99	

The figure (7) displays an example of the result.

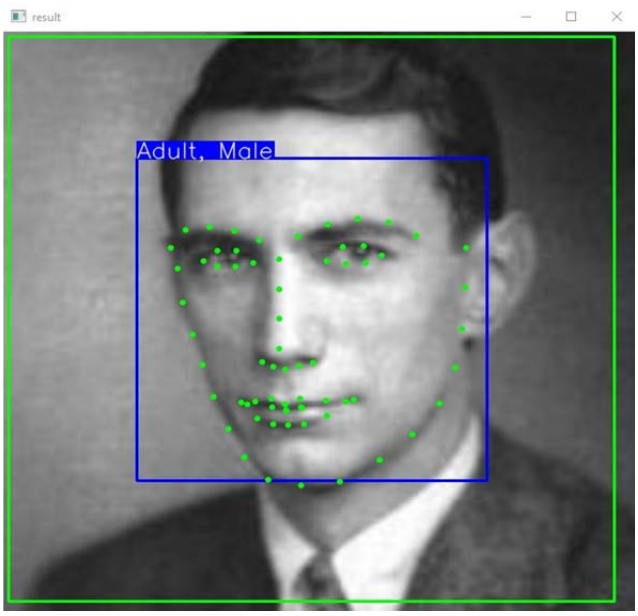


Figure 7 example of result.

7. Conclusion

With the development of information technology, age and gender estimate systems have grown dramatically, making them rich grounds for study and development.

In the scientific study, we introduced an age estimation method and sex determination using a multi-tasking CNN algorithm. Important feature: A system with two outputs works without hybridization. In the suggested system, work was based on exterior appearance and the ability to detect all human races and skin colors. Training the algorithm, UTKFace proved successful. The system offered a substantial change compared to external appearance-based systems. As a future effort, it is feasible to test the (You Only Look Once YOLO 5) algorithm because of its advantages, but it is under development. To improve recognition and distinction, also it should include a child-specific dataset.

References

[1]. Abdolrashidi, A., Minaei, M., Azimi, E., & Minaee, S. (2020). Age and gender prediction from face images using attentional convolutional network. *arXiv preprint arXiv:2010.03791*.

[2]. Abdurrahim, S. H., Samad, S. A., & Huddin, A. B. (2018). Review on the effects of age, gender, and race demographics on automatic face recognition. *The Visual Computer*, *34*(11), 1617-1630.

[3]. Akbari, A., Awais, M., Feng, Z., Farooq, A., & Kittler, J. (2021, January). *A flatter loss for bias mitigation in cross-dataset facial age estimation*. In 2020 25th International Conference on Pattern Recognition (ICPR) (10629-10635). IEEE.

[4]. AL-Jibory, F. K. (2021) Hybrid System for Plagiarism Detection on A Scientific Paper. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, *12*(13), 5707-5719.

[5]. Al-Tuwaijari, J. M., & Shaker, S. A. (2020). Face Detection System Based Viola-Jones Algorithm. In *2020 6th International Engineering Conference "Sustainable Technology and Development"(IEC)*, IEEE, 211–15.

[6]. AL-Tuwaijari, J. M., & Mohammed, S. I. (2017). Face Image Recognition Based on Linear Discernment Analysis and Cuckoo Search Optimization with SVM. In *International Journal of Computer Science and Information Security*, IJCSIS.

[7]. Liu, X., Zou, Y., Kuang, H., & Ma, X. (2020). Face image age estimation based on data augmentation and lightweight convolutional neural network. *Symmetry*, *12*(1), 146. Doi:10.3390/sym12010146

[8]. Guo, G., & Zhang, N. (2019). A Survey on Deep Learning Based Face Recognition. *Computer Vision and Image Understanding, 189*, 102805.

[9]. Hsu, C. Y., Lin, L. E., & Lin, C. H. (2021). Age and Gender Recognition with Random Occluded Data Augmentation on Facial Images. *Multimedia Tools and Applications, 80*(8), 11631–11653.

[10]. Jain, A. K., Ross, A., & Prabhakar, S. (2004). An introduction to biometric recognition. *IEEE Transactions on circuits and systems for video technology*, *14*(1), 4-20.

[11]. Al Jibory, F. K., Mohammed, O. A., & Al Tamimi, M. S. H. (2022). Age Estimation Utilizing Deep Learning Convolutional Neural Network. *International Journal on Technical and Physical Problems of Engineering, 14*(4), 219–24.

[12]. Lal, M., Kumar, K., Arain, R. H., Maitlo, A., Ruk, S. A., & Shaikh, H. (2018). Study of face recognition techniques: a survey. *International Journal of Advanced Computer Science and Applications*, *9*(6), 42–49.

[13]. Lee, S. H., Hosseini, S., Kwon, H. J., Moon, J., Koo, H. I., & Cho, N. I. (2018). Age and gender estimation using deep residual learning network. In *2018 International workshop on advanced image technology (IWAIT)* (1-3). IEEE.

[14]. Marom, N. D., Rokach, L., & Shmilovici, A. (2010).Using the Confusion Matrix for Improving Ensemble Classifiers. *2010 IEEE 26th Convention of Electrical and Electronics Engineers in Israel, IEEEI 2010,* 555–559.

[15]. Qawaqneh, Z., Mallouh, A. A., & Barkana, B. D. (2017). Deep convolutional neural network for age estimation based on VGG-face model. *arXiv preprint arXiv:1709.01664*.

[16]. Ramachandra, R., & Busch, C. (2017). Presentation attack detection methods for face recognition systems: A comprehensive survey. *ACM Computing Surveys (CSUR)*, *50*(1), 1-37.

[17]. Abdolrashidi, A., Minaei, M., Azimi, E., & Minaee, S. (2020). Age and gender prediction from face images using attentional convolutional network. *arXiv preprint arXiv:2010.03791.*

[18]. Ranjan, R., Sankaranarayanan, S., Castillo, C. D., & Chellappa, R. (2017, May). An all-in-one convolutional neural network for face analysis. In *2017 12th IEEE international conference on automatic face & gesture recognition (FG 2017)* (17-24). IEEE.

[19]. Ricanek, K., Bhardwaj, S., & Sodomsky, M. (2015). A Review of Face Recognition against Longitudinal Child Faces. *BioSiG 2015*, 15–26.

[20]. Sendik, O., & Keller, Y. (2019). DeepAge: deep learning of face-based age estimation. *Signal Processing: Image Communication*, *78*, 368-375.

[21]. Singh, S., & Prasad,.S. V.A.V. (2018). Techniques and Challenges of Face Recognition: A Critical Review. *Procedia Computer Science, 143*, 536–43.

[22]. Solanki, K., & Pittalia, P. (2016). Review of face recognition techniques. *International Journal of Computer Applications*, *133*(12), 20-24.

[23]. Wang, H., Wei, X., Sanchez, V., & Li, C. T. (2018, October). Fusion network for face-based age estimation. In *2018 25th IEEE International Conference on Image Processing (ICIP)* (2675-2679). IEEE.

[24]. Wang, W., Shen, J., & Ling, H. (2018). A deep network solution for attention and aesthetics aware photo cropping. *IEEE transactions on pattern analysis and machine intelligence, 41(7), 1531-1544.*

[25]. Zhang, K., Gao, C., Guo, L., Sun, M., Yuan, X., Han, T. X., ... & Li, B. (2017). Age group and gender estimation in the wild with deep RoR architecture. *IEEE Access, 5*, 22492-22503.