Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X Volume 3, Issue 6, 2025

REAL-TIME AGE AND GENDER ESTIMATION USING A FINE-TUNED DEEP LEARNING MODEL AND OPENCV

Asma Khaliq[*1], Abdul Basit[2], Azam Khan[3], Liaquat Ali[4], M. Saeed H. Kakar[5], Raja Asif Wagan[6]

[1,2,3,4,5]*Computer Science and Information Technology, University of Baluchistan Quetta, Pakistan.*
[6]*Department of IT FICT Baluchistan University of Information Technology, Electrical and Management Sciences.*

[*1]asmakhaliq1122@gmail.com, [2]drabasit@um.uob.edu.pk, [6]raja.asif@buitms.edu.pk

DOI: https://doi.org/10.5281/zenodo.15681297

Keywords
Age estimation; Gender recognition; Convolutional Neural Networks (CNN); Face detection; Feature extraction; Image processing; Transfer learning

Article History
Received on 05 May 2025
Accepted on 05 June 2025
Published on 13 June 2025

Corresponding Author: *
Asma Khaliq

Abstract
In today's digital era, automatic age and gender classification plays a vital role in various applications, particularly with the growing use of social media platforms. Despite recent advancements in facial recognition algorithms, analyzing real-world photographs continues to pose significant challenges. This study leverages Convolutional Neural Networks (CNNs) in conjunction with the Caffe deep learning framework and OpenCV to evaluate the accuracy of age and gender detection. By applying the Haar Cascade technique for initial face detection, the proposed model demonstrates improved performance in recognizing multiple faces within an image and accurately estimating their age and gender. The model was trained using both positive and negative facial image datasets, and its performance was thoroughly evaluated.

INTRODUCTION

Gender and age are core elements of human identity, playing important roles in areas such as marketing, social media personalization, healthcare, and security. The ability to estimate these attributes from facial images has become a key area of interest in computer vision research, driven by the growing need for automated systems capable of making accurate, real-time predictions. Traditionally, facial analysis relied on manual assessment or simple machine learning techniques, but these approaches often struggled to handle variations in lighting, pose, and facial expressions.

Convolutional neural networks (CNNs) and deep learning have proven to be powerful tools for identifying and learning patterns in visual data. Their ability to extract hierarchical features directly from raw images has made them effective for tasks such as age and gender estimation. Nonetheless, challenges like overfitting, reduced accuracy in uncontrolled environments, and computational inefficiencies remain, highlighting the need for continued innovation in model architecture and design.

In this paper, we propose an improved CNN-based model for age and gender estimation, designed to enhance accuracy through an optimized network architecture and advanced data augmentation techniques. The model aims to deliver consistent performance across multiple datasets by leveraging real-time face detection and feature extraction, implemented using the OpenCV library and the

Caffe deep learning framework [1]. Sample images from the test dataset are presented in Fig. 1. The proposed approach supports predictions from both static images and live video streams, making it well-suited for real-world applications across various domains.

The remainder of the paper is organized as follows, an overview of related work is provided in Section II, the proposed

Fig. 1. Sample images from the test set used in our model evaluation.

Methodology is described in Section III, the experimental results and analysis are presented in Section IV, and the conclusion along with future research directions is discussed in Section V.

I. RELATED WORK

The ability to estimate age and gender from face image datasets has gained significant popularity in recent years, driven by advancements in deep learning and computer vision. This section reviews the primary models and approaches employed in age and gender estimation, focusing on the evolution of these techniques from traditional machine learning methods to the adoption of Convolutional Neural Networks (CNNs).

Early methods for detecting age and gender were built on handcrafted features and simple learning models. One such approach, proposed by Rizwan et al. [2], utilized Active Appearance Models (AAMs) to simulate the effects of aging on faces by estimating age through shape and texture information. Similarly, techniques for feature extraction and classification, such as Gabor filters [3], Local Binary Patterns (LBP) [4], [3], and Support Vector Machines (SVMs) [5], were commonly used. However, these methods struggled with variations in pose, lighting, and facial expressions.

As machine learning techniques advanced, classifiers such as K-Nearest Neighbors (KNN)[6] and Decision Trees[7] were employed for age prediction and gender classification. However, these conventional methods were limited in accuracy and scalability, particularly when applied to highly variable real-world datasets.

A. The Rise of CNN-based Models

Deep learning has revolutionized image processing tasks, with Convolutional Neural Networks (CNNs) emerging as the standard for feature extraction and classification. CNNs excel in facial analysis due to their ability to automatically extract hierarchical features directly from raw image datasets. To set a new benchmark for accuracy in this domain, Levi and Hassner [8] introduced a CNN-based model for age and gender classification using the IMDB-WIKI dataset. Their research demonstrated that, in unconstrained environments, deep learning models significantly outperform traditional techniques.

Similarly, Benkaddour [9] demonstrated that CNNs trained on large, diverse datasets can achieve state-of-the-art results in tasks such as age estimation and gender recognition. Their work in- corporated data augmentation techniques to improve the model's robustness to variations in pose, lighting, and facial expressions. Another notable study by Sharma et al. [10] introduced a deep learning-based model that enhanced age estimation accuracy by utilizing innovative loss functions tailored for age prediction tasks and leveraging deeper network architectures.

B. Estimating Age and Gender in Real-time

Due to the growing demand for real-time applications, numerous studies have explored combining CNN models with real-time video processing. OpenCV, a widely used computer vision library, is frequently employed in these applications for face detection. Real-time age and gender estimation has been successfully im- plemented using Caffe, a deep learning framework [11], enabling systems to process video feeds and make real-time predictions about the age and gender of individuals. To highlight the practical utility of deep learning models, Uysal and Demiral [12] developed a CNN model integrated with OpenCV for real-time gender and ethnicity estimation.

C. Challenges and Improvements

Despite the successes of CNNs, gender and age prediction still faces several challenges. The limited availability of labeled datasets and the variability in facial features across different populations often lead to overfitting in models. Additionally, factors such as low-resolution images, occlusion, and facial accessories like masks and glasses can significantly impact model performance.

Recent research has focused on improving the generalization and accuracy of models through techniques such as transfer learning, data augmentation, and more advanced network ar- chitectures. Data augmentation methods [13], including rotation, random cropping, and color modification, help mitigate overfitting. To tackle the challenge of limited data availability, transfer learning [14], which involves fine-tuning pre-trained models on specific datasets, has also become increasingly popular.

II. PROPOSED METHODOLOGY

The CNN architecture was designed with multiple convolutional layers to extract features. Dropout layers were added to prevent overfitting, while pooling layers were used to reduce dimensionality. The deep learning framework effectively handled the large-scale dataset during both forward and backward passes. Early stopping was employed to prevent overfitting, and stochastic gradient descent (SGD) was used for optimization. After training, the model was integrated with OpenCV to create a real-time system capable of accurately detecting faces and estimating age and gender from live video streams (see Fig. 2).

A. Image Data Preprocessing

Before classifying images, the pre-processing techniques were applied to ensure optimal feature extraction and improve classi- fication accuracy.

1) The Haar Cascade classifier was utilized to detect and extract faces from the images. After this step, only the relevant facial features were analyzed further.
2) All images were resized to 227 × 227 pixels to comply with the input requirements of the CNN model.

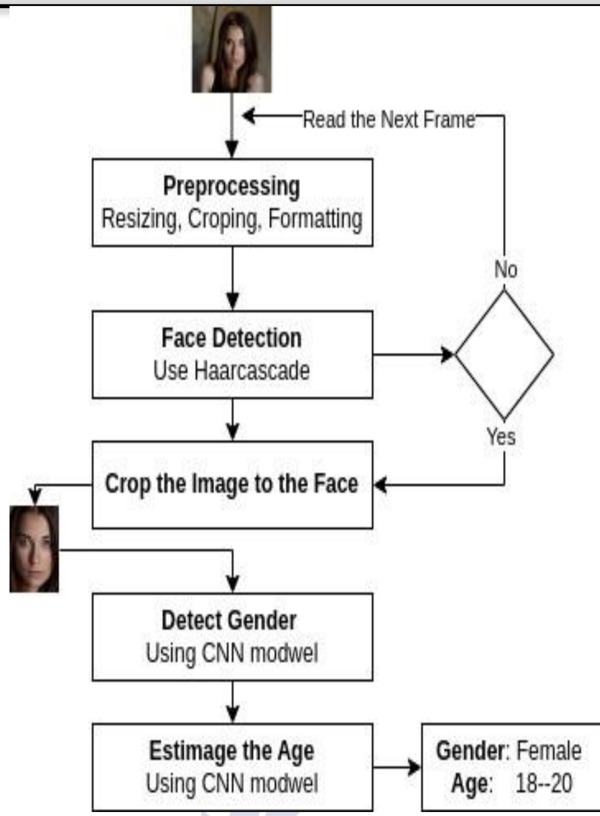

Fig. 2. The algorithm flow diagram. First, face detection followed by gender classification and age estimation.

3) Pixel values were normalized to the range [0, 1] to standardize the input data and speed up convergence during training.

4) Data augmentation techniques, including random rotation, flipping, and zooming, were applied to artificially expand the training set and introduce variations in pose, lighting, and facial expressions, helping to reduce overfitting.

B. Model Architecture

The CNN-based model is designed to extract hierarchical features from the input facial images. The architecture consists of several layers, as outlined below.

- Four convolutional layers were used with 64, 128, 256, and 512 filters, respectively, followed by a ReLU activation function in each layer. The filters helped extract spatial information from the input image, including edges, textures, and shapes.
- The feature maps were downsampled using max-pooling layers with a 2 × 2 pool size, which reduced the dimensionality while preserving important features.
- Subsequently, batch normalization was applied to normalize the output of each convolutional layer, helping to stabilize and accelerate the training process.
- To reduce overfitting, 40% of the neurons were randomly dropped during training by adding dropout layers with a dropout ratio of 0.4.
- Finally, a fully connected layer is used for classification tasks, and softmax activation is applied in the final classification layer, such as for gender. Age estimation can either follow a regression approach, predicting age directly, or use classification into discrete age groups, which would explain the use of eight neurons in your case.

The steps are also summarized and can be seen in the Fig. 3,

C. Real-Time Prediction

After training and validating the model, it was integrated into a real-time application using OpenCV and Python. Live video streams were captured from a webcam, faces were detected in each frame, and the face images were passed through the trained CNN model. The predicted ages and genders were then superimposed in real time on the detected faces.

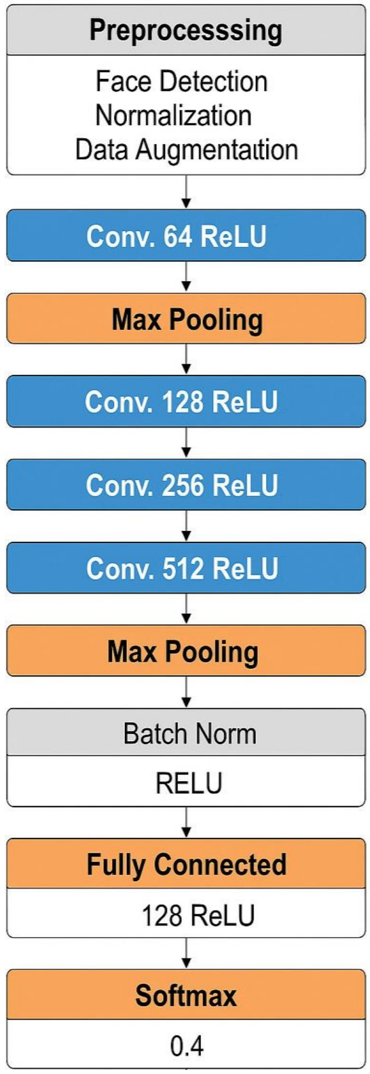

Fig. 3. Model architecture showing various layers, including convolutional layers, pooling layers, fully connected layers, and a softmax layer for classi- fication.

D. **Evaluation Metrics**

The performance of the model was assessed using the following metrics:

- Accuracy is the percentage of correct age and gender predictions out of the total predictions made.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

where TP is true positive, TN is true negative, FP is false positive, and FN is false negative.

- The precision and recall metrics were calculated for gender classification and age estimation to measure the model's ability to accurately identify male and female age and faces.

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

- A confusion matrix was used to assess the performance across different age groups.
- The F1 score, which is the harmonic mean of precision and recall, was computed to provide a balanced measure of accuracy.

$$F1Score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

III. EXPERIMENTS AND RESULTS

To evaluate the performance of the proposed CNN-based model for age and gender estimation, a series of experiments were conducted. These experiments aimed to assess the model's accuracy, robustness, and real-time applicability using both static images and live video streams. This section outlines the experimental setup, presents the results, and discusses key observations.

A. Experimental Setup

The experiments were conducted using the age, gender, and ethnicity face data, which was compiled from various online and offline sources and labeled by age and gender. The model was implemented in Python using the OpenCV library and the Caffe deep learning framework. All experiments were performed on a system with the following specifications:

- Processor: Intel Core i5-7200U @ 2.5 GHz
- GPU: NVIDIA GTX 980 Ti was used with 6GB
- Memory: 8 GB RAM
- Storage: 128 GB SSD
- Software: Python, OpenCV, Caffe, and Keras

B. Experiments

The experiments evaluated the two key aspects of the proposed model, the model's accuracy and real-time performance, to assess its robustness.

1) ***Age and Gender Classification Accuracy:***

The dataset was used to train the Caffe model for age and gender classification. The model's accuracy in predicting age and classifying gender into eight predefined age groups was evaluated.

- (0-2), (4-6), (8-12), (15-20), (25-32), (38-43), (48-53), and (60-100).

In our first experiment, we gathered a diverse image dataset of 100 individuals, encompassing both genders and a broad age range (0 to 80 years), from online sources. Later, gender classification was conducted, achieving an average accuracy of 92%. Subsequently, the age estimation on the same dataset yielded an average accuracy of 80%. The outcomes of this experiment are shown in Figure 4.

The detailed results of the experiments are as follows:

- Gender Classification Performance: The model classified genders across the test set with an accuracy of 90%. Female profiles were predicted with a slightly lower accuracy of 88%, whilst male profiles were predicted with 92% accuracy.
- Age Estimation Performance: The age estimation task proved more challenging due to overlapping facial features across adjacent age ranges. The model achieved an accuracy of 82% across all age categories, with the best performance observed in the (25-32) age group and the lowest performance in the (60-100) age group.

2) ***Real-Time Performance:***

To assess the model's perfor- mance in real-world scenarios, it was integrated into a real-time system using OpenCV to capture video frames from a webcam. The model was then applied to detect faces in the video stream and predict both age and gender in real-time. The Results are:

We have measured the time required to estimate the age and gender of a single person in an image with a resolution of approximately 800 × 600. On our CPU, the model achieves 10 to 14 frames per second (FPS), while on the specified GPU, it processes 48 to 52 FPS.

In real-time video processing, the model faced challenges in low-light environments, resulting in reduced accuracy for both age and gender predictions due to poor visibility of facial features.

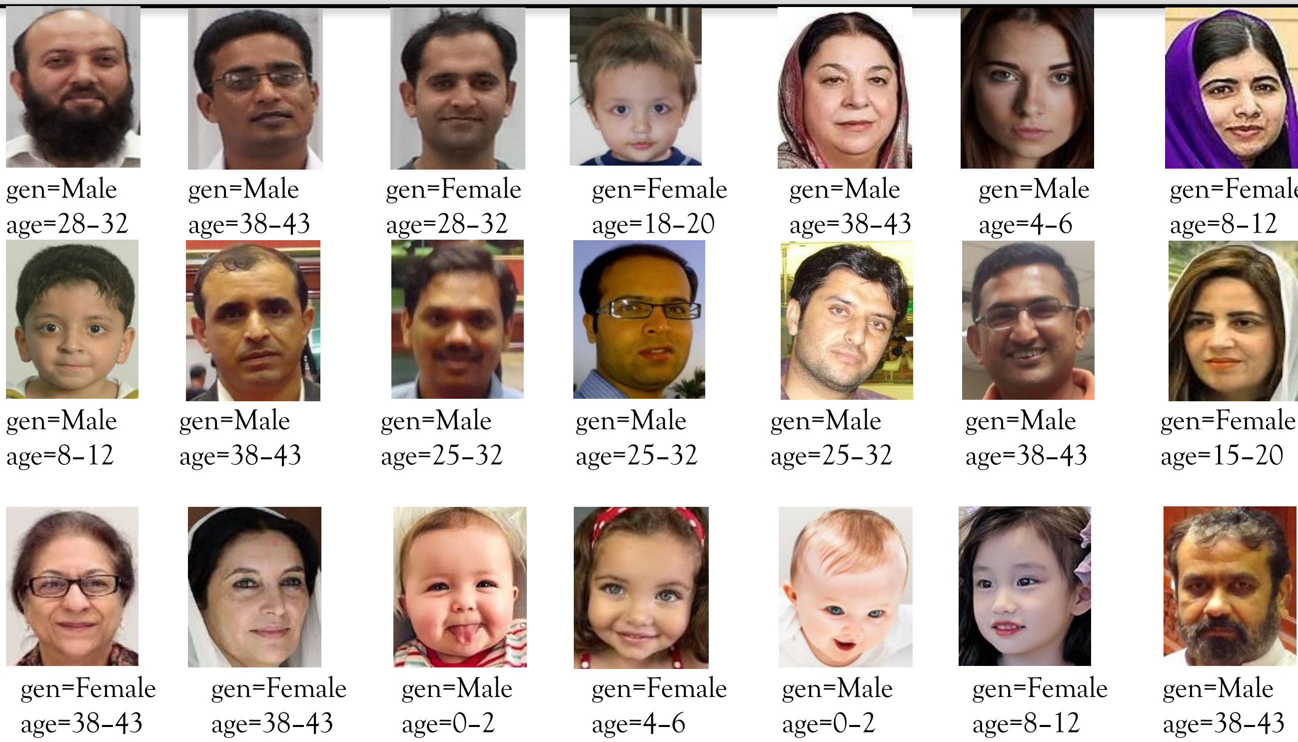

Fig. 4. Proposed model prediction of randomly collected dataset from online resources. The model perform well on the collected dataset.

IV. CONCLUSION

This paper presented an improved CNN-based model for age and gender estimation, achieving strong performance in both static and real-time applications. By integrating OpenCV for face detection and Caffe for CNN implementation, the model demonstrated a high accuracy of 92% for gender classification and reasonable results for age estimation, especially in the 25-32 age group.

Although the model performed well, challenges remain, par- ticularly in predicting age for older individuals and in low-light conditions. Future improvements could focus on addressing these limitations and further refining the model's generalizability for more diverse real-world scenarios.

ACKNOWLEDGMENTS

Asma Khaliq was supported by graduate fellowships from the University of Balochistan and the Higher Education Commission (HEC) of Pakistan.

REFERENCES

Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick,

S. Guadarrama, and T. Darrell, "Caffe: Convolutional architecture for fast feature embedding," in *Proceedings of the 22nd ACM international conference on Multimedia*, 2014, pp. 675–678.

S. A. Rizwan, A. Jalal, M. Gochoo, and K. Kim, "Robust active shape model via hierarchical feature extraction with sfs-optimized convolution neural network for invariant human age classification," *Electronics*, vol. 10, no. 4, p. 465, 2021.

E. Al-Wajih and M. A. Ahmed, "A new application for gabor filters in face-based gender classification." *Int. Arab J. Inf. Technol.*, vol. 17, no. 2, pp. 178–187, 2020.

H. K. Omer, H. A. Jalab, A. M. Hasan, and N. E. Tawfiq, "Combination of local binary pattern and face geometric features for gender classification from face images," in *2019 9th IEEE Inter- national Conference on Control System, Computing and Engineering (ICCSCE)*. IEEE, 2019, pp. 158–161.

S. Kumar, S. Singh, J. Kumar, and K. Prasad, "Age and gender clas- sification using seg-net based architecture and machine learning." *Multimedia Tools & Applications*, vol. 81, no. 29, 2022.

V. Priyadarshni, A. Nayyar, A. Solanki, and A. Anuragi, "Human age classification system using k-nn classifier," in *Advanced Informatics for Computing Research: Third International Conference, ICAICR 2019, Shimla, India, June 15–16, 2019, Revised Selected Papers, Part I 3*. Springer, 2019, pp. 294–311.

S. O¨ ner, M. Turan, and Z. O¨ ner, "Estimation of gender by using decision tree, a machine learning algorithm, with patellar measure- ments obtained from mdct images," *Medical Records*, vol. 3, no. 1, pp. 1–9, 2021.

G. Levi and T. Hassner, "Age and gender classification using convo- lutional neural networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 2015, pp. 34–42.

M. K. Benkaddour, "Cnn based features extraction for age estima- tion and gender classification," *Informatica*, vol. 45, no. 5, 2021.

N. Sharma, R. Sharma, and N. Jindal, "Face-based age and gender estimation using improved convolutional neural network approach," *Wireless Personal Communications*, vol. 124, no. 4, pp. 3035–3054, 2022.

S. Shaban, R. P. Singh, and M. Mehra, "Age and gender prediction using caffe model and opencv," *International Journal of Innovative Research in Computer Science & Technology*, vol. 10, no. 1, pp. 15–21, 2022.

M. Uysal and M. F. Demiral, "Gender and age estimation by image processing," *Dicle U¨ niversitesi Mu¨ hendislik Faku¨ ltesi Mu¨ hendislik Dergisi*, vol. 15, no. 1, pp. 49–59, 2024.

C.-Y. Hsu, L.-E. Lin, and C. H. Lin, "Age and gender recognition with random occluded data augmentation on facial images," *Multi- media Tools and Applications*, vol. 80, no. 8, pp. 11 631–11 653, 2021.

V. Sheoran, S. Joshi, and T. R. Bhayani, "Age and gender prediction using deep cnns and transfer learning," in *Computer Vision and Image Processing: 5th International Conference, CVIP 2020, Praya- graj, India, December 4-6, 2020, Revised Selected Papers, Part II 5*. Springer, 2021, pp. 293–304.