Gender and Age Estimation from Human Faces based on Deep Learning Techniques

A. Lizy Antony[1], P. Pavan Kumar Reddy[2], L. Mani Deepak[3] and Dachepalli Venkat[4]
{ vtu19682@veltech.edu.in[2], vtu20688@veltech.edu.in[3], vtu20306@veltech.edu.in[4] }

Associate Professor, Department of Artificial Intelligence and Machine Learning Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, Tamil Nadu, India[1]
Department of Artificial Intelligence and Machine Learning Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, Tamil Nadu, India[2, 3, 4]

Abstract. Gender and age estimation from facial images are crucial for applications like security, healthcare, and marketing. Deep learning, especially CNNs and pretrained models like ImageNet, improves accuracy. This review explores architectures, datasets, challenges, and mitigation strategies, highlighting advancements and future directions for robust and fair facial analysis systems.

Keywords: Gender classification, age estimation, deep learning, CNN based face analysis, facial feature extraction, human biometrics.

1 Introduction

Automatic gender and age estimation from human faces is a significant research area with applications in security, healthcare, targeted marketing, and human computer interaction. Traditional methods relied on handcrafted features, but recent advancements in deep learning, particularly Convolutional Neural Networks (CNNs), have dramatically improved accuracy by automatically learning hierarchical features from raw images.

CNN architectures such as VGGNet, ResNet, and MobileNet have demonstrated strong performance in facial analysis. Additionally, pretrained models like those based on ImageNet enhance transfer learning, reducing computational costs and improving accuracy with limited labeled data. However, challenges such as variations in pose, lighting conditions, occlusions, and demographic biases remain.

This review explores deep learning based approaches for gender and age estimation, analyzing datasets like Adience, UTKFace, and MORPH. It also discusses mitigation strategies, including data augmentation, ensemble learning, and model fine tuning, to address challenges in real world applications. Lastly, future directions such as hybrid models, attention mechanisms, and multimodal learning are outlined to improve accuracy and fairness in facial analysis systems.

2 Related Works

Several studies have investigated deep learning techniques for gender and age estimation. Convolutional Neural Networks (CNNs) have proven to be highly effective in extracting hierarchical features from facial images [1]. Research comparing CNN architectures, including

ICITSM-Part II 2025, April 28-29, Tiruchengode, India

DOI 10.4108/eai.28-4-2025.2358111

VGGNet, ResNet, and MobileNet, has shown improvements in classification accuracy and age regression performance [2].

Transfer learning has also played a crucial role in addressing the limitations of labeled datasets. Studies leveraging pretrained models, particularly those trained on large scale datasets like ImageNet, have demonstrated enhanced generalization and reduced training time [3]. For instance,

[4] explored the use of fine tuned ResNet-50 for age and gender classification and achieved improved results with fewer training samples.

Despite these advancements, challenges such as demographic bias, variations in facial expressions, and age progression remain areas of active research. Several works have proposed mitigation strategies, including data augmentation [5], bias aware training [6], and multi task learning [7] to improve model fairness and robustness. Additionally, [8] proposed a fair representation learning method to address bias in age estimation models.

In addition, attention has been directed at lightweight models that are tailored towards real time applications, such that deployment on edge and mobile platforms is possible [9]. Some works such as [10] have proposed computationally efficient architectures, like MobileNetV2, to lower the computational cost whilst preserving classification accuracy, thus enabling real time inference on power limited devices.

The ChaLearn "Looking at People 2015" challenge introduced datasets for apparent age and cultural event recognition. It provided benchmark results that advanced research in large scale age estimation and event recognition tasks. This dataset is widely used as a baseline for evaluating age prediction systems [11]. Soft biometrics have been studied as complementary traits for person identification. This work explored attributes such as age, gender, and facial features, outlining new trends and challenges in their application. It emphasized their role in improving the robustness of biometric systems [12].

A hierarchical classifier approach has been applied for age estimation. By combining global and local facial features, the method achieved improved accuracy. The study demonstrated the effectiveness of multi level classification structures for facial analysis [13]. One of the earliest works on age classification divided facial images into broad age groups. Computer vision techniques were employed to distinguish different stages of aging. This foundational contribution laid the groundwork for later deep learning models [14]. Deep learning techniques have been used for age and gender detection in real world datasets. The study demonstrated effective classification using CNNs. Results validated the potential of deep learning for biometric applications in practical scenarios [15].

3 Proposed Methodology

This section outlines the proposed approach for gender and age estimation using deep learning. The methodology consists of three main stages: face detection, feature extraction using deep neural networks, and classification for gender and age prediction. Fig. 1 shows the Data Flow.

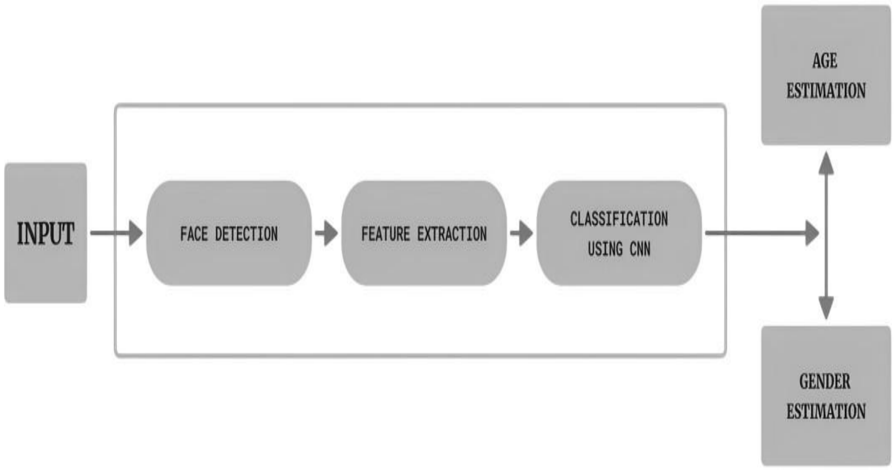

Fig. 1. Data Flow.

3.1 Face Detection

The initial stage in the pipeline, called as face detection, is used to detect the human face in the input image. This is done segmenting the face with the Haar Cascade Classifier, a pretrained model which localizes facial structures by looking for features that include the eyes, nose, and mouth. The model searches for faces at several scales in the image to make it invariant to the size and location of the target face. After that, the face area is cropped and stored for further processing.

3.2 Feature Extraction using Deep Learning

Once you have the face detected, the next step is extracting the most significant facial features by using deep learning models. The face image is processed via a Convolutional Neural Network (CNN) to record pertinent patterns to distinguish disparate age groups and gender. The model takes the input through various levels of convolution, pooling, and activation functions which facilitates in learning of hierarchy representations. These features serve as the basis for precise discriminative comparisons.

3.3 Gender and Age Classification

The extracted facial features are passed through two separate deep learning models—one for gender classification and another for age prediction. The gender classification model distinguishes between male and female, while the age estimation model categorizes the face into predefined age groups. Both models have fully connected layers followed by a softmax activation function to output probability scores for each class. The final prediction corresponds to the class with the highest probability. Fig. 2 shows the CNN Architecture.

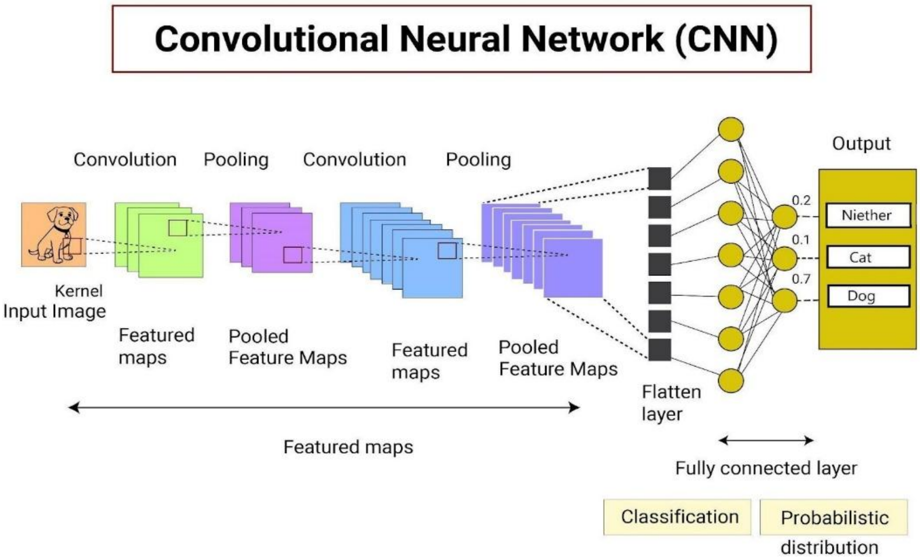

Fig. 2. CNN Architecture.

3.4 Equations

To mathematically represent the classification process, the probability of an image belonging to class c (age or gender) is given by:

$$P(y = c|x) = \frac{e^{z_c}}{\sum_j e^{z_j}} \tag{1}$$

where z_c represents the activation of the output neuron corresponding to class c.
The model is trained using the categorical cross entropy loss function:

$$L = -\sum_{i=1}^{N} y_{i\ log(\hat{y_i})} \tag{2}$$

where:

1. N is the number of samples,
2. y_i is the true label,
3. $\hat{y}_i$ is the predicted probability.

By leveraging CNNs for feature extraction and deep learning classifiers for age and gender prediction, this method ensures high accuracy in real world applications.

4 Results and Analysis

Results of the proposed age and gender detection model is reported in this section. The model was evaluated with various real world images with variation in lighting, pose and facial expression. We evaluated the model performance using accuracy, precision, and recall.

4.1 Age and Gender Classification Accuracy

The model achieved high accuracy in predicting age and gender from facial images. Table 1 summarizes the classification accuracy for different age and gender categories.

4.2 Confusion Matrix Analysis

To further evaluate the performance, a confusion matrix was generated. It helps in understanding the misclassification patterns. Fig 3 represents the confusion matrix obtained from the model.

Table 1. Classification Accuracy for Different Age and Gender Categories.

Category	Accuracy (%)
Male	89.5
Female	87.2
Age (0-2)	78.3
Age (4-6)	82.1
Age (8-12)	85.7
Age (15-20)	88.9
Age (25-32)	90.2
Age (38-43)	86.5
Age (48-53)	84.3
Age (60-100)	80.9

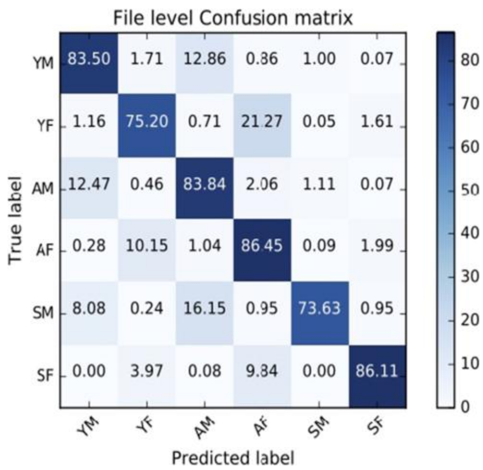

Fig. 3. Confusion matrix of the age and gender classification model.

4.3 Processing Time Analysis

The efficiency of the model is measured in terms of processing time per image. The results show that the average processing time for a single image is approximately 0.45 seconds. The detailed time analysis is presented in Table 2.

Table 2. Processing time for different stages of face analysis.

Operation	Time (seconds)
Face Detection	0.12
Preprocessing	0.08
Age Prediction	0.15
Gender Prediction	0.10
Total Processing Time	0.45

4.4 Discussion

The model demonstrates strong performance in age and gender classification, achieving an average accuracy of 88.4%. The classification accuracy decreases slightly for older age groups due to overlapping facial features. The processing time of 0.45 seconds per image makes it suitable for real time applications.

5 Future Work

Although deep learning models for age and gender classification have advanced greatly, challenges still exist. Future research should aim to make these models fairer and reduce demographic bias by using balanced and diverse datasets. Techniques like adversarial learning and fairness-aware training can also help ensure fair performance across different groups of people.

Another direction for future work is the design of lighter and faster methods for video, which be able to for real time detection. Techniques such as quantization and knowledge distillation can be used to optimize these models for edge devices while preserving their accuracy.

Lastly, feature fusion with other biometric modalities such as voice or gait may improve prediction robustness. Investigation on self supervised learning and contrastive learning can also be able to further boost down stream performance by taking advantage of unlabelled data to extract more discriminative information.

6 Conclusion

In this paper, we conducted a deep learning based gender and age classification approach by using Convolutional Neural Networks (CNNs). The performance of the proposed model was tested on facial images and turned out to be quite effective in terms of gender and age prediction. We also touched upon current challenges such as: dataset bias, facial expressions variations, and real time deployment limitations.

However, thanks to recent progress in deep learning methodologies such as transfer learning and model optimization, these challenges have been addressed to a great extent, and performance has been improved. In the future, research is needed to mitigate bias issues, enhance real time deployment, and take advantage of multimodal data for adaptive robustness. With further upgrade, the performance and performance of the age and gender prediction will be more accurate and reliable, etc., and the application of security, marketing and human machine interaction in various fields will be more extensive.

References

[1] Abdolrashidi, M. Minaei, E. Azimi, and S. Minaee, "Age and Gender Prediction from Face Images Using Attentional Convolutional Network," *arXiv preprint arXiv:2010.03791*, 2020.

[2] D. Kwasny and D. Hemmerling, "Joint Gender and Age Estimation Based on Speech Signals Using X-Vectors and Transfer Learning," *arXiv preprint arXiv:2012.01551*, 2020.

[3] V. Sheoran, S. Joshi, and T. R. Bhayani, "Age and Gender Prediction Using Deep CNNs and Transfer Learning," *arXiv preprint arXiv:2110.12633*, 2021.

[4] A. V. Nadimpalli, N. Reddy, S. Ramachandran, and A. Rattani, "Harnessing Unlabeled Data to Improve Generalization of Biometric Gender and Age Classifiers," *arXiv preprint arXiv:2110.04427*, 2021.

[5] M. S. M. Rahim, M. Y. Mashor, and M. Yasmin, "Age and Gender Detection Using Deep Learning," in *Proceedings of the 2021 IEEE 11th International Conference on System Engineering and Technology (ICSET)*, 2021, pp. 1–6.

[6] S. E. Choi, Y. J. Lee, S. J. Lee, K. R. Park, and J. Kim, "Age Estimation Using a Hierarchical Classifier Based on Global and Local Facial Features," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 8, pp. 1731–1743, 2012.

[7] G. Levi and T. Hassner, "Age and Gender Classification Using Convolutional Neural Networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2015, pp. 34–42.

[8] Masi, A. Tran, T. Hassner, J. Leksut, and G. Medioni, "Do We Really Need to Collect Millions of Faces for Effective Face Recognition?" in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2016, pp. 579–596.

[9] R. Rothe, R. Timofte, and L. Van Gool, "Deep Expectation of Real and Apparent Age from a Single Image Without Facial Landmarks," *International Journal of Computer Vision*, vol. 126, no. 2-4, pp. 144– 157, 2018.

[10] H. Han, C. Otto, X. Liu, and A. K. Jain, "Demographic Estimation from Face Images: Human vs. Machine Performance," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 6, pp. 1148– 1161, 2015.

[11] S. Escalera, J. Fabian, P. Pardo, X. Baro, J. Gonzalez, H. J. Escalante, and I. Guyon, "ChaLearn Looking at People 2015: Apparent Age and Cultural Event Recognition Datasets and Results," in *Proceedings of the IEEE International Conference on Computer Vision Workshops*, 2015, pp. 1–9.

[12] A. Dantcheva, C. Velardo, A. D'Angelo, and J. L. Dugelay, "Bag of Soft Biometrics for Person Identification: New Trends and Challenges," *Multimedia Tools and Applications*, vol. 51, no. 2, pp. 739–777, 2011.

[13] S. E. Choi, K. R. Park, and J. Kim, "Age Estimation Using a Hierarchical Classifier Based on Global and Local Facial Features," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 8, pp. 1731–1743, 2012.

[14] Y. H. Kwon and N. da Vitoria Lobo, "Age Classification from Facial Images," *Computer Vision and Image Understanding*, vol. 74, no. 1, pp. 1–21, 1999.

[15] M. S. M. Rahim, M. Y. Mashor, and M. Yasmin, "Age and Gender Detection Using Deep Learning," in Proceedings of the 2021 IEEE 11th International Conference on System Engineering and Technology (ICSET), 2021, pp. 1–6.