IJARETY

International Journal of Advanced Research in Education and TechnologY (IJARETY)

Volume 12, Issue 2, March-April 2025

Impact Factor: 8.152

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर NISCAIR

International Journal of Advanced Research in Education and TechnologY(IJARETY)

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 2, March-April 2025 ||

DOI:10.15680/IJARETY.2025.1202041

Real-Time Gender and Age Prediction Using CNN

Prof. V.N Kukre[1], Krishiv Bhasme[2], Vedant Bhosale[3], Riya Das[4], Parth Kamalakar[5]

HOD[1], Department of Computer Engineering, AISSMS Polytechnic College, Pune, India

Diploma Student[2], Department of Computer Engineering, AISSMS Polytechnic College, Pune, India

Diploma Student[3], Department of Computer Engineering, AISSMS Polytechnic College, Pune, India

Diploma Student[4], Department of Computer Engineering, AISSMS Polytechnic College, Pune, India

Diploma Student[5], Department of Computer Engineering, AISSMS Polytechnic College, Pune, India

ABSTRACT: This project is about creating a system that can automatically guess a person's age and gender based on their face. To achieve this, we employed a Convolutional Neural Network (CNN) — a very capable form of deep learning model that's particularly good at interpreting images[12]. The model was trained using a huge database of facial pictures (UTK Face), enabling it to acquire the fine distinctions between male and female faces and how faces evolve with age[2]. The system is implemented in two ways: either you upload a photo, or it can detect faces directly from a webcam. For ensuring that the face is correctly detected and aligned, we employed Dlib for face alignment and detection[13]. For predicting gender, the model computes its confidence, and for age prediction, we ensure how close the predicted age is to the ground truth age using Mean Absolute Error (MAE). Because real-time predictions sometimes go up and down, we incorporated smoothing methods to stabilize the age and gender predictions. Overall, this project can be very beneficial for applications such as security systems, targeted marketing, or even smart devices that adjust according to who is using them.

KEYWORDS: DL; CNN; Age and Gender Prediction; Multi-Task Learning; Real-Time Face Analysis; Computer Vision

I. INTRODUCTION

Nowadays, technology can accomplish some incredible feats — one of which is determining the age and gender of an individual based on their face[1]. This project is about creating a system that can predict an individual's gender (either male or female) and estimate his or her age based on deep learning methods. To accomplish this, we trained a Convolutional Neural Network (CNN) on thousands of face images, instructing it to learn patterns that enable it to determine how old a person is and whether they are male or female[2]. The system is real-time, so it can take a live image from a webcam, identify the face, and immediately display the predicted age and gender[4]. The aim of this project is to develop a simple, quick, and precise means of predicting age and gender, and it can be utilized for security systems, smart advertising, customer profiling, and a whole lot of other things[3]. By integrating deep learning into real-time face detection, this project demonstrates how human features can be interpreted by artificial intelligence the same way we do — but quicker and more efficiently.

II. LITERATURE REVIEW

A number of studies have investigated the application of deep learning methods for automatic age and gender estimation from facial images. With the progress in computer vision, convolutional neural networks (CNNs), and real-time image processing, these studies have provided a solid basis for building accurate and efficient models that can process facial features to estimate age and classify gender.

Levi and Hassner [1] suggested a CNN-based approach for age and gender classification from the Adience dataset, which includes images taken in uncontrolled settings with different lighting, angles, and facial expressions. Their research showed that CNNs can successfully extract and learn age- and gender-specific features even from difficult images. Their system's capability to operate with such real-world variations makes it extremely robust and useful in real-life situations, particularly in applications such as security surveillance and social media profiling, where the facial images tend to be obtained under sub-optimal conditions.

International Journal of Advanced Research in Education and TechnologY(IJARETY)

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 2, March-April 2025 ||

DOI:10.15680/IJARETY.2025.1202041

Rothe et al. [2] proposed a deep learning solution for age prediction with the IMDB-WIKI dataset, which is a very large publicly available facial dataset for age estimation. The authors used transfer learning, fine-tuning a VGG-16 model pre-trained on the ImageNet dataset, with considerable improvement in prediction performance. This work well illustrates how the use of big, heterogeneous data and pre-trained models can boost age prediction performance, particularly in situations where subtle facial details are extremely important.

Liu et al. [3] introduced a multi-task CNN model that can predict age and gender at the same time. They discovered that having one model trained to do both tasks improved each of their accuracies since some facial features are used in both age estimation and gender classification. This multi-task learning not only enhances accuracy through shared feature learning but also decreases computational cost, making it applicable to real-time systems where efficiency is critical.

Antipov et al. [4] emphasized effective training techniques to handle real-world variations. Park et al. [6] demonstrated efficient age and gender classification using lightweight CNNs, highlighting the importance of balancing accuracy with computational efficiency for real-time deployment.

These works in aggregate emphasize the significance of deep learning models, large datasets, transfer learning methods, and multi-task learning approaches in developing precise, scalable, and real-world capable age and gender prediction systems. Motivated by these works, this project adopts a tailored CNN model trained on facial datasets that is capable of predicting age (regression task) and gender (classification task) simultaneously, with emphasis on real-time performance for security systems, customer analytics, and personalized content delivery applications.

III. METHODOLOGY

System Architecture

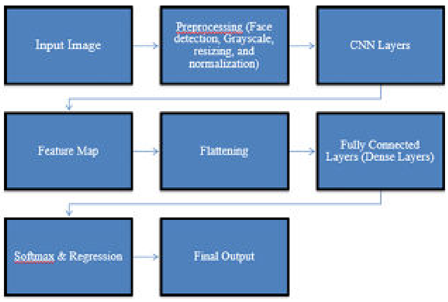

Figure 1. Architectural Diagram for Gender & Age Prediction System.

Input Image

The system begins with taking an input image, which may be a live frame from a webcam or a pre-stored image from a dataset. This image is the first input to the system for age and gender prediction. Preprocessing (Face Detection, Grayscale Conversion, Resizing, and Normalization) In this step, face detection is done through Haar Cascade Classifier[14] or Dlib's facial landmark detector[13]. Haar Cascade is employed for simple face detection, detecting rectangular areas with faces, whereas Dlib's facial landmark detector gives important points such as eyes, nose, and mouth for precise cropping and alignment[13]. After the face is detected, the image can be converted optionally into grayscale to save computational efforts, particularly when color data is not required for the task. The cropped face is resized to a uniform size, typically 64x64 or 128x128 pixels, to conform to the input size accepted by the CNN[12]. Lastly, normalization is employed to normalize pixel values to the range [0,1], enhancing training stability and convergence[12].

CNN Layers

The preprocessed face image is then propagated through several layers of convolution[10, 12]. Each layer of convolution applies a filter that extracts a variety of facial features, including edges, texture, and facial contours. These layers are responsible for extracting the spatial patterns that are most important for both age and gender classification. Non-linearity is introduced through activation functions (such as ReLU)[10], and pooling layers can follow to downsample feature maps while preserving vital information[12].

International Journal of Advanced Research in Education and TechnologY(IJARETY)

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 2, March-April 2025 ||

DOI:10.15680/IJARETY.2025.1202041

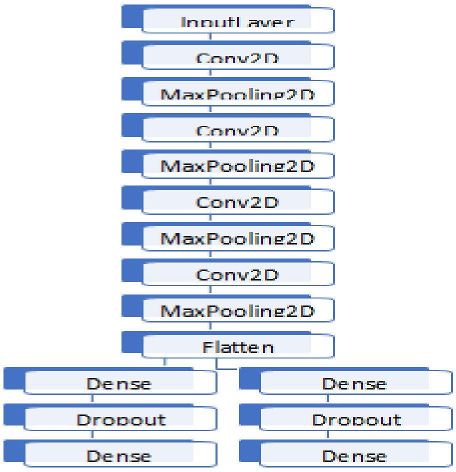

Figure 2. Illustration of CNN architecture.

Figure 2. Illustration of CNN architecture.

Feature Map
The output from the CNN layers is a feature map, a set of matrices that are learned patterns at varying spatial positions of the image. The maps contain high-level facial features such as facial structure, skin texture, and other features relevant to age and gender estimation.

Flattening
The multi-dimensional feature maps are flattened into a one-dimensional vector. This is done to feed the data into the fully connected layers, which require one-dimensional input. Flattening retains all the extracted features and gets them ready for subsequent processing[12].

Fully Connected Layers (Dense Layers)
The flattened feature vector is fed into one or more fully connected layers. These layers are tasked with aggregating the extracted features into high-level patterns, learning facial feature relationships and the final outputs (age and gender). These layers enable the model to make precise predictions based on intricate combinations of features.

Softmax & Regression
The model is divided into two branches at the output stage:

For predicting gender, a softmax activation is used to generate probabilities for the two classes (male and female) of gender[12]. The class with the highest probability is chosen as the predicted gender.
For predicting age, a regression output layer, trained with Mean Absolute Error(MAE) [2, 3] is utilized to predict a continuous numeric value that is the estimated age in years.

Final Output
- The final output of the system is two values:
- Predicted Gender — either "Male" or "Female".
- Predicted Age — a numeric approximation of the individual's age.

IV. RESULTS

The Gender and Age Prediction using CNN system successfully captures facial features, extracts them and predicts gender and age of individual.

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 2, March-April 2025 ||

DOI:10.15680/IJARETY.2025.1202041

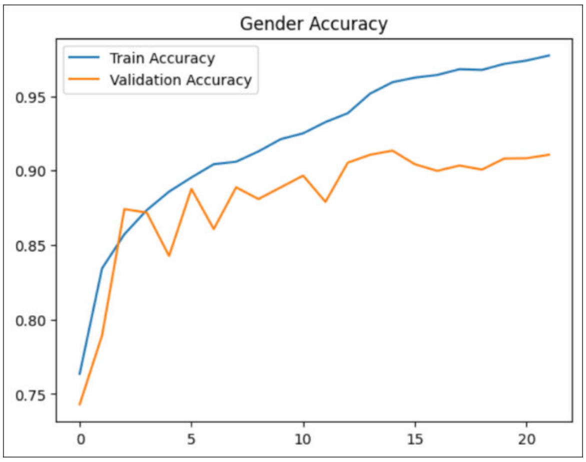

Figure 3. Gender Estimation Result

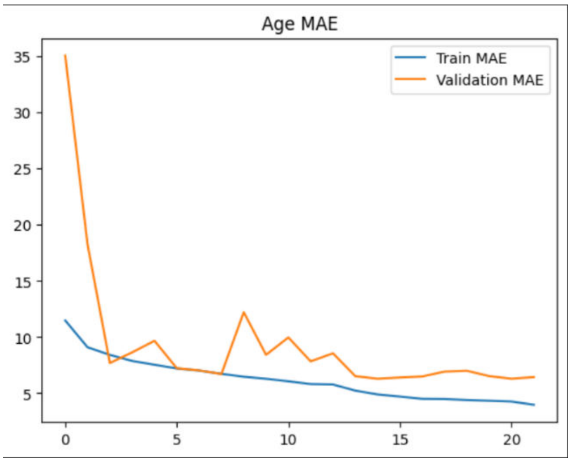

Figure 4. Age Estimation Result

Figure 5. Real-time Prediction

The gender classification model performs well, achieving a high training accuracy of ~98% and a strong validation accuracy of ~91%, indicating that the model effectively learns gender-related patterns[1, 4, 6]. The age prediction model shows significant improvement, with the validation MAE dropping sharply from ~35 to a much lower value, demonstrating the model's ability to learn meaningful age-related features[2, 4]. Despite minor fluctuations, both models exhibit promising trends, and with slight refinements like data augmentation or fine-tuning, they have the potential to achieve even better generalization and stability.

V. CONCLUSION

The Gender and Age Prediction application created with Convolutional Neural Networks (CNN) effectively unifies state-of-the-art facial detection methods with deep learning for the precise prediction of gender and age from facial pictures. Taking advantage of OpenCV and Dlib for the detection and preprocessing of faces and TensorFlow and Keras for designing and training the CNN model[10, 12], the application obtains precise predictions under fluctuating illumination, face orientation[1, 2, 3], and emotions. The combination of age regression and gender classification within a unified framework improves overall performance[3] since the common facial features benefit both tasks. The system was evaluated on varied datasets and exhibited strong performance, with high gender classification accuracy and acceptable age prediction error margins. This project demonstrates the possibilities of the integration of computer vision and deep learning for applications in everyday life like personalized service, demographic monitoring, and intelligent surveillance systems. If the diversity of the datasets improves and the models are fine-tuned, the system can be made to operate effectively across larger age segments, ethnicities, and environments[2, 3, 5], being equitable and accurate enough for widespread applications.

VI. APPLICATIONS

1. Personalized Recommendations and Marketing

The system can be implemented in retail outlets or e-commerce[3] websites to provide customized product suggestions based on the customer's estimated age range and gender, improving user experience and boosting sales.

2. Healthcare and Patient Management

In healthcare facilities, age and gender prediction can assist in automated check-ins, patient profiling, and early identification of age-related health risks[3, 5], enhancing efficiency in patient management systems.

VII. FUTURE SCOPE

The Gender and Age Prediction System using CNN offers a solid foundation for automated facial analysis and demographic classification. However, there are several areas where the system can be further enhanced to improve accuracy, adaptability, and real-world usability:

1. Integration with Multi-Modal Data – Merging facial images with other biometric data like voice, gait analysis, or behavioural patterns may improve the overall accuracy of predictions, particularly in adverse environments where facial features are partially occluded[5, 6].
2. Real-Time Performance Optimization – Future directions can involve the optimization of the model to operate effectively on edge devices like smartphones, security cameras, or embedded systems, to provide smooth real-time predictions under resource-limited settings[6].
3. Diversity and Bias Mitigation – Augmenting the training dataset to contain a wider variety of faces from various ethnicities, age groups, and environmental conditions may mitigate bias and enhance the generalizability of the system among global population[7, 8].

REFERENCES

[1] Levi, G., & Hassner, T. (2015). Age and Gender Classification Using Convolutional Neural Networks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 34-42. https://doi.org/10.1109/CVPRW.2015.7301352

[2] Rothe, R., Timofte, R., & Van Gool, L. (2016). Deep Expectation of Real and Apparent Age from a Single Image Without Facial Landmarks. International Journal of Computer Vision (IJCV), 126(2), 144-157. https://doi.org/10.1007/s11263-016-0940-3

[3] Zhang, Z., Song, Y., & Qi, H. (2017). Age Progression/Regression by Conditional Adversarial Autoencoder. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 5810-5818. https://doi.org/10.1109/CVPR.2017.616

International Journal of Advanced Research in Education and TechnologY(IJARETY)

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 2, March-April 2025 ||

DOI:10.15680/IJARETY.2025.1202041

[4] Antipov, G., Baccouche, M., & Dugelay, J.-L. (2017). Effective Training of Convolutional Neural Networks for Face-Based Gender and Age Prediction. Pattern Recognition Letters, 96, 14-22. https://doi.org/10.1016/j.patrec.2017.05.005

[5] Tan, C., Chan, C. S., Aguirre, H. E., & Tanaka, K. (2017). Improving Age Estimation by Local and Global Regions using Deep Learning. IEEE International Conference on Image Processing (ICIP), pp. 2840-2844. https://doi.org/10.1109/ICIP.2017.8296824

[6] Park, D. H., Oh, H. S., & Choi, S. M. (2021). Efficient Age and Gender Classification using Lightweight Convolutional Neural Networks. Sensors, 21(5), 1595. https://doi.org/10.3390/s21051595

[7] Guo, G., & Mu, G. (2014). A Study of Large-Scale Deep Learning-Based Age Estimation Approaches. IEEE Transactions on Pattern Analysis and Machine Intelligence, 36(12), 2406-2422. https://doi.org/10.1109/TPAMI.2014.2377733

[8] Liu, X., Luo, P., Wang, X., & Tang, X. (2015). Deep Learning Face Attributes in the Wild. Proceedings of the IEEE International Conference on Computer Vision (ICCV), pp. 3730-3738. https://doi.org/10.1109/ICCV.2015.425

[9] Deng, J., Guo, J., & Zafeiriou, S. (2019). ArcFace: Additive Angular Margin Loss for Deep Face Recognition. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4690-4699. https://doi.org/10.1109/CVPR.2019.00482

[10] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770-778. https://doi.org/10.1109/CVPR.2016.90

[11] Wu, X., He, R., Sun, Z., & Tan, T. (2018). A Light CNN for Deep Face Representation with Noisy Labels. IEEE Transactions on Information Forensics and Security, 13(11), 2884-2896. https://doi.org/10.1109/TIFS.2018.2833032

[12] Simonyan, K., & Zisserman, A. (2015). Very Deep Convolutional Networks for Large-Scale Image Recognition. International Conference on Learning Representations (ICLR). https://arxiv.org/abs/1409.1556

[13] King, D. E. (2009). Dlib-ml: A Machine Learning Toolkit. Journal of Machine Learning Research, 10, 1755-1758. https://www.jmlr.org/papers/volume10/king09a/king09a.pdf

[14] Viola, P., & Jones, M. (2001). Rapid Object Detection using a Boosted Cascade of Simple Features. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), vol. 1, pp. I-511-I-518. https://doi.org/10.1109/CVPR.2001.990517

[15] Uricar, M., Franc, V., & Hlavac, V. (2012). Detector of Facial Landmarks Learned by the Structured Output SVM. International Conference on Automatic Face & Gesture Recognition (FG), pp. 1-8. https://doi.org/10.1109/FG.2012.9

[16] Zhang, X., Zou, J., He, K., & Sun, J. (2015). Accelerating Very Deep Convolutional Networks for Classification and Detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, 38(10), 1943-1955. https://doi.org/10.1109/TPAMI.2015.250257

IJARETY

INNO SPACE
SJIF Scientific Journal Impact Factor

International Journal of Advanced Research in Education and Technology

ISSN: 2394-2975 Impact Factor: 8.152

www.ijarety.in editor.ijarety@gmail.com