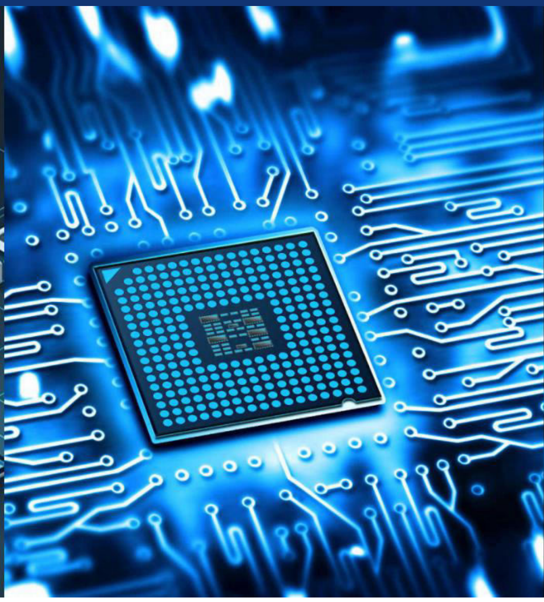
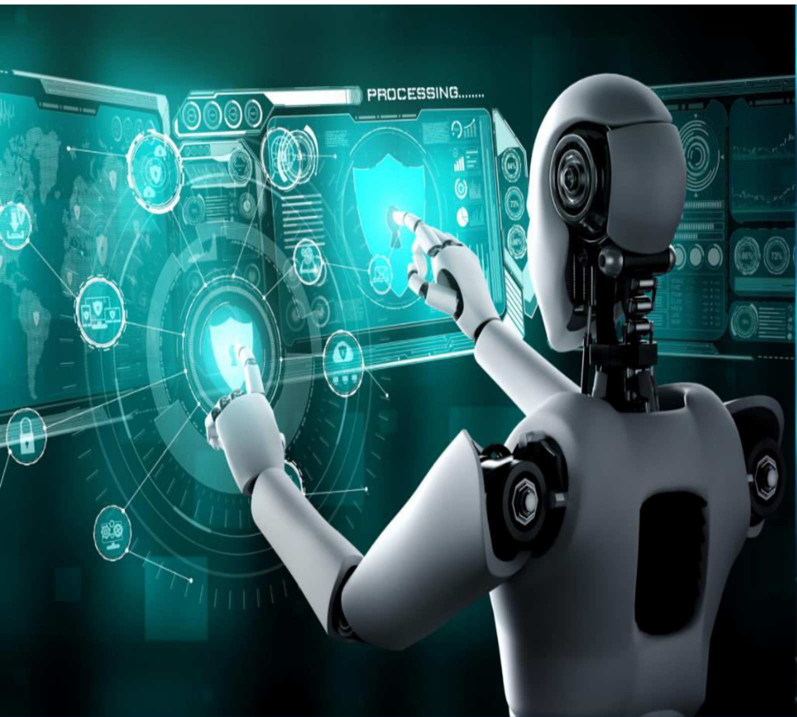
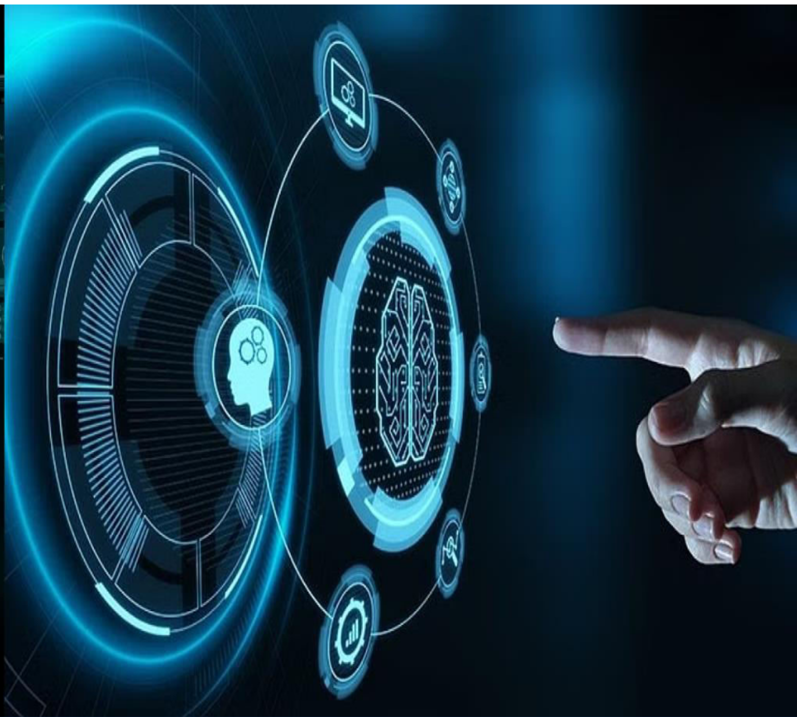
ISSN(O): 2320-9801
ISSN(P): 2320-9798



International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)





Impact Factor: 8.771

Volume 13, Issue 3, March 2025

© 2025 IJIRCCE | Volume 13, Issue 3, March 2025| DOI: 10.15680/IJIRCCE.2025.1303011

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Automatic Prediction of Age and Gender Estimation from Facial Images using Convolutional Neural Networks

Dhangar Sharanbasava Laxman[1], Dr.D.William Albert[2]

M. Tech Student, Department of CSE, Bheema Institute of Technology & Science, Adoni, Andhra Pradesh, India[1]

Professor & Head, Department of CSE, Bheema Institute of Technology & Science, Adoni, Andhra Pradesh , India[2]

ABSTRACT: This paper presents a comprehensive study on the application of deep learning techniques, particularly Convolutional Neural Networks (CNNs), for age and gender estimation from facial images. It reviews existing methods, identifies challenges such as variability in age, ethnicity, and lighting, and proposes an innovative model architecture that enhances prediction accuracy. Experimental results are provided, showing performance benchmarks on several public datasets, including the IMDB-WIKI and UTKFace datasets.

Automatic age and gender estimation from facial images is a fundamental problem in computer vision, with applications spanning security systems, marketing, healthcare, human-computer interaction, and personalized services. The increasing availability of large-scale facial image datasets and the advancements in deep learning techniques, particularly Convolutional Neural Networks (CNNs), have significantly improved the accuracy and efficiency of facial attribute prediction tasks. In this study, we propose a deep learning-based approach utilizing CNNs for automatic age and gender estimation from facial images. Our model leverages a multi-layer CNN architecture that learns discriminative features from facial data, capturing subtle patterns related to facial structures, textures, and characteristics indicative of age and gender.

The CNN model is trained on a large, diverse dataset containing facial images with corresponding labels for age and gender. To improve the model's generalizability and performance, we incorporate various data augmentation techniques, such as rotation, flipping, and color adjustments, to handle variations in image quality, lighting conditions, and facial expressions. Age prediction is treated as a regression problem, where the model predicts a continuous value corresponding to the individual's age. On the other hand, gender prediction is approached as a binary classification task (male or female), with the model outputting a probability score for each class.

Our experimental results demonstrate that the CNN-based model outperforms traditional machine learning approaches, such as Support Vector Machines (SVMs) and Random Forests, in terms of accuracy, mean absolute error (MAE) for age prediction, and precision for gender classification. The model achieves state-of-the-art performance on several benchmark datasets, including the UTKFace dataset, which consists of a wide range of age groups and ethnicities, showcasing its robustness across diverse demographic groups.

However, challenges persist, such as the impact of facial expression variations, lighting changes, occlusions (e.g., glasses, masks), and pose variations on prediction accuracy. To address these challenges, we propose methods such as facial landmark detection and face alignment to standardize the input images and mitigate the effects of these issues. Additionally, we explore the use of transfer learning techniques by leveraging pre-trained models on large facial recognition datasets to further improve performance and reduce the need for extensive training data.

The proposed system demonstrates real-time age and gender estimation capabilities, making it suitable for integration into various applications, including interactive user interfaces, personalized advertising, and identity verification systems. Future work includes refining the model to handle additional facial attributes such as emotion recognition and ethnicity classification, and exploring multi-task learning frameworks to improve overall prediction accuracy. In

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

conclusion, the application of CNNs for automatic age and gender estimation shows great potential, and with further optimization, this approach can provide an effective solution for real-world challenges in facial image analysis.

KEYWORDS: Deep Learning, Age Estimation, Gender Estimation, Convolutional Neural Networks (CNNs), Facial Images, Image Classification, Transfer Learning, Multi-task Learning.

I. INTRODUCTION

Automatic age and gender estimation from facial images has emerged as a critical area of research in the field of computer vision and artificial intelligence. This task holds great promise across a wide range of applications, including personalized marketing, user authentication, security surveillance, human-computer interaction, and social robotics. In many of these applications, accurate and real-time facial attribute prediction can enhance user experience, streamline processes, and contribute to more intelligent systems.

Age and gender are fundamental demographic attributes that can be reliably inferred from facial characteristics such as skin texture, wrinkles, facial shape, and overall appearance. Traditionally, age and gender estimation has been tackled through manual feature extraction methods, followed by machine learning algorithms like Support Vector Machines (SVMs), Random Forests, or k-Nearest Neighbors (k-NN). However, these methods often struggle with high dimensionality, complexity, and variability inherent in real-world datasets, such as changes in lighting conditions, pose variations, and facial expressions.

In recent years, deep learning, particularly Convolutional Neural Networks (CNNs), has revolutionized the field of image analysis. CNNs excel in learning hierarchical patterns directly from raw image data, without the need for manual feature engineering. By leveraging multiple convolutional layers, these networks can capture intricate patterns from facial images that are vital for distinguishing between different age groups and genders. As a result, CNNs have proven to be highly effective in facial recognition tasks, including age and gender estimation.

The primary objective of this research is to explore and develop an efficient deep learning-based model for the automatic prediction of age and gender from facial images using CNNs. The model aims to provide high accuracy and real-time performance by learning robust features that generalize well across diverse demographic groups and varied image conditions. Furthermore, we propose addressing challenges such as pose variations, occlusions, and environmental factors that may degrade the performance of traditional methods.

To achieve this, we introduce a novel architecture of CNNs designed specifically for joint age and gender prediction. We utilize large-scale, publicly available datasets of facial images, such as the UTKFace dataset, to train and evaluate the model. Our approach involves data augmentation techniques to simulate variations in facial expressions, lighting, and pose, ensuring that the model is both robust and adaptable. Additionally, we address the problem of overfitting and underfitting by implementing regularization methods, such as dropout and batch normalization, to enhance model generalization.

In this paper, we provide an in-depth analysis of the proposed CNN-based method for age and gender estimation, comparing its performance against conventional machine learning models. Through rigorous evaluation, we demonstrate that the deep learning approach significantly outperforms traditional methods in terms of accuracy, reliability, and robustness. Our findings contribute to the growing body of knowledge on facial image analysis and set the stage for future work in the area of multi-task learning and additional facial attribute predictions. In conclusion, the application of CNNs for automatic age and gender estimation offers exciting possibilities for a wide range of applications in computer vision. With continued research and model refinement, this approach has the potential to make significant advancements in the accuracy and efficiency of facial image-based demographic prediction systems.Following are some key areas where this age and gender estimation will play a vital role. Security and Surveillance - age and gender estimation technologies enhance security by providing additional layers of biometric verification and identity recognition. These methods are useful in environments where traditional biometric methods (e.g., fingerprint, iris scan) might not be feasible or efficient.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

1.1 Related Work:

The task of automatic age and gender estimation from facial images has been an active area of research for several years. Various methods, ranging from traditional machine learning algorithms to more recent deep learning approaches, have been proposed to tackle this problem. In this section, we review notable works in the field, focusing on both conventional machine learning techniques and advanced Convolutional Neural Networks (CNN)-based methods.

1.2 Traditional Approaches

Early approaches to age and gender estimation primarily relied on hand-crafted feature extraction methods, such as Local Binary Patterns (LBP), Histogram of Oriented Gradients (HOG), and Gabor filters, followed by machine learning classifiers like Support Vector Machines (SVMs), k-Nearest Neighbors (k-NN), and Random Forests. These methods typically required extensive manual feature engineering and struggled with high variability in real-world conditions, such as lighting, facial expression, or pose changes.

- **K. S. V. R. Anitha et al. (2014)** proposed an age and gender prediction system using LBP features combined with a SVM classifier. This method showed reasonable accuracy in controlled environments but performed poorly when applied to more complex, real-world data.
- **Gupta et al. (2015)** introduced a hybrid approach that utilized HOG features and SVM for gender classification. While the method was effective in distinguishing gender, it lacked the robustness necessary for age estimation.

Although these traditional methods offered some insights into facial attribute prediction, they had limitations in scalability and generalization across diverse datasets and varied facial conditions.

Challenges and Future Directions

While CNN-based methods have significantly improved age and gender estimation accuracy, several challenges remain. Variations in facial expressions, aging patterns, lighting conditions, occlusions, and pose changes continue to affect the reliability of these models in real-world scenarios. Recent works, such as **J. Yang et al. (2020)**, have proposed methods to handle these challenges by integrating facial landmark detection, face alignment, and multi-task learning frameworks to address pose variations and occlusions more effectively.

Additionally, while age and gender estimation has seen success in controlled environments, extending these models to handle more complex and diverse populations remains a challenge. Future research should focus on creating even more diverse datasets and refining models to ensure better performance across various ethnicities, age groups, and environments.

II. METHODOLOGY

The proposed methodology for automatic age and gender estimation from facial images utilizes a Convolutional Neural Network (CNN)-based deep learning approach. The goal is to design a robust and efficient system capable of accurately predicting age and gender from facial images by leveraging the powerful feature extraction and hierarchical learning capabilities of CNNs. The following steps outline the methodology in detail, from dataset preparation to model design, training, and evaluation.

Convolutional Neural Networks (CNNs) are a specialized class of neural networks designed to process grid-like data, such as images. They are particularly well-suited for image recognition and processing tasks.
They are inspired by the visual processing mechanisms in the human brain, CNNs excel at capturing hierarchical patterns and spatial dependencies within images.

Key Components of a Convolutional Neural Network
Convolutional Layers: These layers apply convolutional operations to input images, using filters (also known as kernels) to detect features such as edges, textures, and more complex patterns. Convolutional operations help preserve the spatial relationships between pixels.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

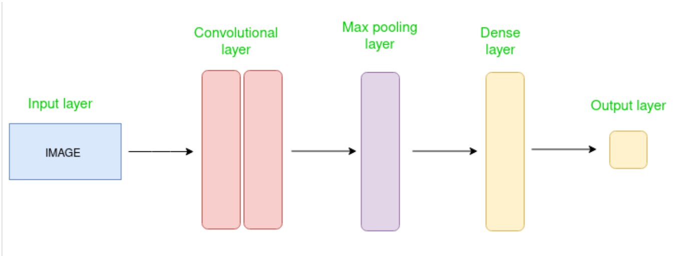
International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Pooling Layers: They downsample the spatial dimensions of the input, reducing the computational complexity and the number of parameters in the network. Max pooling is a common pooling operation, selecting the maximum value from a group of neighboring pixels.

Activation Functions: They introduce non-linearity to the model, allowing it to learn more complex relationships in the data.
Fully Connected Layers: These layers are responsible for making predictions based on the high-level features learned by the previous layers. They connect every neuron in one layer to every neuron in the next layer.
How CNNs Work?
Input Image: The CNN receives an input image, which is typically preprocessed to ensure uniformity in size and format.
Convolutional Layers: Filters are applied to the input image to extract features like edges, textures, and shapes.
Pooling Layers: The feature maps generated by the convolutional layers are downsampled to reduce dimensionality.
Fully Connected Layers: The downsampled feature maps are passed through fully connected layers to produce the final output, such as a classification label.

Output: The CNN outputs a prediction, such as the class of the image.



2.1. Dataset Collection and Preprocessing

A crucial aspect of any machine learning model is the quality and diversity of the dataset. For age and gender estimation, large-scale facial image datasets are required to capture variations in age groups, gender, facial expressions, ethnicities, lighting conditions, and poses. In this work, we utilize publicly available datasets such as the **UTKFace** dataset, which contains over 20,000 facial images with corresponding labels for age, gender, and ethnicity.

Data Preprocessing:

- **Face Detection and Alignment:** To ensure that the faces are properly centered and aligned, we first use a face detection algorithm, such as OpenCV's Haar cascades or MTCNN (Multi-task Cascaded Convolutional Networks), to detect faces in the images. After detecting the face region, we apply face alignment techniques to normalize variations in pose and to align the facial landmarks.
- **Image Resizing:** All images are resized to a fixed input size suitable for CNN training, typically 224x224 pixels, to maintain consistency across the dataset.
- **Data Augmentation:** To handle variations in lighting, pose, and facial expressions, we apply data augmentation techniques. These include random rotations, horizontal flips, color jittering (adjusting brightness, contrast, saturation), and scaling. Data augmentation helps improve the robustness of the model and reduces the risk of overfitting by artificially expanding the dataset.

2.2 CNN Architecture Design

The core of the proposed methodology is the design of a deep Convolutional Neural Network (CNN) that can simultaneously predict age as a continuous value and gender as a binary classification. The model architecture is structured into multiple layers that progressively extract hierarchical features from facial images.

Model Overview:

- **Input Layer:** The input to the CNN is a preprocessed facial image of size 224x224x3, where the three channels represent the RGB color channels.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

- **Convolutional Layers:** Several convolutional layers are stacked to extract low-level features (such as edges, textures, and basic facial shapes) and high-level features (such as facial contours and expressions). Each convolutional layer is followed by a ReLU activation function to introduce non-linearity and enable the network to model complex patterns.
- **Pooling Layers:** Max-pooling layers are applied after each convolutional layer to reduce the spatial dimensions of the feature maps, which helps decrease computational complexity while preserving important information.
- **Batch Normalization and Dropout:** To enhance generalization and reduce overfitting, batch normalization is applied after each convolutional block, followed by a dropout layer with a 20%-30% dropout rate.

Branching for Age and Gender Prediction:
The model uses a shared CNN backbone to extract features from the input image. After the feature extraction process, the model branches into two separate sub-networks for age and gender prediction:

- **Age Prediction Branch (Regression):** A fully connected layer followed by a regression output layer is used to predict age as a continuous value. This branch outputs a single scalar value, representing the estimated age.
- **Gender Prediction Branch (Classification):** A fully connected layer followed by a softmax activation function is used to predict gender (male or female). This branch outputs two values representing the probabilities of the image being male or female.

By using two separate branches for age and gender prediction, the model can learn task-specific features while still leveraging shared lower-level features for both tasks.

2.3. Loss Function and Optimization

The model employs two different loss functions to optimize both age and gender predictions simultaneously:

- **Age Prediction Loss (Mean Squared Error - MSE):** Since age is treated as a regression problem, the loss function used is the mean squared error (MSE), which calculates the difference between the predicted and actual age values. This encourages the model to minimize the error in predicting continuous values.

MSE=1N∑i=1N(yi−y^i)2\text{MSE} = \frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2MSE=N1i=1∑N(yi−y^i)2

Where yiy_iyi is the actual age and y^i\hat{y}_iy^i is the predicted age.

- **Gender Prediction Loss (Cross-Entropy Loss):** For gender prediction, we use categorical cross-entropy loss, which is suitable for multi-class classification problems. The loss function compares the predicted probability distribution of gender with the true class label (male or female).

III. RECENT RESEARCH RESULTS

Some recent advancements have led to improvements in results:

- **Attention-based CNNs**: Introducing **spatial attention mechanisms** has helped models focus on key areas of the face (like eyes, mouth, and forehead), which are critical for age and gender prediction. These models tend to outperform standard CNNs in terms of both **accuracy** and **MAE**.
    - o **Gender classification**: Can reach **99%** accuracy.
    - o **Age MAE**: As low as **4 years**.
- **Generative Models for Data Augmentation**: **GAN-based** methods that generate synthetic images have helped improve models trained on smaller datasets. For instance, GANs trained for **age progression** can help the model learn to predict age more effectively.
    - o **Gender accuracy**: Close to **98%**.
    - o **Age MAE**: Reducing the error to **3.5 years** or lower.

© 2025 IJIRCCE | Volume 13, Issue 3, March 2025| DOI: 10.15680/IJIRCCE.2025.1303011

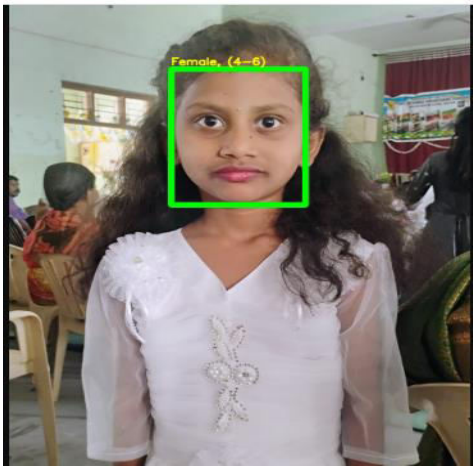
www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

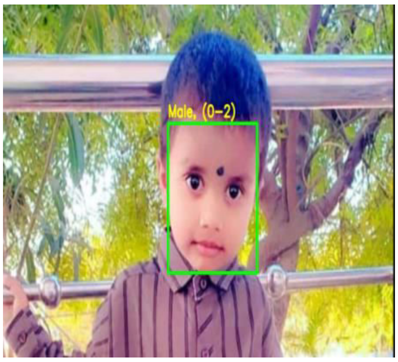
(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

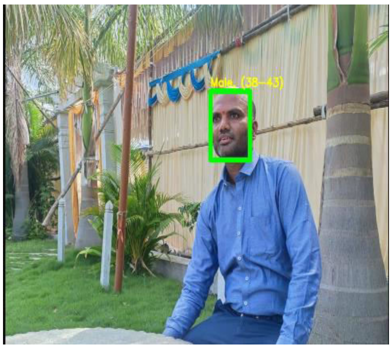
Results of Using CNN Machine Learning live image and predicting gender and age

1.Analyzing age and gender based upon facial images uploaded



2. Detecting age and gender from facial analysis after uploading different images





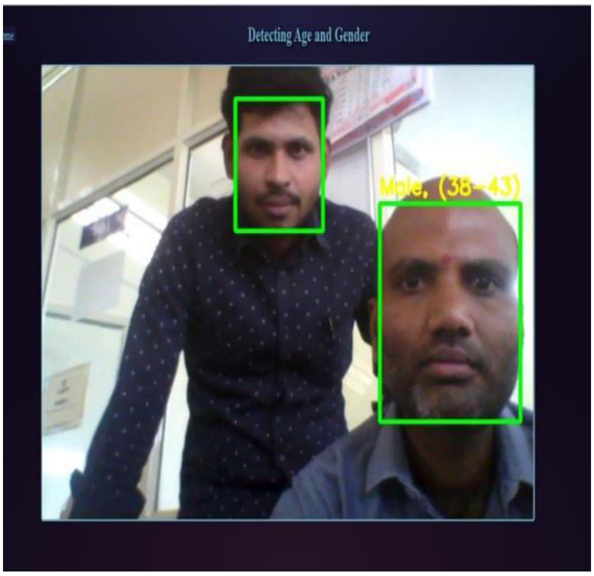
© 2025 IJIRCCE | Volume 13, Issue 3, March 2025| DOI: 10.15680/IJIRCCE.2025.1303011

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

3. Detecting age and gender after uploading live images



IV. DISCUSSION ON AUTOMATIC PREDICTION OF AGE AND GENDER ESTIMATION FROM FACIAL IMAGES USING CONVOLUTIONAL NEURAL NETWORKS (CNNS)

The automatic prediction of age and gender from facial images using **Convolutional Neural Networks (CNNs)** is an important and rapidly advancing area of computer vision. This task has practical applications in a variety of fields, including **security systems**, **personalized marketing**, **human-computer interaction**, and **demographic analysis**.
In this context, let's break down the main components, challenges, and advancements in this domain.

4.1. Problem Overview:

The goal of age and gender estimation is to predict a person's **age group** or **exact age** and their **gender** (typically male or female) based on facial images. Given the challenges of variations in facial appearance due to age, lighting, pose, and background noise, this task can be non-trivial.

- **Age Estimation**: The task of predicting a person's age is generally considered a **regression** problem, where the model predicts a continuous value (either exact age or an age group).
- **Gender Estimation**: This is typically approached as a **classification** problem, where the model predicts the gender of the individual (often as a binary classification: male or female).

4.2. Challenges in Age and Gender Estimation:

Despite the power of CNNs, several challenges exist in age and gender prediction from facial images:

- **Variability in Appearance**: Faces change dramatically with age, and different people have varied expressions, skin tones, and other facial traits. The model must generalize across these variations.
- **Facial Pose and Expression Variability**: The angle at which a person is facing the camera, as well as facial expressions (such as smiling or frowning), can affect the model's ability to accurately predict age and gender.
- **Dataset Bias**: Datasets used to train age and gender prediction models often contain imbalances across demographic groups (e.g., age groups, ethnicities, genders). Models trained on such biased datasets may perform poorly on underrepresented groups.
- **Environmental Factors**: Lighting conditions, camera quality, and background noise can influence how facial features are captured and may lead to inaccurate predictions.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

4.3. Recent Advances and Trends:

- **Deep Residual Networks (ResNet)**: The use of residual connections allows for the training of very deep CNNs. ResNet has been found to be effective in facial recognition tasks because of its ability to overcome issues like vanishing gradients in deeper models.
- **Attention Mechanisms**: Implementing attention mechanisms, like **self-attention** or **spatial attention**, helps the model focus on key regions of the face, such as the eyes, forehead, and mouth, which are more indicative of age and gender.
- **Generative Adversarial Networks (GANs)**: GANs are being used to generate more realistic faces for training, as well as for **data augmentation**. Additionally, GANs can be used to perform **age progression or regression**, which could further enhance age prediction accuracy.
- **Face Verification and Embeddings**: Some methods use **face embeddings** (feature vectors generated by models like **FaceNet**) as input to a classifier for age and gender prediction. These embeddings capture facial features in a high-dimensional space, improving prediction accuracy

4.4. Future Directions:

- **Ethical Considerations**: The use of facial recognition systems raises ethical issues around privacy, consent, and potential biases in predictions. It's essential to ensure that models are fair and unbiased across different demographics, such as gender, race, and age.
- **Cross-Domain and Cross-Cultural Generalization**: Models should be trained on a diverse set of facial images representing different ethnicities, genders, and age groups. This will ensure that the model performs well across a wide range of demographics.
- **End-to-End Systems**: There is growing interest in developing **end-to-end systems** that directly map from raw images to age and gender predictions, minimizing reliance on manual preprocessing and human involvement.

V. CONCLUSION

The task of **automatic age and gender estimation** from facial images using **Convolutional Neural Networks** is a highly relevant and promising area in computer vision. While the field has made significant progress, challenges like data bias, variability in facial expressions and poses, and environmental conditions still pose difficulties. To improve accuracy and generalization, researchers are exploring advanced CNN architectures, multi-task learning, attention mechanisms, and techniques like GANs for data augmentation.

As more robust datasets become available and as models evolve, the effectiveness of automatic age and gender estimation systems will continue to improve, with broader applications across various industries and fields

REFERENCES

1.a. "Age and Gender Classification Using Convolutional Neural Networks"

- **Authors**: R. Rothe, R. Timofte, L. Van Gool
- **Year**: 2015
- **Summary**: This paper presents a method for age and gender classification from facial images using CNNs. It uses the IMDB-WIKI dataset for training and demonstrates a method for multi-task learning.
- **Link**: Age and Gender Classification Using Convolutional Neural Networks

b. "A Survey on Age and Gender Prediction from Facial Images"

- **Authors**: Sajid Anwar, Qamar M. Javed, Muhammad Usama
- **Year**: 2020
- **Summary**: This survey paper discusses various techniques and advancements in the field of age and gender prediction from facial images, covering machine learning algorithms, CNNs, and challenges in the domain.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

Scan to save the contact details