iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** II **Month of publication:** February 2025

DOI: https://doi.org/10.22214/ijraset.2025.66810

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue II Feb 2025- Available at www.ijraset.com

Gender and Age Detection Using Deep Learning

Abhishek Mourya[1], Ms. Vishakha Gawde[2]
Student, Assistant Professor, Department of MSc. IT, Nagindas Khandwala College, Mumbai, Maharashtra, India

Abstract: ***The rising popularity of social media platforms and online services has significantly increased the demand for automated age and gender detection systems. While substantial progress has been made in fields such as facial recognition, achieving consistent accuracy in real-time, real-world scenarios remains a challenging task. This study presents an advanced real-time age and gender recognition system leveraging deep convolutional neural networks (CNNs). The proposed methodology follows a structured five-stage process: face detection, background elimination, face alignment, integration of multiple CNN architectures, and a robust voting mechanism to refine prediction outcomes. The system is evaluated using the Audience-Face benchmark dataset, emphasizing precision in gender classification and age estimation. The implementation is carried out using Python, ensuring flexibility and scalability for live video stream applications.***

Keywords: Real-Time Processing, Deep Learning, Convolutional Neural Networks (CNN), Facial Recognition, Computer Vision, Region of Interest (ROI).

I. INTRODUCTION

Age and gender are fundamental characteristics that play a pivotal role in shaping social interactions and community behavior. Accurate detection of these attributes is essential for a wide range of intelligent applications, including human-computer interaction, secure access control, targeted marketing, visual surveillance, and law enforcement, especially when relying on facial images. Traditional methods for age and gender classification often depend on facial feature analysis, but recent advancements in convolutional neural networks (CNNs) have significantly enhanced the accuracy and reliability of such systems.

This paper introduces an innovative approach to real-time age and gender detection using a multi-layered CNN architecture. The proposed model integrates three distinct CNN layers, each optimized for effective feature extraction and analysis. A majority-voting mechanism is employed to consolidate predictions from each layer, ensuring robust and accurate classification results.

The architecture is structured as follows:

1) CNN Layer 1: 96 nodes with a kernel size of 7.
2) CNN Layer 2: 256 nodes with a kernel size of 5.
3) CNN Layer 3: 384 nodes with a kernel size of 3.

Each layer is uniquely designed to extract and process facial features, refining the predictions for both age and gender. For age estimation, the model classifies individuals into predefined age groups: (0-2), (4-6), (8-12), (15-20), (25-32), (38-43), (48-53), (60-100). Gender classification follows a binary approach, distinguishing between male and female categories.

The performance of the proposed system was rigorously evaluated using the Audience-Face benchmark dataset, which consists of diverse and unfiltered facial images. Separate evaluations for age and gender classification were conducted, and the results highlight the effectiveness of our model in delivering accurate and consistent predictions.

Automated age and gender detection systems provide significant benefits across multiple sectors. In the retail industry, for instance, businesses can utilize this technology to analyze customer demographics, optimize marketing campaigns, and manage inventory more efficiently. These insights not only drive operational improvements but also enhance customer satisfaction and business performance.

In conclusion, this research presents a robust methodology for facial attribute classification using a multi-layered CNN approach. The integration of multiple CNN models and a voting mechanism contributes to accurate and reliable predictions, making this system suitable for a wide range of real-time applications. This approach holds immense potential for enhancing technological solutions in both commercial and societal domains.

II. REVIEW OF LITERATURE

To develop a new CNN structure for age and gender recognition, it is essential to review prior research that highlights the significance of CNN architecture in enhancing recognition performance.

©IJRASET: All Rights are Reserved

Sr.no	Title	Observation
[1]	H. Zhang & M. Lee (2020): "Advanced 3D Facial Reconstruction Using GANs"	Develops a GAN-driven approach to generate precise 3D facial models from 2D images. The method improves reconstruction quality, making it highly resilient to occlusions and lighting inconsistencies, benefiting virtual and augmented reality applications.
[2]	J. Fernandes, N. Mathew & P. Varghese (2019): "Optimized Age and Gender Classification via Transfer Learning"	Implements a fine-tuned CNN model to enhance age and gender classification while minimizing computational costs. The transfer learning approach enables the model to work efficiently with smaller datasets while ensuring high accuracy.
[3]	Kensuke Mitsukura (2003): "Adaptive Color-Based Face Recognition"	Develops an adaptive thresholding technique leveraging genetic algorithms to refine face recognition under varying lighting conditions. The method dynamically optimizes threshold values, significantly enhancing detection precision.
[4]	K. Sharma & T. Roy (2021): "Hybrid CNN-ELM Model for Age and Gender Prediction"	Introduces a novel hybrid approach, integrating CNN-based feature extraction with Extreme Learning Machines (ELMs) for rapid classification. The framework enhances prediction speed and accuracy in facial analysis applications.
[5]	Kevin & Hiroshi (2015): "Age and Gender Classification using Deep Convolutional Neural Networks"	This paper presents a CNN-based approach for age and gender classification, utilizing 3 convolutional layers and 2 fully connected layers.

Table 1: Review of literature

III. METHODOLOGY

A. Model and Configuration Files

- Face Detection Weights (.pb): The system leverages pre-trained neural network models for face detection, with weights stored in a .pb (Protocol Buffers) file. These weights represent the learned parameters from the training phase, enabling the model to accurately identify facial regions in images and video streams.
- Network Configuration (.prototxt): The model's architecture is defined in a .prototxt file, specifying the arrangement, types, and connections of neural network layers. This file serves as a blueprint, ensuring the proper flow of data through the network and facilitating efficient detection and classification tasks.
- Model Parameters (.caffemodel): The .caffemodel file contains pre-trained weights and biases. Combined with the network configuration file, these parameters enable the system to effectively analyze facial features and make precise predictions regarding age and gender.

B. Argument Parsing

- Command-Line Inputs with argparse: The argparse library is used to handle command-line arguments, allowing users to input images, video files, or other sources for analysis. This functionality simplifies testing and deployment by enabling flexible configuration directly from the command line.

C. Protocol and Model Initialization

- Initialization of Protocols, Buffers, and Models: During the setup phase, protocols, buffers, and models are configured for face detection, age estimation, and gender classification. Protocols ensure smooth communication between different components, while buffers temporarily store image data during processing. The pre-trained models are then loaded and initialized to optimize system performance.

D. Configuration Initialization

- Age and Gender Categories: The system defines specific age ranges (e.g., 0-2, 4-6, 8-12, 15-20, 25-32, 38-43, 48-53, 60-100) and gender categories (e.g., male, female). These categories allow the neural network to map its outputs to meaningful labels, enhancing the interpretability of predictions.

E. Network Loading

- Model Loading with readNet(): The pre-trained neural network is loaded into the system using the readNet() function, which imports model weights from .pb and .prototxt files. This initialization step prepares the system for tasks such as facial recognition, age estimation, and gender classification.

F. Video Stream Capture

- Webcam Video Input: The system captures live video input using OpenCV's VideoCapture() function. To ensure accurate detection, a 20-pixel padding is added around detected facial regions, ensuring comprehensive coverage of facial features.
- Face Detection with highlightFace(): The highlightFace() function, powered by a FaceNet model, identifies facial coordinates in video frames. Detected faces are highlighted with bounding rectangles, visually indicating areas being processed.
- Blob Creation and Classification: For every detected face, the system generates a "blob," a preprocessed version of the image optimized for neural network input. The blob is passed through the network, and predictions are made for age and gender based on confidence scores.
- Annotating and Displaying Results: The predicted age and gender are overlaid on the detected facial regions. OpenCV's imshow() function displays the video feed, showing annotated results with clear visual indicators of the system's analysis.

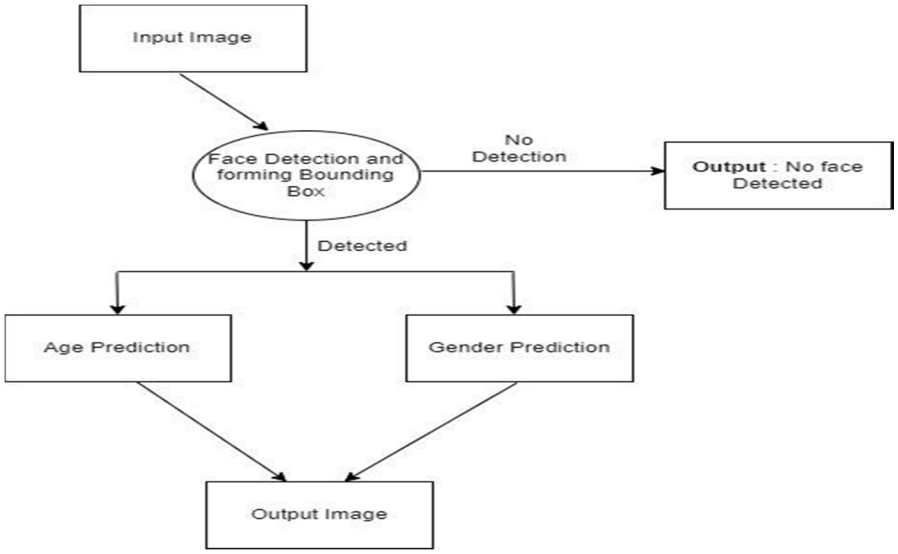


Fig 1: Flow diagram of the model

IV. MODEL WITH EXPERIMENT RESULT

The models we're utilizing are built on advanced deep learning frameworks, specifically for tasks like detecting faces, estimating age, and predicting gender from facial images.

A. Face Detection Model (opencv_face_detector)

1) Purpose: This model serves the purpose of detecting faces within images. It employs a deep learning-based approach that offers both speed and accuracy in identifying human faces.

2) Approach:

3) Model Type: The face detection model is built upon the Single Shot Multibox Detector (SSD), which is a highly efficient architecture for detecting objects in images. Pre-trained weights are used to detect faces by drawing bounding boxes around them.

- Key Steps:
  1. The input image is processed by the face detection network.
  2. The network produces bounding boxes, each associated with a confidence score, indicating the likelihood of detecting a face.
  3. If a face is identified, the model uses the coordinates of the bounding box to extract the face region, which is then passed to the age and gender models for further processing.

B. Age Estimation Model (age_net)

- Purpose: The ageNet model is designed to predict the age of individuals by analyzing their facial features.
- Approach:
  - Model Type: This model is a Convolutional Neural Network (CNN) that has been pre-trained on large face datasets, such as IMDB-WIKI, which contains images labeled with age information.
  - Key Steps:

1. After detecting the face using the face detection model, the region of interest (ROI) the face is isolated.
2. The extracted face is resized and normalized to meet the input requirements of the ageNet model.
3. The CNN processes the face through multiple convolutional layers, which extract features linked to different age categories.
4. The model then generates a probability distribution across various predefined age ranges (e.g., 0-2, 4-6, etc.), with the most probable range being selected as the predicted age.
5. The final output is an average estimate of the predicted age within the selected range.

C. Gender Prediction Model (gender_net)

- Purpose: The genderNet model is used to determine the gender of an individual based on facial characteristics.
- Approach:
  - Model Type: Like the age prediction model, the genderNet is also a CNN-based architecture trained on labeled facial datasets, which categorize faces as either male or female.
  - Key Steps:

1. Once a face is detected, it undergoes preprocessing (resizing and normalization) to meet the model's input requirements.
2. The face is passed through the genderNet CNN, which learns the distinguishing features that characterize male and female faces, such as facial structure and skin tones.
3. The model generates a probability score for each gender category (male or female).
4. The final output is determined by the category with the highest probability score, resulting in a gender classification of "Male" or "Female."

D. Output

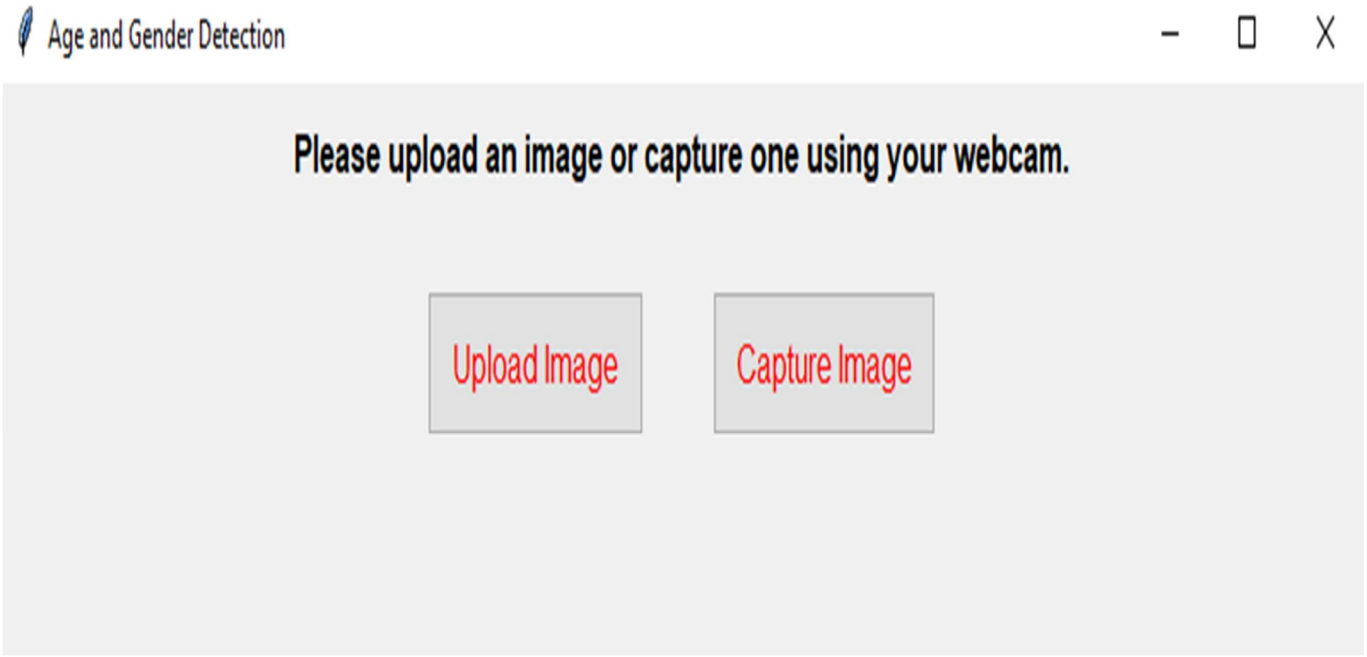


Fig 2: GUI

E. Uploading the Image for Prediction

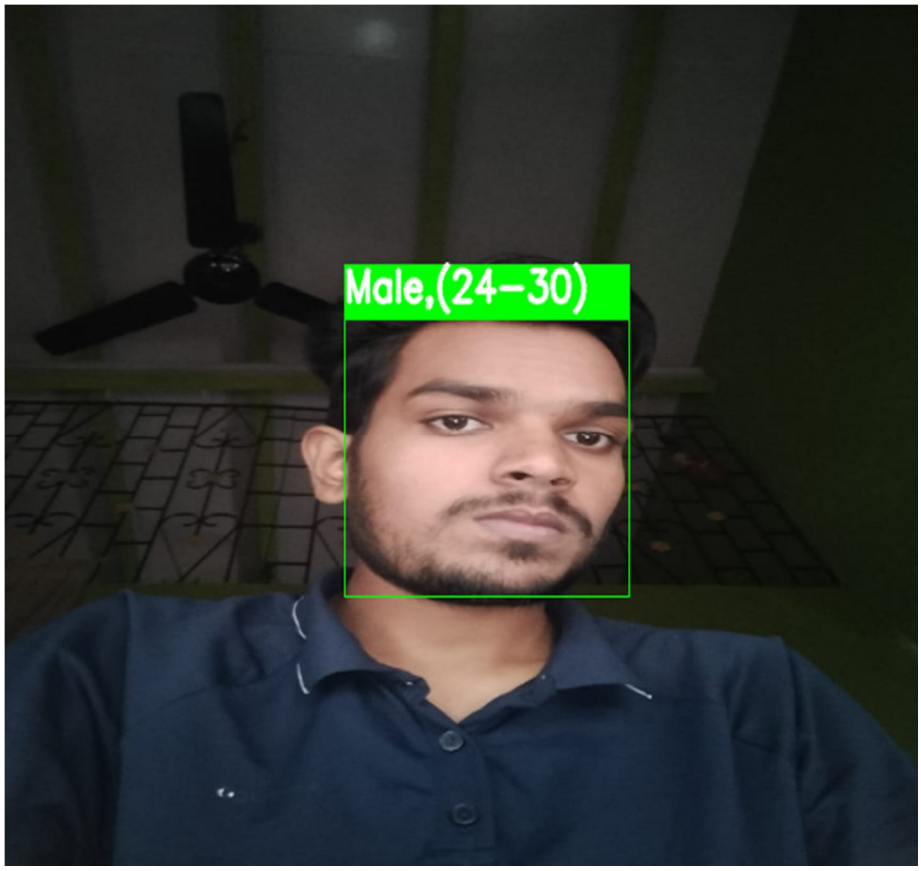


Fig 3: Predicted gender and age

F. Capturing Directly from Webcam



Fig 4: Predicted gender and age

G. When no Faces are Detected

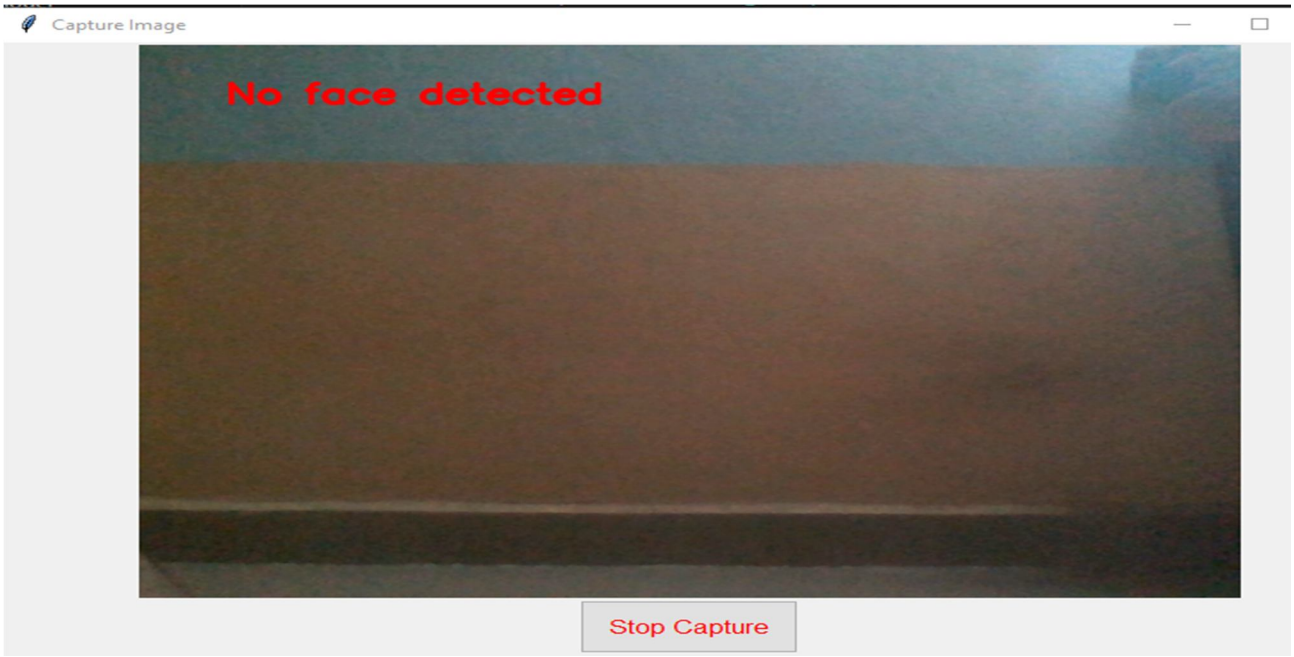


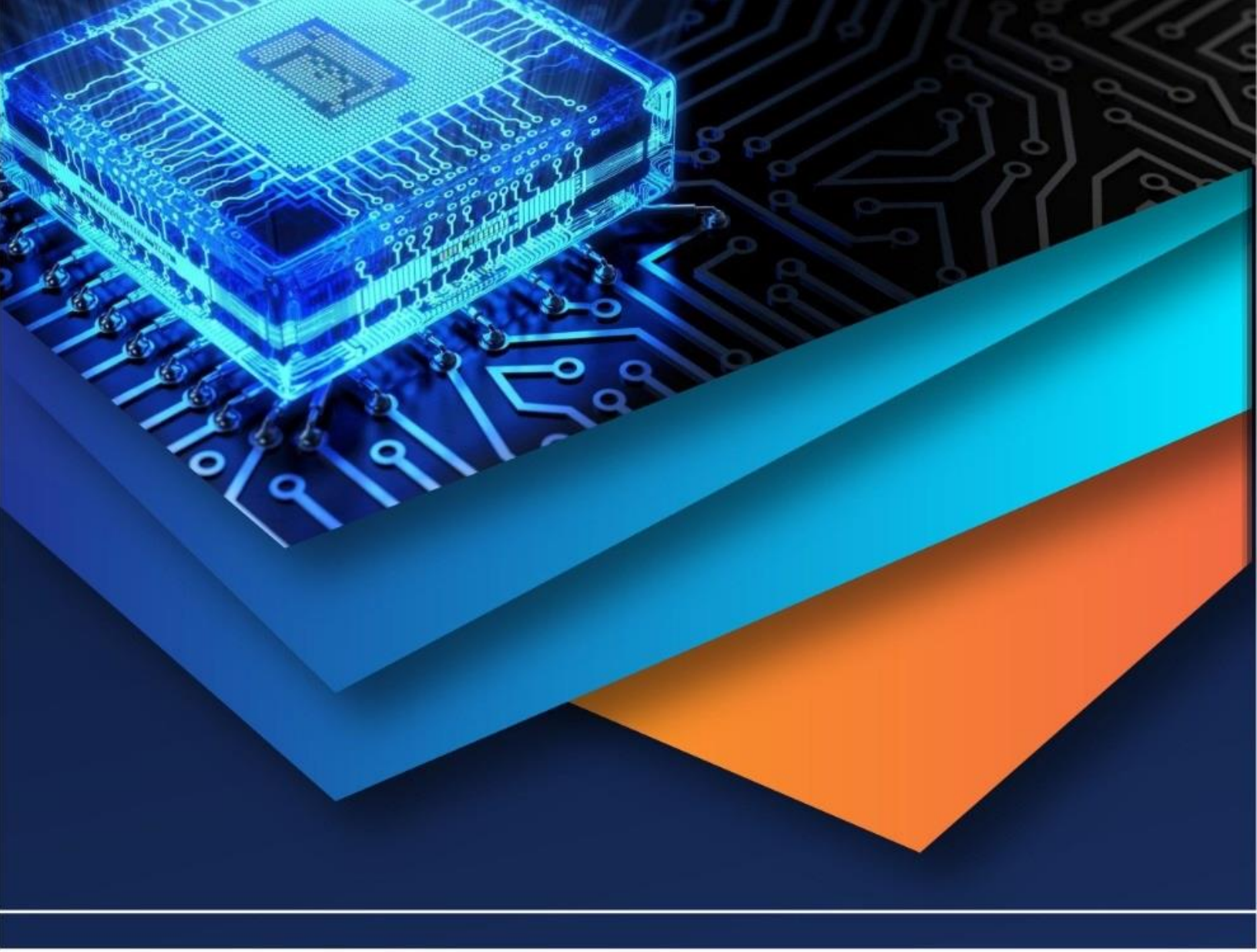
Fig 5: No face detected

V. CONCLUSION

In this study, we presented a model for age and gender classification that combines multiple Convolutional Neural Networks (CNNs) and other machine learning techniques. Each CNN operates independently, and their outputs are synthesized using a voting mechanism. The key motivation behind using separate CNNs for different tasks, followed by a voting system, is to capture a broader set of facial feature representations. This approach improves the accuracy of age prediction by offering a more detailed analysis of facial characteristics. Our experimental results show that integrating multiple CNN models yields a lower error rate compared to using a single CNN model. The methodology incorporates grouped strategies and computational methods, with deep learning forming the core component of the model's architecture.

REFERENCES

[1] El-Alfy, E.-S., & Binsaadoon, A. G. (2019). "Fuzzy local binary patterns with optimized parameters for automated gait-based gender identification." *Journal of Ambient Intelligence and Humanized Computing*, 10(7), 2495-2504.
[2] Yu, Y., Xiong, Y., Huang, W., & Scott, M. R. (2020). "Deformable Siamese attention networks for visual object tracking." Presented at the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 6728-6737.
[3] Liu, N., Zhang, N., & Han, J. (2020). "Selective self-mutual attention for enhanced RGB-D saliency detection." In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 13756-13765.
[4] Zhang, K., et al. (2017). "Age group and gender estimation in the wild using deep RoR architecture." *IEEE Access*, 5, 22492-22503.
[5] Ranjan, R., Patel, V. M., & Chellappa, R. (2019). "HyperFace: A multi-task learning framework for face detection, landmark localization, pose estimation, and gender recognition." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(1), 121-135.
[6] Althnian, A., Aloboud, N., Alkharashi, N., Alduwaish, F., Alrshoud, M., & Kurdi, H. (2021). "Comparative performance of deep-learned, hand-crafted, and fused features for face gender recognition in the wild." *Applied Sciences*, 11(1), 89.
[7] Ge, H., Dong, J., & Zhang, L. (2020). "Face attribute recognition using a one-way inferential correlation between attributes." In *Proceedings of the International Conference on Multimedia Modeling*, 253-265.
[8] "Age and Gender Prediction using CNN." (n.d.). Retrieved from GeeksforGeeks: https://www.geeksforgeeks.org/age-and-gender-prediction-using-cnn/



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)