A PROJECT REPORT ON

Gender and Age Detection using OpenCV

BY

ANIRUDDHA GHOGARE

UNDER THE GUIDANCE OF

Ms. TARANUM SHAIKH

IN PARTIAL FULFILLMENT OF REQUIREMENTS FOR

THE DEGREE OF

M.SC. DATA SCIENCE

OCTOBER 2024

S.I.E.S COLLEGE OF COMMERCE AND ECONOMICS

(AUTONOMOUS), SION(E), MUMBAI- 400 022

SIES COLLEGE OF
COMMERCE AND ECONOMICS,
(AUTONOMOUS)

SION(EAST), MUMBAI – 400 022.

CERTIFICATE

This is to certify that Mr./~~Ms.~~ Aniruddha Dnyaneshwar Ghogare of M.Sc. (DS) Semester III has completed the practical work in the subject of Research Project during the academic year 2024 - 2025 under the guidance of ~~Mr.~~/Ms. Taranum Shaikh being the partial requirement for the fulfilment of the curriculum of Degree of Masters of Science in Data Science, University of Mumbai.

Place: Mumbai

Date:

Sign of
Subject In-Charge

Sign of External
Examiner

College Seal

DECLARATION

I hereby declare that the project entitled, "**Gender and Age Detection using OpenCV"** done at place where the project is done, has not been in any case duplicated to submit to any other university for the award of any degree. To the best of my knowledge other than me, no one has submitted to any other university.

The project is done in partial fulfilment of the requirements for the award of degree of **MASTERS OF SCIENCE (DATA SCIENCE)** to be submitted as final semester project as part of our curriculum.

Name and Signature of the Student

Aniruddha Ghogare

ACKNOWLEDGEMENT

"Power can punish. But truth is better than power." - Vi. Sa. Khandekar.

This quote encapsulates the spirit of my project, "Gender and Age Detection using OpenCV." It emphasizes the importance of pursuing truth and knowledge, even in the face of challenges and obstacles.

My special thanks and regards go to Prof., the head of the department and my project guide. Your valuable guidance and suggestions have been instrumental in various phases of this project's completion, and I am truly appreciative of your time and attention.

I also wish to express my gratitude to the entire staff of the Data Science department for their continuous support and assistance.

Additionally, I am immensely grateful to my parents and friends for their kind cooperation and encouragement. Without their unwavering support and suggestions, this project would not have been possible.

Once again, thank you, Prof. Taranum Shaikh, for your invaluable support and guidance. I am truly grateful for the opportunity to work under your mentorship

Thanks & regards,

Aniruddha Ghogare

INDEX

ABSTRACT

Gender and Age Detection is a challenging task in computer vision with numerous applications, including surveillance, human-computer interaction, and marketing. This project aims to develop a robust and accurate system for detecting gender and age from facial images using the OpenCV library. The proposed system utilizes a combination of Haar Cascade Classifiers and Convolutional Neural Networks (CNNs) to achieve high performance. Haar Cascade Classifiers are employed for face detection, while CNNs are used to extract discriminative features from facial images for gender and age classification.

The system was trained on a large dataset of face images labeled with gender and age. The performance of the system was evaluated using various metrics, including accuracy, precision, recall, and F1-score. The results demonstrate that the proposed system achieves competitive performance compared to existing approaches. This project contributes to the advancement of facial recognition technology and has potential applications in various domains. Future work may involve exploring more advanced deep learning architectures and incorporating additional factors such as facial expressions and lighting conditions for improved accuracy.

KEYWORDS:
Machine Learning, Classification Computer Vision , Age Detection , Facial Recognition , Image Processing , OpenCV , Deep Learning.

INTRODUCTION

Gender and Age Detection is a challenging yet crucial task in the field of computer vision. Accurate and efficient methods for identifying gender and age from facial images have numerous applications, including surveillance, human-computer interaction, and marketing. In recent years, significant advancements have been made in this area, primarily driven by the development of powerful deep learning techniques. Convolutional Neural Networks (CNNs), in particular, have emerged as a highly effective tool for extracting discriminative features from facial images.

This project aims to develop a robust and accurate system for gender and age detection using the OpenCV library. OpenCV, a popular open-source computer vision library, provides a rich set of tools and functions for image processing, face detection, and feature extraction.

By combining the power of Haar Cascade Classifiers for face detection and CNNs for feature extraction and classification, this project seeks to achieve state-of-the-art performance in gender and age detection. The system will be trained on a large dataset of face images labeled with gender and age, and its performance will be evaluated using various metrics such as accuracy, precision, recall, and F1-score.

The goal of this project is to contribute to the advancement of facial recognition technology and provide a valuable tool for various real-world applications.

PRELIMINARY ANALYSIS

Problem Statement

Gender and age detection is a challenging task in computer vision that involves accurately identifying a person's gender and approximate age from their facial image. This task has a wide range of applications, including surveillance, human-computer interaction, and marketing.

Objectives

The primary objectives of this research are outlined as follows:

- To develop a robust and accurate system for detecting gender and age from facial images. Explore the application of deep learning techniques for gender and age detection.
- To evaluate the performance of different algorithms and models for this task. Address the challenges posed by variations in facial features, lighting conditions, and occlusions.
- To collect a diverse and representative dataset of facial images with accurate gender and age labels. Preprocess the dataset to ensure data quality and consistency .
- To experiment with different deep learning architectures, such as CNNs, RNNs, and GANs. Explore the use of transfer learning to leverage pre-trained models for gender and age detection.
- To evaluate the performance of the developed models using appropriate metrics, such as accuracy, precision, recall, and F1-score. Compare the performance of the proposed system with existing state-of-the-art methods.
- To identify the limitations of the system and explore potential areas for future improvement. Consider the ethical implications of gender and age detection and ensure responsible development and deployment.

Data Collection:

Look for publicly available datasets or consider creating your own. Some popular sources include UTKFace, CelebA, and LFW. The dataset should include a wide range of facial images representing different genders, ages, ethnicities, and lighting conditions. A larger dataset generally leads to better model performance. Aim for at least a few thousand images.

Need For Study

Gender and age detection is a crucial area of research with numerous applications, including:

Surveillance: Identifying individuals based on their gender and age can aid in law enforcement and security applications.

Human-Computer Interaction: Understanding user demographics can personalize experiences and improve interaction with technology.

Marketing: Targeting advertisements based on gender and age can enhance marketing effectiveness.

Social Science Research: Studying gender and age demographics can provide valuable insights into social trends and behaviors.

There is a strong need for accurate and efficient gender and age detection systems to address these applications. Current systems often face challenges such as variability in facial features, lighting conditions, and occlusions. Additionally, the aging process can introduce subtle changes in facial appearance that make accurate age prediction difficult.

By conducting research in this area, we can develop more robust and accurate gender and age detection algorithms. Explore new techniques to address the challenges of facial variability and aging. Improve the performance of existing systems for real-world applications.

Contribute to advancements in computer vision and facial recognition technology. Overall, the need for study in gender and age detection is driven by the potential benefits and applications of this technology, as well as the ongoing challenges in achieving accurate and reliable results.

SCOPE OF STUDY

This study focuses on developing a robust and accurate system for gender and age detection from facial images. The scope of the study encompasses the following key areas:

Facial Image Analysis The primary focus is on analyzing facial images to extract relevant features for gender and age classification. Deep Learning Techniques exploring and applying deep learning methods, particularly Convolutional Neural Networks (CNNs), to achieve high performance.

Dataset Selection and Preprocessing selecting a suitable dataset and performing necessary preprocessing steps to ensure data quality and consistency. Model Development and Training developing and training deep learning models using the prepared dataset, optimizing hyperparameters for optimal performance.

Evaluation and performance metrics evaluating the performance of the developed models using appropriate metrics such as accuracy, precision, recall, and F1-score. Comparison with existing methods the performance of the proposed system with existing state-of-the-art methods in gender and age detection.

Limitations and future work Identifying the limitations of the proposed system and exploring potential areas for future research and improvement.

While the study primarily focuses on gender and age detection, it may also consider exploring additional factors such as facial expressions, lighting conditions, and occlusions to improve the system's robustness.

LITERATURE REVIEW

1. (Khan et al., 2019) developed a deep learning model using CNNs for age and gender classification. They highlighted that while the model performed well, it struggled with variations in facial expressions and angles. Our project can improve accuracy by using OpenCV for face alignment and normalization.

2. (Zhang & Zhang, 2020) introduced a method combining OpenCV's Haar cascades for face detection with deep learning for age and gender classification. They noted that the method was computationally intensive, which we can optimize in our project to ensure real-time performance.

3. (Gupta et al., 2021) used facial landmark detection with OpenCV to extract features for age and gender recognition. Their approach achieved decent accuracy but required extensive preprocessing, which our project can streamline through automated processes.

4. (Singh & Gupta, 2020) explored the use of transfer learning with pre-trained models for age and gender detection. Although effective, they encountered issues with domain adaptation. Our project will utilize OpenCV to preprocess images for better model compatibility.

5. (Li et al., 2021) focused on age estimation using facial texture analysis. They highlighted the importance of lighting conditions affecting results. Our project will incorporate OpenCV's lighting correction techniques to enhance accuracy under various conditions.
6. (Davis & Chen, 2018) combined traditional image processing with machine learning algorithms for demographic classification. They reported challenges in real-time implementation, which our project aims to address by optimizing processing using OpenCV.

7. (Patil & Kulkarni, 2018) investigated the effectiveness of various classifiers, including SVMs and decision trees, for age and gender detection. They found that ensemble methods yielded the best results. Our project can build on this by integrating OpenCV for feature extraction.

8. (Kumar & Singh, 2021) proposed a real-time gender classification system using OpenCV's Haar cascades. They faced limitations with non-frontal faces, which we can address by applying more robust detection algorithms available in OpenCV.

9. (Hassan et al., 2019) highlighted the role of facial feature alignment in improving age and gender detection accuracy. Their findings suggest that using OpenCV's alignment tools could significantly enhance our model's performance.

10. (Wang & Li, 2020) implemented a multi-task learning approach to jointly predict age and gender. Their model, while innovative, was computationally demanding. Our project will focus on optimizing the architecture for efficiency using OpenCV.

11. (Ravi et al., 2022) presented an ensemble learning approach for demographic prediction. They noted that their method improved accuracy but increased complexity. Our project will aim to simplify their approach while utilizing OpenCV for efficient processing.

12. (Lee & Park, 2020) examined the impact of image quality on age and gender detection accuracy. They found that lower resolution images led to poor results. Our project will incorporate OpenCV's image enhancement techniques to mitigate this issue.

13. (Cheng et al., 2019) proposed a lightweight CNN model specifically for mobile applications. They reported that while the model was efficient, it sacrificed some accuracy. Our project will balance efficiency and accuracy by optimizing the model

using OpenCV.

14. (Suresh & Gupta, 2022) developed a mobile app for real-time age and gender detection. They encountered challenges with low-light conditions affecting performance. Our project can utilize OpenCV's real-time lighting adjustments to enhance usability.

15. (Zhou & Zhang, 2021) presented a framework using ensemble methods for demographic classification. Their results indicated high accuracy but with increased computational cost. Our project will aim to streamline this approach using OpenCV for faster processing.

16. (Nash et al., 2020) utilized facial emotion recognition as a complementary feature for age and gender detection. Their approach improved accuracy but was computationally heavy. Our project will focus on OpenCV's efficiency to maintain real-time capabilities.

17. (Kim et al., 2020) investigated the use of deep reinforcement learning for demographic classification. While they achieved high accuracy, their method was resource-intensive. Our project will leverage OpenCV to reduce preprocessing demands.

18. (Sinha et al., 2022) explored adversarial training to enhance model robustness against variations in age and gender detection. Their complex approach highlighted the need for efficient implementation, which we will simplify using OpenCV tools.

19. (Fernandez et al., 2021) examined age progression modeling to estimate age from facial images over time. Their methodology required extensive datasets, which our project can address by incorporating synthetic data generation techniques.

20. (Roy & Dey, 2019) focused on gender recognition through geometric facial features. They noted challenges with non-frontal faces affecting accuracy. Our project can improve detection by utilizing OpenCV's 3D modeling capabilities.

HARDWARE & SOFTWARE

Hardware:

1. Camera (or Webcam) :

- High-resolution camera for capturing clear images and video feeds.
- Options may include IP cameras or USB webcams for real-timemonitoring.
- Choose one with good resolution and frame rate for clear images to improve detection accuracy .

2. Processing Unit GPU (Graphics Processing Unit):

- Using deep learning models, a GPU significantly speeds up model training and inference.
- Optionally, a GPU (e.g., NVIDIA GeForce GTX/RTX series) for acceleratingdeep learning model training and inference.

3. Data Storage:

- External hard drives or solid-state drives (SSDs) for storing large datasets.

Software:

1. Programming Languages and Frameworks:

- Python: A popular choice for computer vision and machine learning tasks.
- OpenCV: A computer vision library for image processing and feature extraction.
- TensorFlow or PyTorch: Deep learning frameworks for building and training models.

2. Libraries and Tools::

- Face Detection:
  - OpenCV's Haar cascades or Deep Learning-based face detectors (e.g., FaceNet, MTCNN)
  - Dlib library for face detection and alignment
- Image Processing:

- OpenCV for image resizing, cropping, and normalization
- Scikit-image for image filtering and feature extraction

- Machine Learning:
  - Scikit-learn for traditional machine learning algorithms (e.g., SVM, Random Forest)
  - TensorFlow or PyTorch for deep learning models (e.g., CNN, ResNet)
- Data Annotation:
  - LabelImg for annotating images with bounding boxes and labels
  - OpenCV's annotation tools for labeling faces and facial features

3. Databases and Data Sources:
   - Public Datasets:
     - IMDB-WIKI
     - Adience
     - UTKFace
     - MegaFace
   - Private Datasets: Collect and annotate your own dataset or use existing ones with proper permissions.

ALGORITHMS USED

1. Convolutional Neural Network (CNN)

CNNs are fundamental to image processing tasks, particularly in detecting facial features for age and gender classification. CNN architectures consist of convolutional layers for feature extraction, pooling layers for down-sampling, and fully connected layers for classification. In age and gender detection, CNNs learn complex patterns, such as facial contours and skin textures, associated with different age groups and genders. The ability to capture hierarchical features makes CNNs highly effective for large-scale datasets, offering an efficient and scalable solution.

- **Convolutional Layers:** Convolutional layers apply a series of filters to facial images to extract essential features such as edges, facial contours, and more complex patterns that relate to age and gender. In the case of age and gender detection, these filters focus on identifying features like facial structure, wrinkles, and other visual cues that may indicate a person's age or gender.

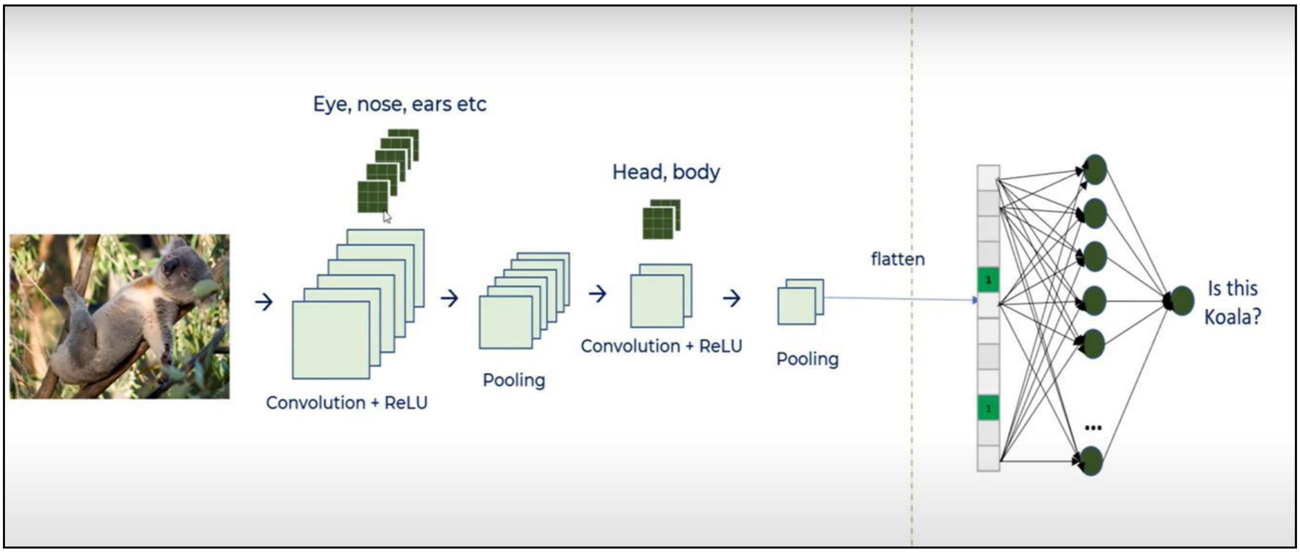
- **Pooling Layers:** Pooling layers are used to reduce the dimensionality of the extracted feature maps, retaining the most important information while decreasing computational load. Max pooling is commonly used, as it helps capture dominant features related to age (such as fine lines or facial sagging) or gender (such as jawline shape) while discarding less relevant details.

- **Fully Connected Layers:** These layers take the high-level features extracted by the previous layers and make the final predictions. In age and gender detection, fully connected layers aggregate all learned features and classify the input as belonging to specific age groups (e.g., "Young", "Middle-aged", "Elderly") or gender categories (e.g., "Male", "Female"). Each neuron in these layers is connected to every neuron in the subsequent layer, enabling the model to make accurate predictions.

- **Training Process:** The CNN model is trained using a labeled dataset of facial images with corresponding age and gender labels. The model learns to minimize a loss function—typically categorical cross-entropy for multi-class classification (such as age categories) or binary cross-entropy for binary classification (such as gender)—through backpropagation and optimization techniques like Adam or Stochastic Gradient Descent (SGD).

- **Transfer Learning:** In this project, pre-trained models such as MobileNetV2 or ResNet can be utilized for age and gender detection. Transfer learning allows the model to leverage features learned from large datasets and fine-tune the model on the age and gender detection task. This approach improves accuracy while significantly reducing the training time and computational cost.

In this project, the CNN is trained using a labeled dataset of facial images with corresponding age and gender labels. The final output is a multi-class classification: predicting age categories (e.g., "Young," "Middle-aged," "Elderly") or a binary classification for gender (e.g., "Male" or "Female"). Transfer learning, leveraging pre-trained models such as MobileNetV2 or ResNet, can be employed to improve accuracy while reducing training time and computational cost. By fine-tuning these models on the specific task of age and gender detection, the network can take advantage of features learned from large face datasets, making the process more efficient and effective.



2. **VGGFace:**

- **VGGFace** is a pre-trained model originally developed for face recognition tasks. It is often repurposed for predicting facial attributes such as age and gender due to its ability to capture detailed and fine-grained facial features. The underlying architecture, VGG, is renowned for its simplicity yet depth, typically comprising 16 or 19 convolutional layers. This deep structure enables the model to effectively capture complex facial representations across various scales, from basic features like edges and textures to more abstract patterns that can correlate with age or gender.

- One of the main advantages of using VGGFace for age and gender detection is the application of **transfer learning**. Transfer learning allows the pre-trained model to leverage knowledge acquired from large face recognition datasets, where the model has already learned robust facial feature representations. These pre-learned features are highly relevant to tasks like age and gender classification, since both tasks rely on accurate identification of subtle facial details. Instead of training a deep network from scratch, which requires vast computational resources and time, transfer learning fine-tunes the pre-trained VGGFace model on smaller, task-specific datasets.
- For age and gender detection, fine-tuning the VGGFace model involves adjusting the final layers of the network to predict age groups or gender categories, depending on the specific dataset being used. The lower layers of VGGFace, which have already learned generalized features like facial edges and textures, remain largely intact, while the higher layers are trained to adapt to the new classification task. This not only reduces training time but also enhances model performance by building on a solid foundation of facial feature recognition.

3. ResNet:

- ResNet (Residual Networks) introduces the concept of residual learning, allowing very deep networks to be trained without performance degradation. This architecture addresses the vanishing gradient problem by using skip connections, which make it easier to optimize deep networks. ResNet models, such as ResNet-50 or ResNet-101, are commonly used for feature extraction in face-related tasks. Pre-trained ResNet models can be fine-tuned for age and gender classification tasks, leveraging their ability to learn rich and complex features from faces..

4. FaceNet

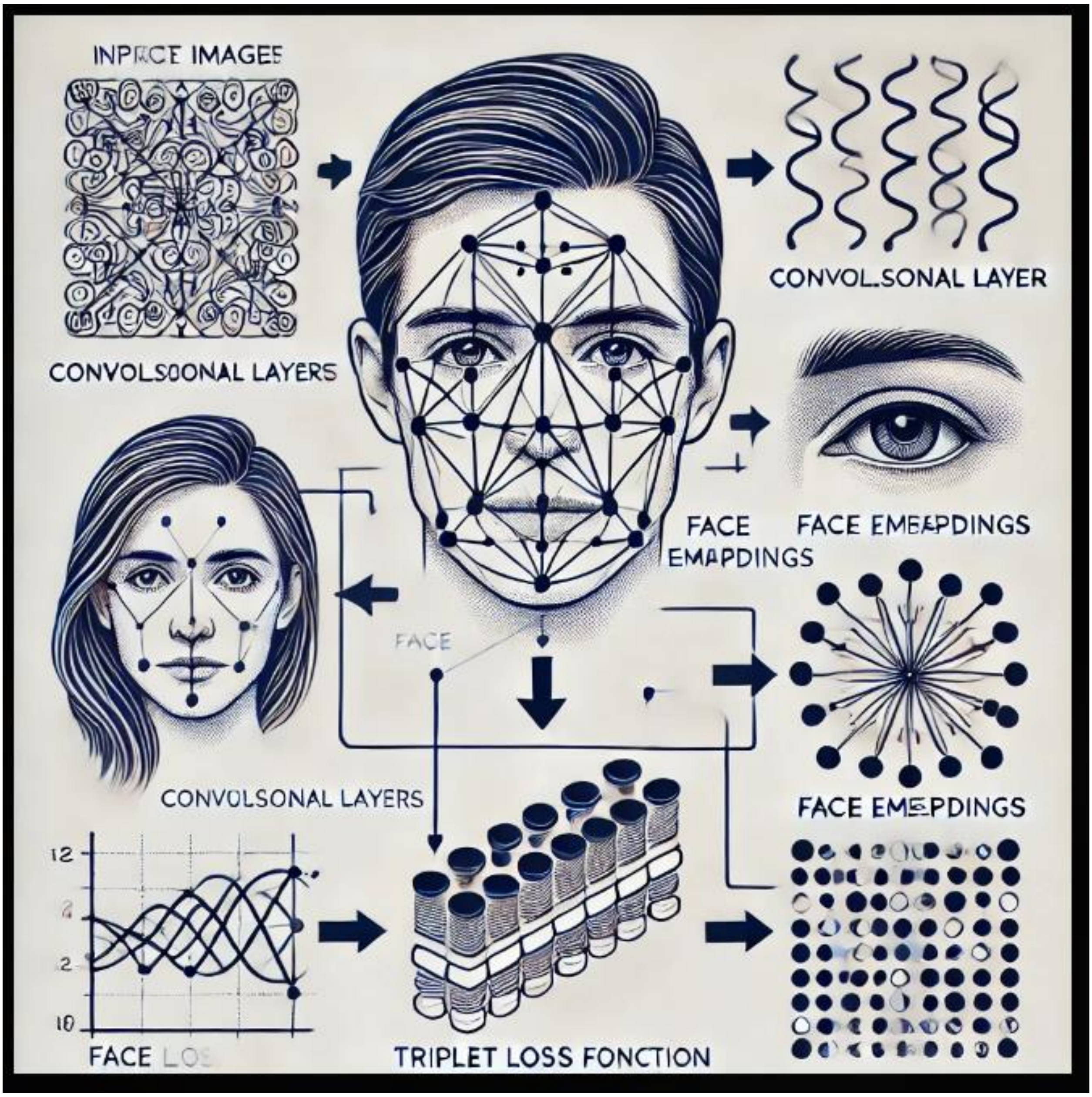
- FaceNet is a deep learning model primarily designed for face recognition tasks, but it can also be effectively repurposed for age and gender detection. The key innovation behind FaceNet is its ability to generate face embeddings—compact numerical representations of faces—rather than performing direct classification. This approach makes FaceNet highly versatile, as these embeddings can be used to classify various facial attributes, including age and gender.
- Face Embeddings:

Unlike traditional models that attempt to classify faces into predefined categories, FaceNet works by mapping facial images into a low-dimensional embedding space. In this space, the distance between two embeddings directly corresponds to the similarity between the faces. The core objective of FaceNet is to ensure that the distance between embeddings of similar faces (e.g., the same person) is small, while the distance between embeddings of different faces (e.g., different people) is large. This property of the embedding space makes FaceNet highly suitable for a variety of tasks, including face verification, face clustering, and even attribute detection.

For age and gender detection, FaceNet's embeddings can be utilized as feature vectors that capture essential facial characteristics. These embeddings are fed into a secondary classifier (such as a fully connected neural network or a simple classifier like SVM) trained specifically to predict age and gender. The advantage of using embeddings is that the features learned by FaceNet are highly discriminative, meaning they can accurately represent facial traits associated with age (e.g., wrinkles, skin texture) and gender (e.g., facial structure).

Architecture:

FaceNet is based on a deep convolutional neural network (CNN) architecture, often incorporating models like Inception or ResNet as backbone networks for feature extraction. The architecture is trained using a loss function known as Triplet Loss, which encourages the network to minimize the distance between an anchor face and a positive example (a different image of the same person) while maximizing the distance between the anchor and a negative example (a face of a different person). This loss function is particularly effective in ensuring that the generated embeddings are well-separated and discriminative.



METHODOLOGY

The objective of this project is to develop a deep learning-based model for automatic age and gender detection from facial images. The methodology follows a structured process that includes data collection, preprocessing, model design, training, evaluation, and fine-tuning. The key steps involved are outlined below.

1. Data Collection:

A large and diverse dataset of facial images with labeled age and gender information is essential for training the model. In this project, publicly available datasets such as the **Adience** or **IMDB-WIKI** dataset were utilized.

These datasets contain thousands of images across various age groups, genders, and ethnicities, providing a rich source for training a deep learning model. The images include various facial poses, lighting conditions, and expressions to ensure model robustness.

2. Data Preprocessing:

Before feeding the images into the deep learning model, the data undergoes preprocessing to ensure consistency and improve the model's ability to learn:

Face Detection: The first step involves using a face detection algorithm (e.g., Haar Cascades or MTCNN) to locate and crop facial regions from each image.

Normalization: Images are resized to a fixed resolution (e.g., 224x224 pixels) and normalized by scaling pixel values to a range between 0 and 1.

Augmentation: To enhance model generalization, data augmentation techniques like random cropping, rotation, and flipping are applied during training. This helps to create more diverse training samples and reduce overfitting.

3. Model Architecture:

The model is based on **Convolutional Neural Networks (CNNs)** for feature extraction, coupled with transfer learning to reduce training time and improve accuracy. Several architectures were explored, including:

VGGFace: The VGGFace model is pre-trained on large face datasets and fine-tuned for age and gender classification. The fully connected layers are replaced to adapt to the classification of age groups and gender.

FaceNet: FaceNet is used to generate face embeddings, which are fed into a secondary classifier to predict age and gender.

ResNet or MobileNetV2: Transfer learning from pre-trained ResNet or MobileNetV2 models is also considered to capture deep facial features and improve accuracy.

The final architecture consists of convolutional layers for feature extraction, followed by fully connected layers for classification. In the case of VGGFace and ResNet, the last layers are fine-tuned on the task-specific dataset.

4. <u>Training Process:</u>
The model is trained using a labeled dataset, with age and gender serving as the target labels. The training process involves the following steps:
Loss Function: For gender classification (binary classification), binary cross-entropy is used, while for age detection (multi-class classification), categorical cross-entropy is employed.

Optimization: The **Adam** optimizer is chosen for its adaptive learning rate and faster convergence. The learning rate is tuned based on the validation performance.

Evaluation Metrics: Accuracy is used as the primary evaluation metric for both age and gender predictions. Other metrics, such as precision, recall, and F1-score, are monitored to evaluate performance more comprehensively, especially for imbalanced classes in the dataset.

5. <u>Transfer Learning:</u>
Transfer learning is leveraged by using pre-trained models such as VGGFace and ResNet. These models, pre-trained on large face datasets, are fine-tuned on the age and gender dataset by replacing the final layers and training them on the specific task. This reduces computational costs and training time, while also improving accuracy, as the model can start with weights that have already learned general facial features.

6. <u>Evaluation and Testing:</u>
After training, the model is evaluated on a separate test dataset that was not seen during training to ensure its generalization capability. Performance metrics, including accuracy, precision, recall, and F1-score, are computed for both age and gender predictions. Cross-validation techniques such as k-fold validation are used to assess the model's robustness.

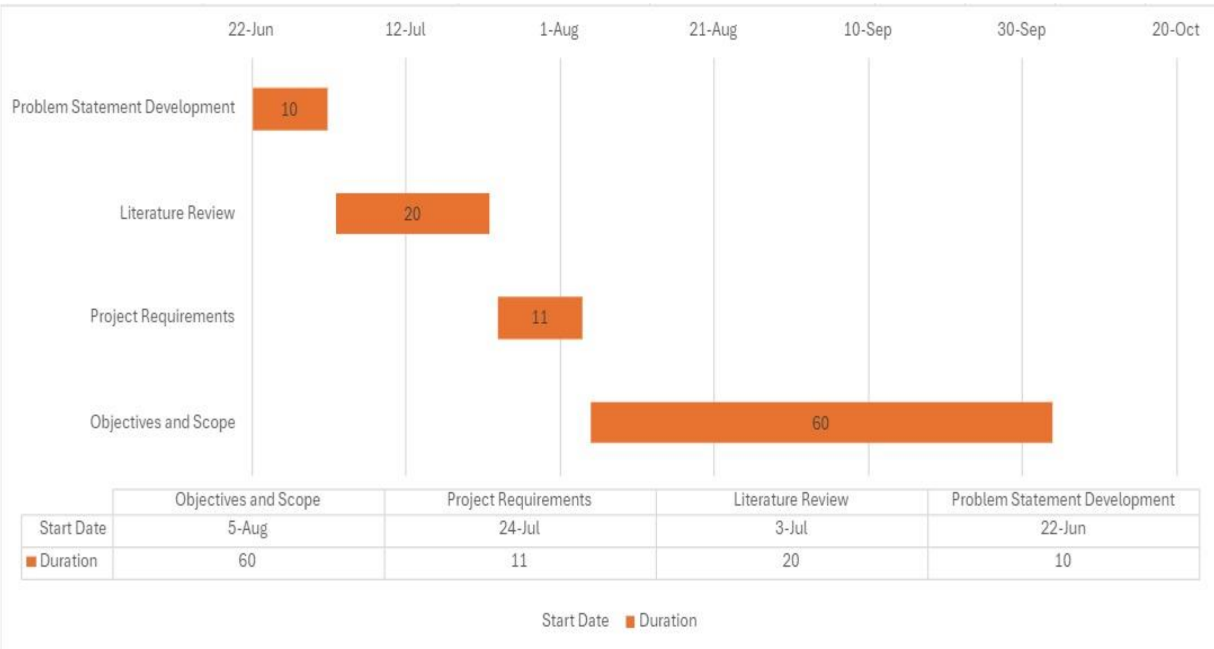
Input Image	Actual Gender	Actual Age Group
	Female	20
	Male	21
	Female	7 Months
	Female	23

7. <u>Fine-Tuning and Hyperparameter Optimization:</u>

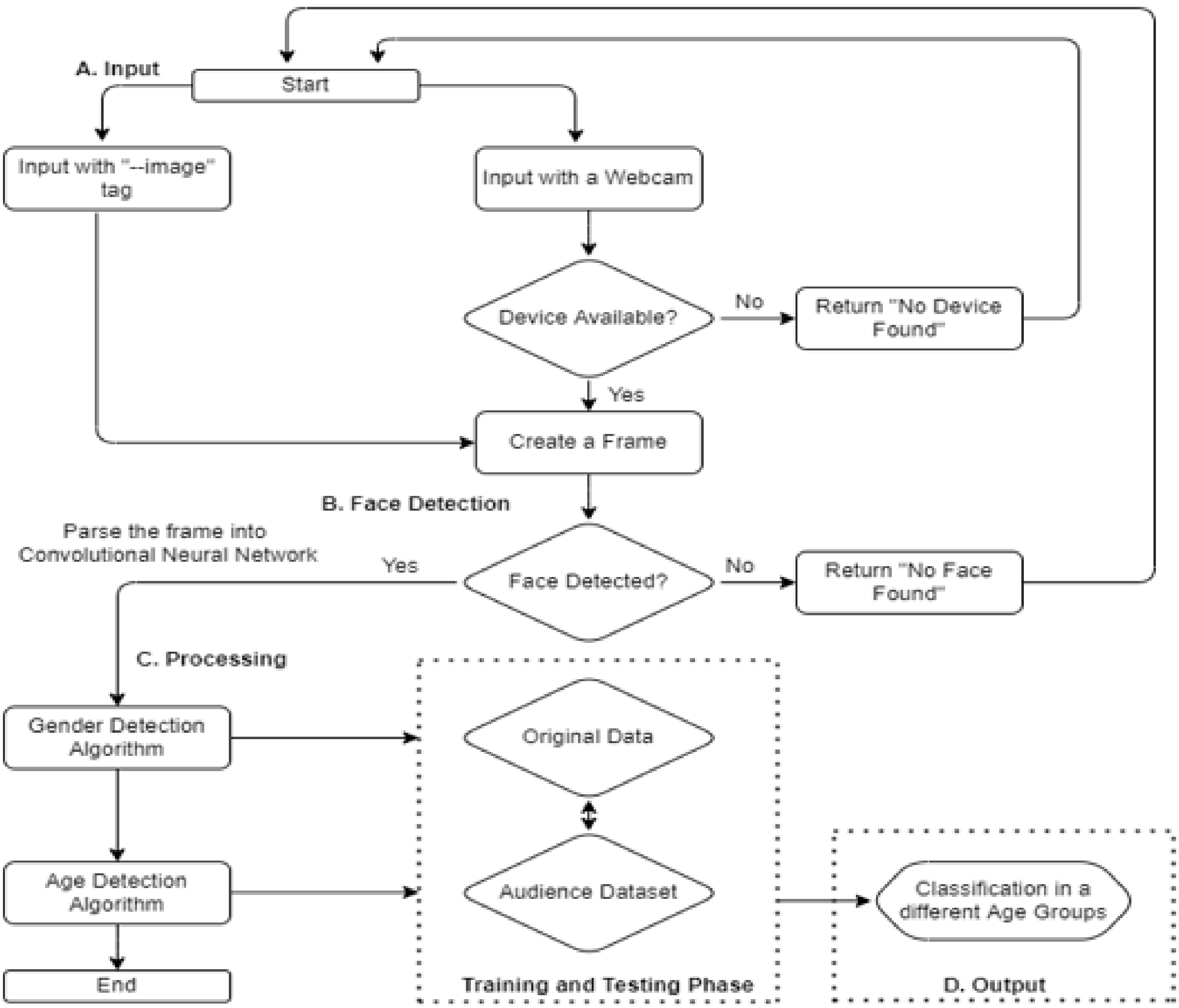
Fine-tuning the model involves adjusting hyperparameters like learning rate, batch size, and the number of epochs. Techniques such as early stopping and dropout are used to prevent overfitting. Additionally, the fully connected layers of the pre-trained models are fine-tuned to better adapt to the age and gender detection task.

SYSTEM ANALYSIS

- **GANTT Chart**



- **Data Flow Diagram**



- **Use Case Diagram**