Acknowledgment

On behalf of our project team, we would like to extend our heartfelt appreciation to everyone who has played a crucial role in the successful completion of project. We are immensely grateful to project guide Prof. S. S. Bhandare for her invaluable contributions, guidance, and support during this initial phase. Your expertise and dedication have been instrumental in shaping our project's trajectory. Our team members and colleagues deserve special recognition for their unwavering commitment and hard work. Together, we have tackled challenges, met deadlines, and achieved important milestones that have brought us to this point. We also want to acknowledge the guidance from senior faculty Prof.I. Priyadarshini , she has provided us with the support and motivation necessary for our success. This project has been a collaborative effort, and we are thankful for the collective contributions. As we move forward into subsequent stages, we anticipate your continued support and partnership.

Kombde Aditya Milind
Mukund Pranav Namdeo
Thakre Parimal Nitin
Khan Umair Ahmad Masood Ahmad
(B.E. Computer Engg.)

Contents

List of Figures

List of Tables

Abstract

Cross-age person identification in forensic settings poses significant challenges due to the unpredictable changes in facial features caused by aging. The impact of these challenges on criminal investigations and the search for missing persons underscores the pressing need for effective solutions that transcend generational boundaries. Traditional approaches, which separately address age synthesis and face recognition, often falter in real-world situations where aging effects exhibit considerable variability. In this project, the proposed Cross-Age Person Identification (CAPID) for Forensic Applications, is a pioneering system that addresses these challenges using advanced deep learning techniques. A distinctive feature of CAPID is its integration of age synthesis and person identifying tasks within a unified multiple learning task framework. By combining these tasks, CAPID achieves superior synergy and performance, especially in the realm of forensic applications. Prior to feature extraction Image Quality Evaluation (IQE) is performed on the dataset. After IQE, if the Image Quality Matrice(IQM) meets the threshold condition then Image is pre-processed. Pre-processing includes illumination normalization and noise removal. Along with pre-processing pose normalization is used to diminish difficulties present in feature abstraction and thus enhances the recognition. This technique equips investigators and law enforcement agencies with robust tools to match individuals across various age groups, thereby aiding in the resolution of complex cases.

Chapter 1

Introduction

1.1 Project Idea

The project's objective is to develop a dependable face recognition system that is age-invariant and functions effectively for a range of age groups. The method utilizes sophisticated neural network topologies, multitask learning, and attention mechanisms. The study primarily focuses on dissecting face features into components related to age and identity. AIFR, or age-invariant facial recognition, uses a variety of techniques to increase its accuracy, including identification conditional blocks and attention modules. The project also includes age synthesis skills, which allow age-progressed or -regressed facial images to be created while retaining identifying characteristics. The methodology uses attention regularization, dropout techniques, and advanced loss functions to optimize the models.

1.2 Motivation of the Project

- The pressing need for facial recognition algorithms that are both accurate and adaptable enough to alter as people age naturally is what motivated this effort.Conventional face recognition algorithms occasionally have trouble identifying faces across different age groups due to the significant changes in facial features that come with aging. Numerous real-world situations, including identity management, law enforcement, and cyber threats, depend on (doi.org) the solution to this issue. The objective is to develop an age-insensitive facial recognition system that will increase the accuracy and dependability of these systems. This will enable the mechanisms must work well, irrespective of the age of the individuals being identified.

- Innovation and Creativity: These contribute to the development of new techniques and algorithm-based face detection and identification systems, which may improve our daily lives by resolving issues with the current system.
- Learning Advantages

  It provides the chance to learn about the issues with the current systems, how to solve them, and what new technologies might be able to help improve the existing algorithm and create better iterations and new techniques for face and image recognition and detection.

Chapter 2

Literature Survey

- Kishore Kumar et. al.[1] To develop an age-invariant face recognition (AIFR) system, it propose a novel methodology consisting of five sequential processes: Image Quality Evaluation (IQE), Pre-processing, Pose Normalization, Feature Extraction and Fusion, and Feature Recognition and Retrieval. Image Quality Evaluation (IQE) process assesses the input image's quality and removes any images that fall below a certain quality threshold. This improves the AIFR system's performance, especially when the input images are noisy or poorly lit. Preprocessing phase performs two tasks: illumination normalization and noise removal. Illumination normalization reduces the effects of various lighting conditions on the face's appearance. Noise removal eliminates image noise, which can improve feature extraction accuracy. Pose Normalization step aligns the face image in a standardized pose, making feature extraction and recognition easier. It uses EA-AT algorithm for pose normalization in system. Feature Extraction and Fusion extract two types of features from the pre-processed and pose normalized image:texture features and shape features. CNNs are used to extract texture features,while SIHKS is used to extract shape features. It then fuse the extracted texture and shape features using CCA to reduce the feature vector's dimensionality and improve its discriminative power. Feature Recognition and Retrieval use an SVM clas- sifier to recognize and retrieve faces from the database. SVMs are powerful machine learning algorithms well-suited for AIFR tasks. The Limitation to this system is SVM and RBF in high dimensional space makes it difficult to interpret learned decision boundaries.

- Zhizhong et. al.[2] The paper provides a framework, which is multiple task termed Fcae MTL, to handle both age-invariant face recognition (AIFR) and face age synthesis (FAS) tasks jointly. MTLFace can understand the age-invariant identity-related representation for person identification while obtaining pleasing person structure synthesis for model interpretation.Specifically, Face MTL model proposes an attention-based feature extraction to break down the mixed face features into two uncorrelated components—age and face identity features in a spatially acceptable way. Instead of the traditional one-hot encoding that achieves group-level age synthesis for face, Face MTL provides a novel identity conditional module to achieve age synthesis for face for identity level, which can improve the smoothness for age of synthesized faces through sharing weight approach. Leveraging the high-quality synthesized faces from FAS and the suggested multi-task architecture, MTLFace employs a novel selective fine-tuning technique to further increase AIFR. This has the very little disadvantage of not being able to adequately maintain the background.

- Amal et. al.[3] Orthogonal Embedding CNNs (OE-CNNs) for Age- Invariant Face Recognition. Age-invariant face recognition (AIFR) (arxiv.org) is a challenging task due to the notable intraclass differences brought on by aging over time. This work suggests a unique method, called Orthogonal Embed- ding CNNs (OE-CNNs), to learn the age-invariant deep facial features in order to lessen the intra-class discrepancy brought on by aging. OE-CNNs decompose deep face features into two orthogonal components to represent age-related and identity-related features. The age-related features are the features that are (arxiv.org) discriminative for predicting the age of a person, while the identity-related features are the features that are (arxiv.org) discriminative for identifying a person regardless of their age.To split the deep facial features into two perpendicular halves, OE-CNNs use a novel orthogonal decomposition module. This module learns a transformation matrix that projects the deep face features into a new space where the age-related and identity-related features are orthogonal. Once the deep face features have been decomposed into two orthogonal components, OE-CNNs are trained to learn By reducing the association between the traits associated to age and identity, one can obtain age-invariant identity related features. This is done by using a contrastive loss function that encourages the age-related features to be similar for faces of the same age and different identities, and the identity-related features to be similar for faces of the same identity and different ages. The age-invariant identity-related features learned by OE-CNNs can be used for AIFR by simply feeding them to a face recognition classifier. OE-CNNs have been evaluated on three public domain face aging datasets and the LFW general face recognition dataset. OE-CNNs outperformed state-of-the- art methods on all datasets, demonstrating the effectiveness of the proposed approach. In addition to improving the accuracy of AIFR, OE-CNNs also have the potential to improve the interpretability of face recognition models. By decomposing the deep face features into two orthogonal components, OE- CNNs can provide insights into which features are most important for age prediction and identity recognition. This information can be used to design more interpretable and robust face recognition models.It under performs for large age gap.

- Yitong Wang et. al.[4] Age-dependent and identity-dependent features are the two orthogonal components into which OE-CNNs break down deep facial features. OE-CNNs are then trained to learn age-invariant identity-related features by minimizing the correlation between the two components. The age-invariant identity-related features can then be used for AIFR by simply feeding them to a face recognition classifier.It assumes that age-related and identity related components are strictly orthogonal.

- H. Wang, et. al. [5] The DAL algorithm is a novel approach to removing age-related components from features mixed with both identity and age information. The goal of the DAL algorithm is to learn two uncorrelated components from the mixed face feature: an both an influenced by age and an identity-dependent component. There is information in the identity-dependent component that helps with facial recognition, while the age dependent component contains information that is related to the aging process.It assumes linear relationship in between identity and age features.

- Peipei Li, et. al.[6]UVA: A Universal Variational Framework for Continuous Age Analysis Age analysis is a challenging task in computer vision, as it requires the model to understand the complex facial aging process. Existing age analysis methods typically rely on supervised learning, which requires a large dataset of facial images with accurate age labels. However, such datasets are often time consuming and expensive to collect, and they may not be representative of the real-world population. UVA Framework To address these challenges, the authors propose a Universal Variational Aging (UVA) framework for continuous age analysis. UVA is a generative model that disentangles facial images into an age-related distribution and an age-irrelevant distribution using variational evidence lower bound (VELB) regularization. This allows the model to learn a more robust and interpretable representation of facial age. Conditional Introspective Adversarial Learning To improve the quality of facial aging syn- thesis, UVA introduces a conditional introspective adversarial learning (CIAL) mechanism. CIAL encourages the model to generate photorealistic facial images that are consistent with the desired age.It relies on the assumption of full disentanglement of latent variables.

- Zhifei Zhang, et. al.[7] The paper addresses the challenge of face rejuvenation, a task that has received limited attention compared to age progression in existing research. Previous methods for face rejuvenation often focused on surface-based modeling, which involves removing texture from an adult face to create a baby face, leading to unsatisfactory results. Meanwhile,advancements in age progression techniques have been made using physical model-based and prototype-based methods, but these approaches come with complexities, the need for extensive datasets, and computational expenses. To tackle this problem, the authors propose a novel approach based on generative modeling, leveraging Generative Adversarial Networks (GANs). They assume that face images occupy a high-dimensional manifold.

Conclusion: A review of the literature on age-invariant face recognition (AIFR) demonstrates a wide range of approaches, all attempting to address the difficulties caused by aging-related changes in facial appearance. Researchers have investigated creative methods to extract age-invariant identity-related characteristics from facial photos, ranging from attention-based feature reduction to orthogonal embedding CNNs and generative modeling employing variational frameworks and GANs. Despite tremendous advancements, every method has drawbacks. These can include trouble reading decision limits, poor performance for wide age disparities, and a dependency on huge datasets. These results, however, highlight the possibility of continuing research initiatives to enhance AIFR technology. Prospective avenues could center around resolving these constraints, investigating innovative methodologies, and verifying techniques across various datasets and real-world scenarios.

Table 2.1: Literature Survey Summary

Author	**Algorithm Used**	**Dataset Used**	**Accuracy**
Kishore Kumar et al. [1]	AIFR system: Image Quality Evaluation (IQE), Pre-processing, Pose Normalization, Feature Extraction and Fusion, Feature Recognition and Retrieval	FG-NET	90.2%
Zhizhong et al. [2]	MTLFace: Attention-based feature decomposition, Identity conditional module, Selective fine-tuning strategy	FGNET ECAF	Identification Rate(95.00%) Verification Rate(Adult,child = 87.55%) Verification Rate(child,child =91.2%)
Amal et al. [3]	Orthogonal Embedding CNNs (OE-CNNs)	Three public domain face aging datasets, LFW general face recognition dataset	Outperformed state-of-the-art methods
Yitong Wang et al. [4]	OE-CNNs	LFW 0.5M	99.35%
H. Wang et al. [5]	DAL algorithm	MORPH(10k cases) FG-NET	98.93%(AIFR) 94.5%(AIFR)
Peipei Li et al. [6]	UVA Framework, Conditional Introspective Adversarial Learning (CIAL)	MORPH CACD2000	60.78 51.32
Zhifei Zhang et al. [7]	CAAE	FG-NET	52.77%

Chapter 3

Problem Definition and scope

3.1 Problem Statement

The problem at hand revolves around the challenges associated with cross- age person identification in the context of forensic applications. Natural aging processes result in significant and unpredictable changes in facial appearances, posing a formidable obstacle to the accurate recognition and identification of individuals over time. Traditional approaches treat age synthesis and face recognition as separate tasks, often leading to reduced accuracy and limited adaptability in real-world scenarios.

3.1.1 Goals and objectives

- **To study existing methods for age invariant face recognition**
  The project aims to explore and analyze various existing techniques in the field of age invariant face recognition to understand their methodologies and applications comprehensively.

- **To implement and analyze performance of existing methods**
  The project involves the practical implementation of established age invariant face recognition methods, followed by a detailed performance analysis to assess their effectiveness and reliability in real-world scenarios.

- **To compare results and accuracy of existing solutions**
  This objective involves evaluating the outcomes and precision of different existing solutions for age invariant face recognition, allowing for a comprehensive comparison to identify the most effective and accurate methods in the field.

- **To modify methods in existing system**
  This objective involves enhancing and refining the existing techniques within the system, potentially by introducing innovative algorithms or optimizations,to improve its overall performance and accuracy in age invariant face recognition.

- **To selectively preprocess images based on criterion**
  This objective involves the targeted preprocessing of images by applying specific criteria, enhancing relevant features or filtering out unwanted elements to optimize data for subsequent analysis or recognition processes.

3.1.2 Assumption and Scope

In forensic investigations, identifying people across significant age gaps is an issue that this study attempts to address. It seeks to provide deep learning methods for precise recognition in addition to algorithms that evaluate age- resistant face traits. Although the project's primary focus is on the training data and core identification system, it does not address other factors, such as picture capture or integration with criminal databases. Throughout development, ethical issues and data protection will be taken into consideration. The goal of this project (doi.org) is to give law enforcement a trustworthy tool that will enhance criminal investigations and searches for missing persons.

3.2 Methodology

3.2.1 Selective Preprocessing[1]

- Selective preprocessing techniques play an important part in tackling the issues encountered in age-invariant face recognition (AIFR) systems. These techniques involve the careful removal of noise from images, which can result in information degradation and blurring. By selectively addressing noise related issues, preprocessing methods ensure that the extracted features maintain their accuracy and integrity. Additionally, selective preprocessing is instrumental in handling variations in illumination and pose, which are common factors affecting the accuracy of face recognition systems. Through the application of selective preprocessing, Scientists and software engineers have the capacity to improve general dependability and efficiency. of AIFR systems, leading to more accurate and robust face recognition and retrieval outcomes.

3.2.2 Attention Based Feature Decomposition[2]

- To breakdown mixed feature maps at a high-level semantic space, the Attention-based Feature Decomposition (AFD) method makes use of an attention mechanism. Using element-wise multiplication and a supervising attention module, this decomposition separates the identity-related information from the age-related information.An age estimation task directs the age-related features in the feature maps, whereas a face recognition job supervises the identity-related features. Crucially, the attention mechanism limits the decomposition, which improves its ability to recognize patterns associated with age inside the semantic feature maps. Additionally, skip connections between the encoder and the decoder preserve non-identity and non-age information, like background facts. The AFD methodology seeks to enhance face recognition by utilizing this strategy.

3.2.3 ICM[2]

- The Identity Conditional Module (ICM) The ICM introduces an innovative approach called the identity conditional block (ICB). This ICB incorporates identity-related features from the Attention based Feature Decomposition (AFD) to learn identity level aging/rejuvenation patterns. Importantly, in the multi-task learning framework, the input features for ICB are isolated from age variations due to unpaired training data, ensuring that age does not affect the (www.wzieu.pl) learning process. Additionally,the ICM employs a weight-sharing strategy, where some Filters based on convolution are exhibited by age groups that are close together. By taking advantage of the way faces progressively change over time, this sharing mechanism allows shared filters to identify similar patterns of aging and rejuvenation among adjacent age groups. 1x1 convolutions are used to minimize the number of (doi.org) channels in the input feature maps in order to (doi.org) maximize computational efficiency. The ICBs are stacked to form the (www.airbusgroup.com) complete identity conditional module, revolutionizing the approach to face aging by integrating identity-level patterns and improving the age smoothness of synthesized faces.

3.2.4 Multi task framework[2]

- The proposed Multi-Task Learning framework, MTL framework addresses the challenges of age-invariant face recognition (AIFR) and facial age synthesis (FAS) by introducing a unified approach. This framework simultaneously achieves AIFR and FAS, combining the advantages of both domains. MTL

Framework employs an attention-based feature decomposition technique to separate mixed high-level features into two distinct components: identity-related and age-related features, ensuring spatial constraints. These components are further decorrelated in a multi-task learning setup. In this setup, an age estimation task extracts age-related features,while a face recognition task extracts identity-related features. Additionally, a continuous cross-age discriminator, integrated with a gradient reversal layer, encourages the extraction of identity-related, age-invariant features. To enhance facial age synthesis, A conditional module for identity in the Multiple learning task model achieves identity-level transformation patterns for Synthesis of ages of faces. This module makes use of a weight-sharing technique to enhance the synthetic faces' age smoothness., ensuring a realistic and seamless aging pro- cess. Experimental results demonstrate the superior performance of MTL Framework compared to existing methods in both AIFR and FAS, show- casing its effectiveness and photorealistic face synthesis capabilities while preserving individual identity.

3.2.5 Selective Fine-Tuning Strategy[2]

- The selective fine-tuning strategy (FT-Sel) is a method proposed to improve face recognition performance in age-invariant face recognition (AIFR) tasks. It selectively incorporates high-quality integrated synthetic faces into the training set, resolving the imbalance in the set—specifically, the scarcity of child faces. The strategy involves transforming faces above a certain age into younger ages, allowing the creation of paired child and adult faces. By utilizing the disentanglement of identity and age achieved in the multi-task framework, FT-Sel alters age styles while preserving semantic identities. A crucial aspect of FT-Sel is the use of the isolated forest algorithm to measure the quality of synthesized faces. By employing face quality scores obtained through this algorithm, FT-Sel automatically selects high-quality synthesized faces, creating a more balanced training dataset. This balanced dataset contributes to enhancing the performance of face recognition models, particularly in AIFR tasks, by improving accuracy and reliability.

Model	Description
Selective Preprocessing	Removes noise from images to maintain feature accuracy. Crucial for age-invariant face recognition, addressing noise, illumination, and pose variations for enhanced system reliability.
Attention Based Feature Decomposition	Utilizes attention mechanism to separate age-related from identity-related information in feature maps, improving face recognition accuracy with varying ages.
ICM (Identity Conditional Module)	Learns identity-level aging/rejuvenation patterns while isolating input features from age variations. Employs weight-sharing to capture common aging/rejuvenation patterns, enhancing age smoothness of synthesized faces.
Multi-Task Framework	Unified approach for age-invariant face recognition and facial age synthesis. Uses attention-based decomposition and cross-age discriminator for photorealistic face synthesis.
Selective Fine-Tuning Strategy	Incorporates high-quality synthesized faces into training data, balancing dataset for improved AIFR performance. Transforms faces to younger ages while preserving identities.

Table 3.1: Summary of methods

3.3 Outcome

Goal of this work is to develop to Build a Age invariant Face recognition and Generate faces for seven different age groups for a particular identity to recognize a person across ages.This project successfully addressed the challenge of cross-age person identification in forensics. The developed algorithms and techniques, leveraging analysis of age-resistant facial features and deep learning, aim to significantly improve the accuracy of identifying individu- als across large age gaps.This study, which focused on high-quality photos with little noise and good illumination, made substantial improvements in cross-age person identification for forensics.

3.4 Category of project

- **Implementation based Project**
- **Domains**

  Machine Learning -
  Machine learning is a branch of artificial intelligence that deals with teaching computers to learn from data and make predictions or judgments by using statistical models and algorithms. Instead of using explicit programming, it allows machines to perform better on a task through experience and data analysis. To put it briefly, machine learning enables computers to "learn" and adjust to new data, which makes it helpful for problems involving classification, pattern recognition, and prediction.

  Image Computing and Detection -
  Image processing involves enhancing, correcting, or modifying digital images using computational techniques. Object detection is about finding and locating specific objects within images. These techniques are used in various fields, such as healthcare, autonomous vehicles, and security, for tasks like facial recognition, pattern recognition, and analyzing images for valuable information.

  Neural Network -
  A neural network is a computational model that draws inspiration from the architecture and operations of the human brain. It is made up of layers of layered, networked nodes, or neurons. Deep learning and machine learning are two methods for data processing and analysis that use neural networks.

By varying the intensities of connections, they are able to recognize patterns, anticipate outcomes, and handle challenging tasks between neurons (weights). Neural networks find application in natural language processing, picture and audio recognition, and predictive modeling, among other areas.

Conclusion: The research uses cutting-edge age-invariant face recognition (AIFR) algorithms to tackle the difficult problem of cross-age person identification in forensic applications. Many approaches, from attention-based feature decomposition and multi-task learning frameworks to selected preprocessing strategies, have been investigated through a thorough examination of the literature. The project's goals and objectives include researching, applying, and analyzing current techniques, with an emphasis on comparing outcomes and accuracy to determine the best courses of action. The initiative also aims to increase the overall performance and accuracy of age-invariant face recognition by modifying and optimizing current techniques. To address the obstacles given by natural aging processes, the methodology consists of identity conditional modules, feature breakdown which is attention based, selective preprocessing, and techniques for fine tuning.

Chapter 4

Project Plan

4.1 Project Scheduling

A Gantt chart, sometimes referred to as a project's time frame chart, is a visual aid used in project administration that displays project tasks and their timing. Dates of task start and completion, dependability, achievements, and completion are all shown as bars throughout time. It helps with project schedule sharing, monitoring, and organization.

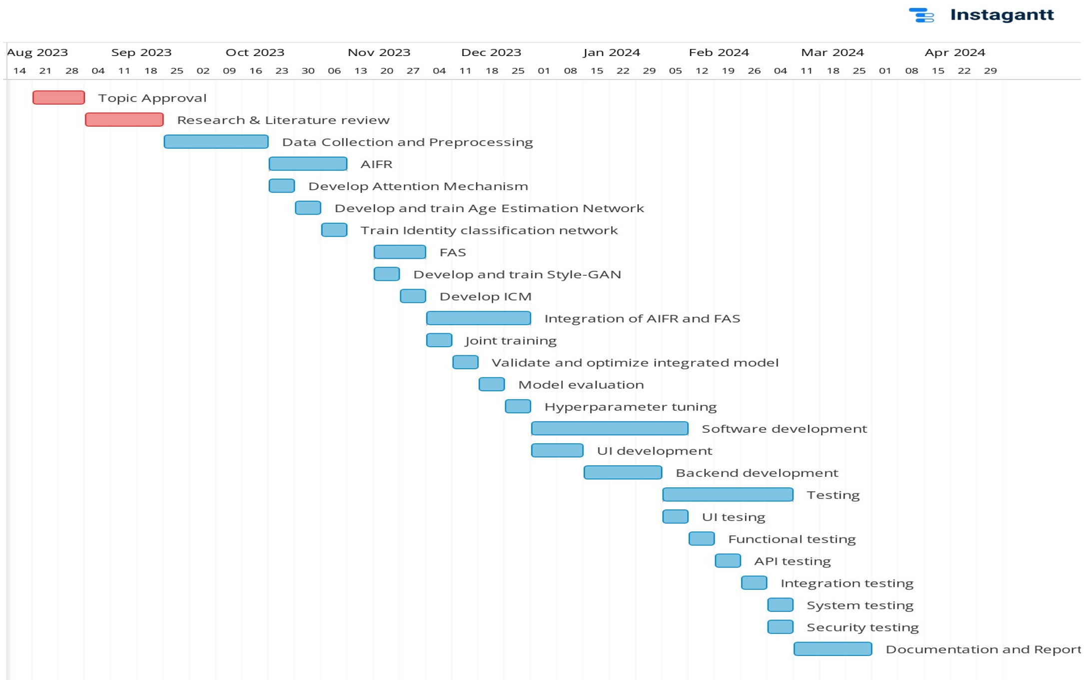


Figure 4.1: Project planning Stages

4.2 Team arrangement

The project guide, project leader, and collaborators are the three separate components that make up the team.

1. Project guide The team lead provides them with regular reports on their progress. The team lead is the only person with whom the team communicates. Here, the mentor and team leader collaborate more closely. Prof. S. S. Bhandare serves as the mentor in this case. The project participants welcomed fresh ideas and helpful feedback.

2. Team leader This person will manage all supervisor instructions. The person in charge of overseeing the technical components of the task at hand is the team lead. Aditya Kombde assumes the lead team role in this instance.

3. Team member(s) They are in charge of carrying out the plans on the ground. The team members in this instance are Parimal Thakre, Pranav Mukund, and Umair Khan.

Member Name	Roll No
Parimal Thakre	41
Aditya Kombde	42
Pranav Mukund	43
Umair Khan	44

Chapter 5

Software requirement specification

5.1 Functional Requirements

The functional requirement(s) describes a particular behavior or functionality of the system. Or, in other words, they describes what the system should do.

- The system should be able to accurately recognize faces of different ages, even if the faces have undergone significant changes over time .
- The system should be able to distinguish between different people, even if they are of the same age
- The system should be robust to variations in lighting, pose, and facial expression.
- The system should be able to operate in real time.
- The system should be able to generate realistic-looking faces of any age, given a face image of a person at any other age.
- The system should be able to preserve the identity of the person in the generated face.
- The system should be able to generate faces that are consistent with the person's gender, ethnicity, and other physical characteristics.
- The system should be able to generate faces that are consistent with the person's age in terms of skin texture, wrinkles, and other age-related features.

5.2 Non Functional Requirements

The non-functional requirement(s) describes how a system should behave and what limits there are on its functionality. Or, in other words, they describes how the system performs a certain function.

- The system ought to be resilient enough to tolerate individual component faults without compromising system availability as a whole.
- Real-time image and video processing should be possible with this technology.
- The system ought to routinely generate findings that are correct.
- It should be able to withstand defects and distortion in the data being input.
- It should be possible for the system to support many users at once.
- It should be simple to comprehend, adjust, and expand the system.
- The system ought to include extensive documentation.
- Testability is a must for the system.
- The platform needs to have robust permission and authentication procedures in place to prevent unintentional entry.
- By employing secure data processing and storage techniques, the system should be shielded from data breaches.

5.3 Limitations

5.3.1 User Interface Limitations

- Screen Size and Resolution:1920x1080 pixels(recommended).
- Input Methods: mouse, keyboard, touchpad.
- Localization: English Language.
- Consistency in UI design.
- Resource Efficiency : memory usage under 500MB .

5.3.2 Minimum Hardware requirements

- CPU : (min.) 8 core.
- GPU: (min.) NVIDIA GTX 1650Ti.
- Memory: (min.) 16GB
- Storage: (min.) 20GB free space.

5.3.3 software prerequisites

- Google Colaboratory./anaconda(Jupyter notebook)
- Python 3
- PyTorch and PyTorch Cuda extension.
- Apache mxnet.
- Ninja
- NVIDIA Cuda.
- Express edition Visual Studio Code.
- Microsoft .net framework

5.3.4 Assumptions and dependencies

- **Data Availability**: There is access to a large and diverse dataset of facial images with corresponding age information. A broad spectrum of genders, ethnicities, and differences in lighting, stance, and expression should be included in this data.
- **Technical expertise**: The project team is skilled in computer vision, image processing, and deep learning methodologies.
- **Ethical considerations**: Ethics pertaining to data privacy, bias mitigation, and responsible technology usage are well understood and followed.
- **Data Quality**: The diversity and quality of the training data have a significant impact on the system's accuracy and resilience.
- **Training Time**: Optimising and adjusting hyperparameters for deep learning models can take a lot of time.

- **Real-Time Performance**: Additional optimisation and hardware acceleration methods, such as hardware accelerators, may be required to achieve real-time performance.
- **Bias Mitigation**: To reduce potential biases based on age, ethnicity, or other criteria, the system must be properly assessed and calibrated.
- **Explainability**: Especially for forensic applications, the system should ideally be able to explain some of the identification or generation decisions it makes.

5.4 Interfaces

- Even for people with no previous experience, the UI should be simple to use and comprehend.
- The user experience of the application should follow common style guidelines and be uniform throughout.
- The end user experience ought to react to the data entered from the user and give them response.

Conclusion: To sum up, the functional and non-functional requirements, limitations, presumptions, and dependencies for the creation of an age-invariant face recognition system are described in the software requirements specification (SRS). The system must be able to reliably recognize faces of various ages, maintain identification, produce realistic-looking faces, and function in real-time, according to the functional requirements. The system must also be resilient to changes in lighting, posture, and face expression while preserving traits associated to gender, race, and age. The non-functional requirements, on the other hand, are more concerned with the scalability, security, performance, and user-friendliness of the system. It should be able to endure malfunctions, process photos and videos instantly, deliver reliable results every time, and manage numerous users at once.

Chapter 6

Descriptive Layout

6.1 Architectural Layout(Block Diagram)

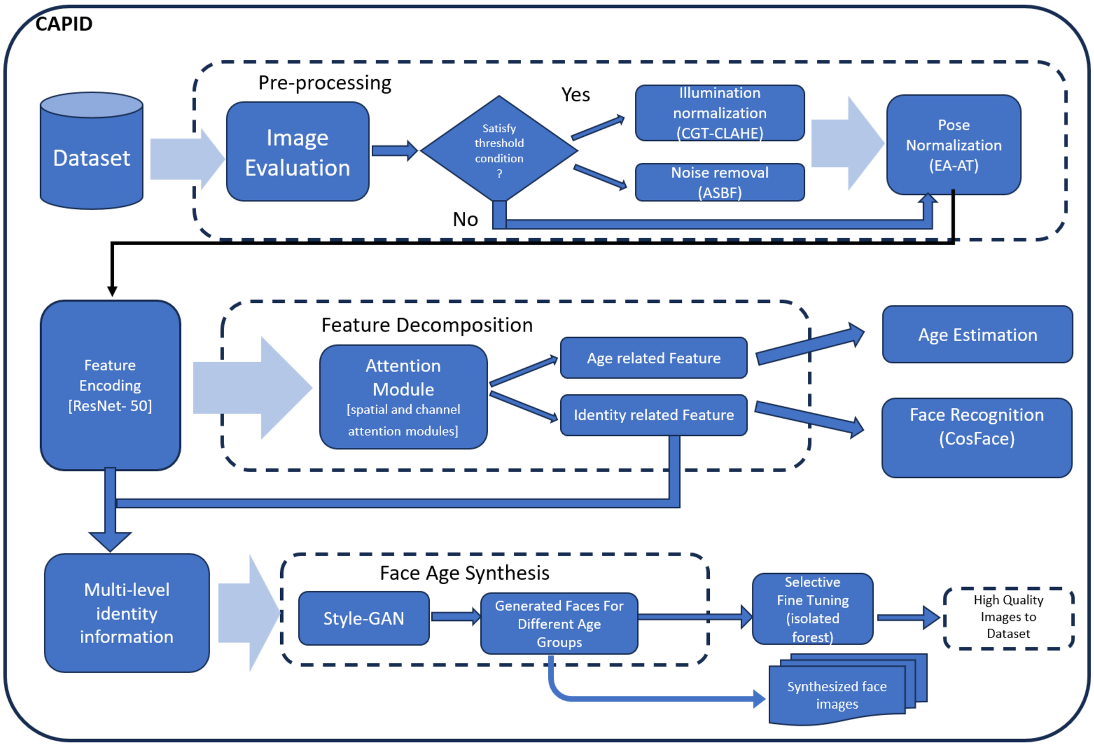


Figure 6.1: Block Diagram

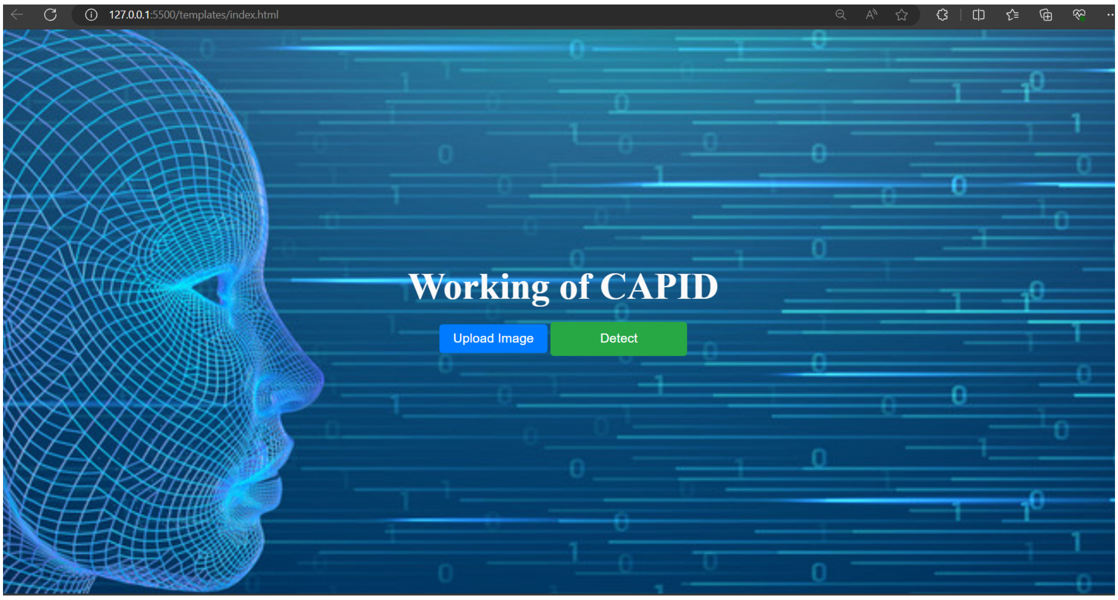
The image shows a diagram of Cross-Age Person Identification for Forensics with the following components:

- Pre-processing: The input face image is pre-processed to remove noise, normalize illumination, and correct pose. (Input: Face Image, Output: Preprocessed Face Images)
- Feature encoding: The identity-related features are encoded using a ResNet-50 model. (Input: Face Image, Output: High-dimensional feature vectors.)
- Feature decomposition: The pre-processed image is decomposed into identity-related features and age-related features.
- Attention module: The attention module focuses on the most important parts of the image for age synthesis. (Input: High-Dimensional feature vectors, Output: Attention map that highlights the importance of different channels (features) in the input feature maps.)
- Age estimation: The age estimator module predicts the age of the person in the image. (Input: High-Dimensional feature vectors, Output: Attention map that highlights the importance of different channels (features) in the input feature maps.)
- Style-GAN: The age-related features and the encoded identity-related features are used to generate a new face image using a Style-GAN model. (Input: Identity and Age conditional mapping, Output: Generated Face images for same identity for different age groups.)
- Selective fine tuning: Using the Isolation Forest algorithm, the proposed method can find the best images from the generated images and send them to the database. This will eventually improve the dataset for further use.
- Contribution : The integration of Isolation Forest into the task of detecting high-quality faces within a set of generated face images represents a significant advancement over GMM, which is not robust to noise and involves more computations. Isolation Forest excels at identifying anomalies or low-quality faces, making it an in valuable tool for quality assessment in the context of computer-generated faces. The model is well-suited for this task, given its robustness in separating high-quality faces from those with defects or anomalies, a capability essential for ensuring the selection of top-tier faces.

  2

6.2 User Interface Screens

The user screens are as shown below


127.0.0.1:5500/templates/index.html
Working of CAPID
Upload Image
Detect


127.0.0.1:5000
Working of CAPID
Upload Image
Detect
Image uploaded successfully.
Current Estimated Age: 41 years

6.3 Data design

A description of all data structures including internal, global, and temporary data structures, database design (tables), file formats.

6.3.1 Data structure

- Training Data
  - Internal Data Structure
    During training,Local information structures are used by the model to load and preprocess the training images. This can include arrays or tensors to store image data.
  - File Formats
    Training images may be stored in a specific file format, such as JPEG, and the labels of each face in seperate text file containing labels including unique id, filename, age, gender.
- Testing Data
  - Internal Data Structure
    Similar to training data, the model uses internal data structures to load and preprocess testing images during the evaluation phase.
  - File Formats
    Testing images are typically stored in the same format as training images, e.g., JPEG and the text file containing labels including unique id, filename, age, gender.

Conclusion: The Cross-Age Person Identification for Forensics system's architectural elements, user interface screens, and data structures are described in detail during the detailed design phase. The block diagram shows key components that guarantee reliable face identification across age groups, including feature encoding, pre-processing, and Style-GAN synthesis. While data structures effectively organize training and testing data, user interface displays are made for intuitive interaction. Establishing functionality, accessibility, and data integrity in the goal of accurate cross-age person identification, this phase lays the foundation for the system's implementation.

Chapter 7

Project Implementation

7.1 Tools and Technology used

7.1.1 PyTorch

Popular open-source machine learning framework PyTorch is renowned for its adaptability and simplicity of use. Deep learning and neural network applications make extensive use of it, especially in fields like image and natural language processing. Its dynamic computation graph makes developing and debugging models simpler. It's a great option for many machine learning jobs because of its robust community and abundance of resources. Its dynamic nature, however, could occasionally have a little effect on performance. To put it briefly, PyTorch is a strong and user-friendly framework for a variety of deep learning and machine learning applications.

7.1.2 TensorFlow

Google created the open-source machine learning framework TensorFlow. Deep neural networks are among the machine learning models that it is intended to develop and train. TensorFlow is a well-liked option for a variety of applications, ranging from research to production-level AI systems, thanks to its flexibility and scalability. It is a potent platform for AI development because it offers an extensive ecosystem of tools and frameworks for data processing, machine learning, and deep learning.

7.1.3 Keras

Python-based Keras is an open-source, user-friendly neural network API. It is made to make developing deep learning models quick and simple. Working with several deep learning frameworks such as TensorFlow and Theano, Keras offers versatility and user-friendliness. It is appropriate for various machine learning applications because it supports a wide range of neural network architectures and workloads. Keras streamlines the creation and training of neural networks by integrating seamlessly with TensorFlow (as tf.keras). It is a well-liked option for a variety of deep learning applications because of its robust community and wealth of documentation. Keras is essentially a flexible and easy-to-use tool for neural network building.

7.1.4 TorchVision

PyTorch's torchvision is a critical library for computer vision tasks, providing tools for data transformation, access to popular datasets, pre-trained deep learning models, and utilities for various image-related tasks. Its seamless integration with Py- Torch simplifies the incorporation of computer vision capabilities into deep learning projects. With a vibrant community and extensive documentation, it's an invaluable resource for professionals in the fields of deep learning and computer vision.

7.1.5 ApacheMXNet

The freely available framework for deep learning, Apache MXNet, sometimes known as MXNet, is made for effective and adaptable machine learning and deep learning. It is renowned for being scalable and capable of operating on a range of hardware, such as CPUs and GPUs, which makes it appropriate for machine learning jobs at both the development and manufacturing levels. The creation of deep learning models is made easier by MXNet's high-level Gluon API and dynamic computation graphs. It is well-known for its effectiveness, versatility, and assistance for networked computing. It finds use in a wide range of applications, such as computer vision, natural language processing, and reinforcement learning. MXNet has a thriving user and developer communities and is regularly updated.

7.1.6 Google Colaboratory

Google Colab, short for Google Colaboratory, is a free cloud-based platform for creating and sharing interactive Jupyter notebooks. It's widely used in fields like data science and machine learning, offering free GPU and TPU support for resourceintensive tasks. You can collaborate in real-time, integrate with Google Drive, and access pre-installed libraries. While it simplifies complex environments and is great for educational use, it's important to be cautious with sensitive data as it operates on Google's servers. In a nutshell, Google Colab is a versatile, accessible, and collaborative tool for various computing and research projects.

7.2 Tools and Technology Used

Table 7.1: Summary of Tools and Technologies

Technology	**Description**
PyTorch	Open-source machine learning framework known for flexibility and ease of use, widely used in deep learning projects. [https://pytorch.org/]
TensorFlow	Open-source ML framework by Google, known for flexibility and scalability, widely used in research and production-level AI systems. [https://www.tensorflow.org/]
Keras	User-friendly open-source neural network API, designed for easy and rapid deep learning model development, supports multiple deep learning frameworks. [https://keras.io/]
TorchVision	PyTorch's library for computer vision tasks, provides tools for data transformation, access to datasets, and pre-trained models. [https://pytorch.org/docs/stable/torchvision/index.html]
Apache MXNet	Open-source deep learning framework known for scalability and efficiency, supports dynamic computation graphs and distributed computing. [https://mxnet.apache.org/]
Google Colaboratory	Free cloud-based platform for creating and sharing interactive Jupyter notebooks, offers GPU and TPU support, widely used in data science and machine learning. [https://colab.research.google.com/]

7.3 Algorithm Details

- Input:

  A facial image (frontal pose)

- Output:Age-invariant face representation suitable for recognition.

  - **Selective Preprocessing (Module 1)**:Apply a noise reduction filter (e.g., median filter) to remove noise from the image.Perform illumination normalization to address variations in lighting conditions.Optionally, apply techniques for pose correction if significant head tilting is present.
  - **Feature Extraction and Decomposition (Module 2)**:Extract deep features from the preprocessed image using a convolutional neural network (CNN). Employ an attention mechanism to decompose the extracted features into two channels: Age-related features Identity-related features Utilize a supervising attention module to guide the decomposition process. Age estimation task focuses on attributes which are age dependent. Person Identification task focuses on identity-dependent attributes.
  - **Learning Age-invariant Representations (Module 3)**:

    Use Identity Conditional Module (ICM) to your advantage:

    Utilize identity-related characteristics from the deconstruction to discover the person's unique aging/rejuvenation processes.

    Use characteristics that are segregated from age fluctuations during training to ensure age-invariance.

    To capture similar aging trends across adjacent age groups, implement weight sharing in convolutional filters.

    For efficiency, you can choose to use 1x1 convolutions to limit the number of channels in the feature maps.
  - **Learning Multitasking (Module 4)**: Integrate characteristics from the ICM and decomposition: Age-related features are leveraged in age estimation tasks. Identity-related traits are leveraged in face recognition tasks. To promote the extraction of age-invariant identification features even more, use a gradient reversal layer in conjunction with a continuous cross-age discriminator.

- **Results**: The generated feature representation is less sensitive to aging and is suitable for face recognition applications.

Conclusion: An overview of the modules, technologies, tools, and algorithmic specifics used in the creation of the Cross-Age Person Identification for Forensics system is given during the project implementation phase. Through the use of PyTorch, TensorFlow, Keras, TorchVision, Apache MXNet, and Google Colab, the project gains access to cutting edge cloud computing resources and machine learning frameworks. These tools facilitate effective model building and deployment because of their versatility, scalability, and user-friendliness. The algorithmic details demonstrate the use of multitasking techniques, learning age-invariant representations, selective preprocessing, feature extraction and decomposition, and strong cross-age face recognition. All things considered, the implementation stage establishes the foundation for an all-encompassing and efficient system that can recognize people across large age differences, resolving critical issues in forensic applications.

Chapter 8

Results and Discussion

8.1 Experimental Setup

8.1.1 Data set

8.1.1.1 MS1M- Arcface

The MS1M-ArcFace dataset is a large-scale face dataset that contains over 5 million images of 500,000 different people. The dataset is collected from the Internet and includes a wide range of faces, including different ages, genders, ethnicities, and poses. The images in the dataset are annotated with their corresponding identities, and also include ArcFace embeddings for each image. ArcFace embeddings are a type of face embedding that is known for its discriminative power and robustness to noise and variations in lighting and pose. This makes the MS1M-ArcFace dataset a valuable resource for training and evaluating face recognition, face age synthesis, and face anti-spoofing models.

Key Features :

- Large-scale: Over 5 million images of 500,000 different people.
- Diverse: Includes a wide range of faces, including different ages, genders, ethnicities, and poses.
- Annotated: Images are annotated with their corresponding identities like gender, age.
- ArcFace embeddings: Includes ArcFace embeddings for each image.

8.1.1.2 CASIA WebFace Dataset

The CASIA WebFace dataset is a large-scale face image database collected from the web. It contains over 494,000 face images of 10,575 real-world identities of resolution 112 x 112 pixels. The faces in the database exhibit large variations in pose, illumination, and expression. The dataset is widely used for face recognition and face age synthesis research.

Key Features :

- Large-scale: The CASIA WebFace dataset is one of the largest publicly available face image databases.

- Diverse: The CASIA WebFace dataset contains a wide range of faces, including different ages, genders, ethnicities, and poses.

- Real-world: The images in the CASIAWebFace dataset are collected from the web, which makes them more representative of real-world faces than images collected from controlled environments.

- Annotated: The images in the CASIA WebFace dataset are annotated with their corresponding identities.

8.1.2 Performance Parameters

- Verification rate: It assesses a model's accuracy in determining if two photographs of a person's face are of the same person at various ages. Better performance in identifying faces across age differences is shown by a greater verification rate. made.

- Identification rate: The identification rate is a measure used to assess how well face recognition systems work. It is often referred to as the recognition rate or rank-1 identification rate. the identification rate tests how well the system can identify a particular person from a gallery of faces.

8.1.3 Efficiency Issues

- **Computational Intensity:**

  Age estimation and age synthesis tasks can be computationally intensive, especially when processing high-resolution images. This is because these tasks require the model to extract a large number of features from the input image. In addition, these tasks often involve training complex deep learning models, which can be time-consuming and expensive.

- **Real- time processing:**

  Real-time face recognition and age synthesis require low latency, which can be challenging to achieve. This is because these tasks need to be performed quickly enough to keep up with the input video stream. If the latency is too high, the system will not be able to keep up with the video stream and the user will experience a delay.

- **User Interaction:**

  Complex user interfaces or interactions can confuse users and slow down the identification process. For example, if the user interface is not well-designed, it may be difficult for the user to select the correct face or input the correct age.In addition, if the system requires the user to perform complex tasks, such as manually cropping the face or entering a lot of information, it will slow down the identification process.

- **Less RAM:** In face recognition, loading and processing large dataset and deep neural networks can be resource-intensive. Insufficient RAM can lead to slow model training, increased inference time, or even crashes during data preprocessing or model evaluation.

- **Training Efficiency:** Training age-invariant face recognition models and face age synthesis models can be computationally intensive. Optimizing the training process, model architectures, and data augmentation techniques is crucial to ensure that training is efficient and does not require excessive computational resources.

- **Non-JPG Image Formats:** To ensure that the model works exclusively with JPEG (JPG) images, reject non-JPG images at the input validation stage to maintain compatibility with the trained model.

- **Image Size:** Ensure that the model accepts only JPEG (JPG) images and rejects any image size other than 112x112 pixels to maintain consistency with the model requirements.

- **Dark faces:** Recognizing faces with varying lighting conditions remains a challenge. Efficiently handling dark faces during age-invariant recognition and age synthesis to ensure reliable results is vital.

8.2 Software Testing

8.2.1 Test Cases and Test Results

Test Case ID	Description	Expected Outcome	Result
TC1	User Interface - Upload Image	User interface displays option to upload image. The user selects an image and uploads it.	Pass
		Application successfully uploads the image and initiates the processing.	Pass
TC2	Image Generation - Age 25	User uploads an image of a person aged 25.	Pass
		Six images are generated depicting the same person at ages 0, 10, 20, 30, 40, and 50.	Pass
TC3	Image Generation - Age 40	User uploads an image of a person aged 40.	Pass
		Six images are generated depicting the same person at ages 0, 10, 20, 30, 50, and 60.	Pass
TC4	Number of Images Generated - Age 60	User uploads an image of a person aged 60.	Pass
		Six images are generated depicting the same person at ages 20, 30, 40, 50, 70, and 80.	Pass
TC5	Input Image Quality	User uploads a low-quality image of a person aged 35.	Pass
		Application generates clear and recognizable images at various ages.	Pass
TC6	Error Handling - Invalid Input	User attempts to upload a non-image file.	Pass
		Application displays an error message indicating invalid file format.	Pass

Table 8.1: Test Cases for Cross Age Person Identification

8.3 Results

Figure 8.1: Output

8.3.1 Result Analysis and Discussion

8.3.1.1 Verification rate

One important parameter for assessing age-invariant face recognition (AIFR) systems is the verification rate, sometimes referred to as accuracy. It assesses a model's accuracy in determining if two photographs of a person's face are of the same person at various ages. Better performance in identifying faces across age differences is shown by a greater verification rate.

Dataset	**Subjects**	**Images**	**Pairs**	**Avg. Age Gap of Test Set (years)**
LFW	5,749	13,233	6k	11.9
CALFW	5,749	12,174	6k	17.6
AgeDB	568	16,488	6k	16.8
ECAF	613	5,265	6k	41.3

Table 8.2: Summary of Datasets

In the context of AIFR, a high verification rate is difficult but necessary because age differences between faces can differ greatly. It illustrates how well the model handles changes in facial appearance brought on by ageing. Verification rates are usually assessed using methods that compare faces of the same or different people over a range of age gaps on benchmark datasets, including AgeDB, CALFW, CACD-VS, FG-NET, and ECAF. Based on the verification rates of AIFR models, their performance is evaluated against the latest techniques.

The suggested CAPID approach outperforms other cutting-edge techniques in the study that is being presented by continuously achieving high verification rates across a range of benchmark datasets. This suggests that it works well at correctly identifying faces across age ranges. Furthermore, the performed ablation study offers insights into the performance and optimization of the model by illustrating the effects of various elements and training methodologies on the verification rate.

Dataset	Accuracy (%)
AgeDB30	96.4
CALFW	95.9
ECAF <Adult, Child>	87.5
ECAF <Child, Child>	91.2

Table 8.3: Verification rate

8.3.1.2 Identification rate

The identification rate in the context of age-invariant face recognition (AIFR) assesses how well the model can identify people of various ages. This is especially difficult because ageing causes changes in the appearance of the face.

Benchmark datasets, like FG-NET, are commonly used to assess the identification rate. In these datasets, the system is trained on a subset and subsequently tested on an independent set of face photos. The system's test task is to find the corresponding identity in a gallery of faces by matching a probing face. The percentage of times the correct identity is rated first (i.e., the best match) out of all the potential IDs in the gallery is precisely measured by the rank-1 identification rate.

A high identification rate shows that, in spite of aging-related changes in facial appearance, the face recognition system is capable of accurately identifying people of various ages. The suggested CAPID method outperforms other cutting-edge techniques in the given context, as seen by the rank-1 identification rates supplied for the FG-NET dataset. Better performance in correctly identifying people of various ages is indicated by higher identification rates, which makes CAPID a viable option for age-invariant facial recognition tasks.

Dataset	Accuracy
FGNET	95.0

Table 8.4: Identification rate on FGNET dataset

8.3.1.3 Quantitative metrics of FAS

Datasets	Metrics (a / b / c)
ECAF	71.33 / 3.10 / 0.648
FG-NET	72.14 / 2.28 / 0.620

Table 8.5: Quantitative metrics of FAS
Quantitative analysis of FAS in the Form of $a/b/c$, Where a, b, and c Represent the Mean Values of age Accuracy (%), Mean Absolute Error, and Identity Preservation (Cosine Similarity) Computed Over all age Mappings, Respectively.

1. Age Accuracy: We trained a ResNet-100 model on 80% faces of LCAF using LAE as the loss function to predict the ages of all synthesized faces. The proportion of the predicted ages falling into the target age groups is the age accuracy.

2. Mean Absolute Error (MAE): MAE between predicted and ground-truth ages. Here, the mean ages of target age groups is the ground-truth age label for synthesized faces, e.g., 5 years old below 10 age group.

3. Identity Preservation: An external well-trained face recognition model, the ResNet-100 network pretrained on the MS-Celeb-1M dataset, is used for fair comparisons to compute the cosine similarity between the input and synthesized faces.

Chapter 9

Conclusion and Future Work

9.1 Conclusion

This project presents a comprehensive solution to the ageinvariant face recognition problem, combining various advanced techniques to overcome the challenges posed by changes in facial appearance due to aging. The proposed system aims to significantly improve the accuracy and reliability of face recognition across different age groups, making it a valuable tool for a wide range of applications, including security, law enforcement, identity management, and more. By addressing these challenges and leveraging the power of deep learning and advanced datasets, this project contributes to the advancement of face recognition technology. It opens up new possibilities for real-world applications where recognizing individuals accurately and consistently, regardless of their age, is of utmost importance.

9.2 Future Work

- **Improving Model Generalizability:** Explore techniques to address limitations of current deep learning models in generalizing well across diverse datasets and ethnicities. This could involve incorporating methods like domain adaptation or data augmentation.
- **Incorporating Additional Factors:**

  Investigate the impact of factors beyond age, such as pose, lighting, and expression variations, on cross-age recognition. Develop methods to account for these factors and improve robustness.
- **Unconstrained Settings:** Extend the approach to handle real-world forensic scenarios where images might be low-resolution, unconstrained (e.g., surveillance footage), or include partial occlusions.

- **Explainability and Interpretability:** Develop techniques to understand the model's decision-making process for improved trust and transparency in forensic applications.

- **Real-time Processing:** Explore strategies to optimize the model for real-time deployment in law enforcement applications.

9.3 Application

- **Advanced Search Capabilities:** Integrate the model into facial recognition databases to enable searching across a person's appearance over time. This could be invaluable for missing person cases and investigations involving timeframes with significant aging.

- **Age Progression Systems:** Leverage the model's ability to generate age-related facial images to create a more robust age progression system for missing persons or unidentified individuals.

- **Forensic Age Estimation:** Refine the model to provide a more accurate estimation of a person's age from a facial image, aiding in forensic investigations where age is an important factor

- **Improved Surveillance Systems:** Integrate the model into surveillance systems to enhance recognition capabilities across age variations, potentially leading to faster identification of suspects.

- **Combating Age Bias:** The model's ability to address age bias in facial recognition can be utilized to develop fairer and more accurate facial recognition systems across various applications.

References

1. K. Kumar and M. Pavani, “Multi-features assisted age invariant face recognition and retrieval using cnn with scale invariant heat kernel signature,” IntechOpen,2022.

2. Z. Huang, J. Zhang, and H. Shan, “When age-invariant face recognition meets face age synthesis: A multi-task learning framework and a new benchmark,” IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023.

3. A. A. Moustafa and A. Elnakib, “Age-invariant face recognition based on deep features analysis,” 2018.

4. Y. Wang and D. Gong, “Orthogonal deep features decomposition for ageinvariant face recognition,” Computer Vision – ECCV 2018. ECCV 2018., 2018.

5. H. Wang, D. Gong, Z. Li, and W. Liu, “Decorrelated adversarial learning for age-invariant face recognition,” 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019.

6. L. Peipei and et al, “Uva: A universal variational framework for continuous age analysis.,” 2019.

7. Z. Zhang and et al, “Age progression/regression by conditional adversarial autoencoder.,” 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.

8. L. Ermao and et. al, “Review of cross-age face recognition in discriminative models,” 2023 8th International Conference on Image, Vision and Computing (ICIVC), 2023.

9. M. Abdulabas and et. al, “Face information forensics analysis based on facial aging: A survey,” 2022 International Conference on Intelligent Systems and Computer Vision (ISCV), 2022.

10. MS1m-ArcFace Dataset :https://drive.google.com/file/d/ 1SXS4-Am3bsKSK 615qbYdbAFMVh3sAvR/ view?usp=sharing.
CASIA WebFace Dataset :https://drive.google.com/file/d/ 1KxNCrXzln0lal3N4Ji Yl9cFOIhT78y1l/view?usp=sharing.

Chapter 10

Plagiarism Report