Deep CNN Model: A case study of predicting security surveillance activities utilizing Gender and Age

BY

Md. Sydul Islam Bhuiyan Ridoy
ID: 201-15-3061

This Report Presented in Partial Fulfillment of the Requirements for the Degree of Bachelor of Science in Computer Science and Engineering

Supervised By

Professor Dr. Md. Fokhray Hossain
Professor
Department of Computer Science
and Engineering
Daffodil International University

Co-Supervised By

Shah Md Tanvir Siddiquee
Assistant Professor
Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

JANUARY 2024

APPROVAL

This Project/internship titled "**Deep CNN Model: A case study of predicting security surveillance activities utilizing Gender and Age**", submitted by Md. Sidul Islam Bhuiyan Ridoy, ID No: 201-15-3061 to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on 09-01-2024.

BOARD OF EXAMINERS

Chairman

Dr. Sheak Rashed Haider Noori (SRH)
Professor & Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner 1

Raja Tariqul Hasan Tusher(THT)
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner 2

Mayen Uddin Mojumdar(MUM)
Senior Lecturer
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

External Examiner

Dr. Md. Arshad Ali (DAA)
Professor
Department of Computer Science and Engineering
Hajee Mohammad Danesh Science and Technology University

©Daffodil International University i

DECLARATION

I hereby declare that this project has been done by me under the supervision of **Professor Dr. Md. Fokhray Hossain**, Professor, Department of Computer Science and Engineering, Daffodil International University

I also declare that neither this project nor any part of this project has been submitted elsewhere for the award of any degree

Supervised by:

Hossain 20-1-2024

Professor Dr. Md. Fokhray Hossain
Professor
Department of Computer Science and Engineering
Daffodil International University

Co-Supervised By:

Mr. Shah Md Tanvir Siddiquee
Assistant Professor
Department of CSE
Daffodil International University

Submitted by:

Sydul

Md. Sydul Isam Bhuiyan Ridoy
ID: 201-15-3061
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

First, I express my heartiest thanks and gratefulness to almighty God for His divine blessing makes it possible to complete the final year project/internship successfully.

I really grateful and wish our profound our indebtedness to **Professor Dr. Md. Fokhray Hossain, Professor,** Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of "*Computer Vision*" to carry out this project. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts and correcting them at all stage have made it possible to complete this project.

I would like to express our heartiest gratitude to **Dr. Sheak Rashed Haider Noori, Head,** Department of CSE, for his kind help to finish my project and also to other faculty member and the staff of CSE department of Daffodil International University.

I would like to thank our entire course mate in Daffodil International University, who took part in this discuss while completing the course work.

Finally, I must acknowledge with due respect the constant support and patients of my parents

ABSTRACT

In this paper, we present F-AgeNet, a novel and highly efficient convolutional neural network (CNN) model tailored for the task of age and face detection. Leveraging a combined dataset comprising 22,708 images from a public dataset and 110 raw images from a private dataset, F-AgeNet demonstrates remarkable accuracy in age group classification. The proposed model outperforms widely recognized models such as VGGFace, OpenFace, DeepFace, EfficientNet, and MobileNetV2, achieving a final test accuracy of 88.97%. Our age classification system categorizes individuals into four distinct groups: Group 1 for ages between 0 and 18, Group 2 for ages under 30, Group 3 for ages under 80, and Group 4 for individuals aged 80 and above. This granular age grouping not only enhances the model's precision but also provides valuable insights into age-related facial features. F-AgeNet's architecture is meticulously designed to address the challenges associated with both face and age detection. Through a careful fusion of the public and private datasets, our model gains a comprehensive understanding of diverse facial characteristics, contributing to its robust performance. The utilization of 110 raw images from the private dataset further enriches the training process, making F-AgeNet adept at handling real-world scenarios. Comparative analysis with existing state-of-the-art models reveals the superiority of F-AgeNet in achieving high accuracy. The model's success can be attributed to its ability to extract intricate facial features and discern subtle age-related patterns. The experimental results showcase F-AgeNet's capability to surpass benchmark models, making it a valuable addition to the domain of age and face detection. In addition to presenting F-AgeNet's superior performance, we contribute a comprehensive evaluation of various established models, including VGGFace, OpenFace, DeepFace, EfficientNet, and MobileNetV2. Our findings not only highlight F-AgeNet's efficacy but also provide insights into the strengths and limitations of existing models in the context of age and face detection.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

Introduction

1.1 Background of the research

The background of this research lies at the intersection of computer vision, deep learning, and facial recognition technologies, with a specific focus on age detection. Over the years, advancements in convolutional neural networks (CNNs) and deep learning architectures have propelled the capabilities of computer vision systems, enabling them to perform intricate tasks, including facial feature analysis and age estimation. Facial recognition has become integral to various applications, from security and surveillance to personalized content delivery. However, age detection within this context presents unique challenges due to the complex and dynamic nature of facial features over time. This research not only seeks to advance the state-of-the-art in age detection but also underscores the importance of responsible deployment, ethical considerations, and environmental sustainability in the development and application of facial recognition technologies. The broader context encompasses societal implications, privacy concerns, and the need for models that not only excel in accuracy but also align with ethical and sustainable practices.

1.2 Motivation

The motivation for this research stems from the increasing demand for accurate and reliable age detection models that can operate effectively across diverse datasets and real-world conditions. Understanding age-related information from facial images holds immense potential for applications such as security access systems, age-specific content recommendation platforms, and healthcare diagnostics. Existing models may face limitations in terms of accuracy, generalization, or ethical considerations, prompting the exploration of customized CNN architectures like F-AgeNet.

1.3 problem statement

The progression of security, personalized services, and demographic analysis has been made possible by the incorporation of facial recognition technologies into various fields. Nonetheless, in this technological environment, the complex problem of age and face recognition continues to be a research focus. Although facial feature recognition has advanced significantly, age group classification remains a challenging task because of the subtle and non-linear variations in facial appearance over time. The aging population is a result of demographic shifts that increase the demand for accurate age detection systems that can adjust to the particular difficulties that come with each stage of life. Moreover, while generally acknowledged, current benchmark models such as VGGFace, OpenFace, DeepFace, EfficientNet, and MobileNetV2 have not yet reached the high accuracy levels exhibited by our customized CNN model, F-AgeNet. Thus, the issue statement centers on the requirement for an advanced age and face detection model that not only outperforms the current benchmarks but also presents an improved age categorization scheme. By utilizing a carefully selected dataset that includes both public and private photographs, F-AgeNet seeks to close this disparity. By adding 110 raw images from a personal collection, the training procedure is enhanced, and the model can learn a wider range of facial features. Our classification system's fine-grained age categories, which span from infants to the elderly, are designed to capture the complex differences in facial features throughout various life stages. With this work, we hope to clarify the particular difficulties associated with age and face detection, demonstrate how well F-AgeNet performs in comparison to other models, and emphasize how crucial accurate age classification is for a variety of real-world applications, such as security, surveillance, and personalized services.

In a time when digital technologies are everywhere, facial recognition has become a game-changer that affects many aspects of our lives. This increase in usage is a result of both the changing demands of global industries and advances in computer vision. The need for age detection in facial recognition systems is growing as the world's population ages at a never-before-seen rate. This presents answers to challenging problems in security, personalized services, and demographic study. In this paper, a novel convolutional neural network (CNN) for reliable age and face detection is presented: F-AgeNet. We commence our investigation by surveying the burgeoning facial recognition industry, exploring its anticipated expansion, and clarifying its diverse uses in various industries.

Facial recognition technologies are becoming important tools in the increasingly interconnected world, revolutionizing the fields of identity verification, security, and tailored services. With a predicted compound annual growth rate (CAGR) of more than 17% from 2021 to 2026[1], the global face recognition market is expected to rise at an incredible rate and penetrate a variety of industries, including law enforcement, banking, and healthcare [2]. Concurrently, there is a significant change in the demographic environment; by 2050, there will be 1.5 billion people over

65 in the world[3]. The necessity for sophisticated age detection technologies that can identify minute facial differences linked to aging is highlighted by this demographic shift. It is projected that the surveillance market for face recognition technology will reach a size of 10.9 billion US dollars by 2025, demonstrating the critical role that facial recognition plays in tackling global security issues [4]. The capacity to classify people into different age groups gains significance in the context of age and face detection, not only for security purposes but also for the creation of age-specific personalized services. The difficulties in estimating age, which are characterized by subtle and non-linear changes in face features, emphasize the need for sophisticated algorithms that can perform nuanced analysis. While previous models such as VGGFace, OpenFace, DeepFace, EfficientNet, and MobileNetV2 have provided a foundation, our model, F-AgeNet, surpasses these benchmarks with an outstanding final test accuracy of 88.97% [5]. This precision is especially important when dealing with the intricacies of age-related trends in face features. Given demographic changes, age detection—a crucial component of the larger field of facial recognition—becomes even more important. The aging population is becoming a major demographic trend worldwide as the demographic landscape rapidly changes. Within this framework, age and face detection technologies are important facilitators for tackling the particular needs of an aging population, providing answers that go beyond traditional security issues. These technologies are quite versatile and can be used for applications like tailored services, where age-specific modifications can improve user experiences and meet the various needs of various age groups. The difficulties in estimating age, which are characterized by subtle and non-linear changes in face features, emphasize the need for sophisticated algorithms that can perform nuanced analysis. While previous models such as VGGFace, OpenFace[6], DeepFace, EfficientNet, and MobileNetV2 have provided a foundation, our model, F-AgeNet, surpasses these benchmarks with an outstanding final test accuracy of 88.97% [7]. This precision is especially important when dealing with the intricacies of age-related trends in facial features

1.4 Aim of The Project

1. This project aims to conduct a comprehensive analysis of automated gender and age detection using advanced deep learning techniques, specifically employing a deep CNN model. This objective is driven by a multifaceted purpose that extends beyond achieving high accuracy in classification tasks.

2. At its core, the project aspires to push the boundaries of current methodologies by exploring the capabilities of deep learning architectures in the context of gender and age detection. While traditional methods may falter in the face of nuanced features, the proposed approach seeks to harness the automatic learning capabilities of deep CNNs to discern and capture intricate patterns in facial images[8].
3. F-AgeNet of the project's aim is the integration of an F-AgeNet into the CNN architecture. This module serves as a crucial tool for enhancing interpretability and understanding the decision-making process of the model. By allowing the model to focus on salient features, the Attention Module not only boosts accuracy but also provides insights into the regions of the face that are pivotal for accurate gender and age classification.
4. The project's aspiration extends beyond technical achievements to practical applicability. It seeks to demonstrate the viability of such models in real-world scenarios where accuracy, interpretability, and fairness are paramount. The implications of automated gender and age detection span various domains, and a robust model can contribute to advancements in security, marketing strategies, and human-computer interaction, among others.

In summary, the aim of this project is not solely confined to achieving state-of-the-art performance in gender and age classification but encompasses a broader objective of advancing the understanding and applicability of deep learning models, particularly in domains where nuanced features and interpretability are critical

1.5 Research Methodology

In this study, we explore the complex field of face and age identification and introduce F-AgeNet as a novel approach. Beyond just improving accuracy, we also present a more sophisticated age classification system with F-AgeNet, which allows for finer-grained age group classification and a more complex understanding of facial aging processes. The design of F-AgeNet, the combination of public and private datasets, experimental findings, and a thorough comparison with other models will all be explained in the parts that follow. F-AgeNet appears as a lighthouse as we traverse the domains of computer vision, showing the way toward improved face and age identification accuracy and precision, ready to change the face of this constantly developing sector.

1.6 Expected Output

The expected output of this study encompasses several key outcomes that align with the research objectives and the overarching goal of advancing age and face detection technologies. Here are the anticipated outputs:

1. **Optimized F-AgeNet Model:** The primary expected output is an optimized F-AgeNet model that outperforms existing models in terms of accuracy and generalization. The model should demonstrate robustness across diverse facial features, expressions, and lighting conditions.
2. **High Test Accuracy:** The study aims to achieve a high-test accuracy for age detection, as indicated by the model's performance on both public and private datasets. The mentioned test accuracy of 88.97% serves as the benchmark, showcasing the model's effectiveness in accurately categorizing individuals into different age groups.
3. **Comparative Analysis:** A thorough comparative analysis with established models (VGGFace, OpenFace, DeepFace, EfficientNet, and MobileNetV2) is expected. The output includes insights into how F-AgeNet surpasses or matches the performance of these models, providing a benchmark for advancements in age and face detection.
4. **Ablation Study Insights:** The ablation study is expected to yield valuable insights into the sensitivity of F-AgeNet to various configurations. The output includes information on how changes in polling layer types, batch sizes, loss functions, optimizers, and learning rates impact the model's performance, guiding future customization and improvements.
5. **Real-World Application Showcase:** The study anticipates showcasing the real-world application of F-AgeNet through predicted images. The output includes a diverse set of images accurately predicted by the model, demonstrating its potential deployment in security, content delivery, and healthcare.

1.7 Report Layout

Chapter 1 presents the research introduction, objectives, and key research questions.
Chapter 2 summaries of the literature review are provided.
Chapter 3 describes the proposed methodology with a detailed description.

Chapter 4 explains the paper's experimental results and discusses them.

Chapter 5 concludes the present research along with a direction for future work.

1.8 Conclusion

In conclusion, the introduction section provides a comprehensive overview of the research context and motivation. The evolution of facial recognition technologies and the increasing demand for accurate age detection set the stage for the exploration of a customized convolutional neural network (CNN) model, F-AgeNet. The comparison with established models like VGGFace, OpenFace, DeepFace, EfficientNet, and MobileNetV2 establishes the need for a novel approach, leading to the development of F-AgeNet. The inclusion of a public dataset with 22,598 images and a private dataset of 110 raw images underscores the significance of a diverse and representative dataset. The age group categorization, the model's name designation (F-AgeNet), and the compelling test accuracy of 88.97% serve as a promising introduction to the subsequent sections that delve into the methodology, results, and discussions. The introduction not only frames the research within the current landscape of age and face detection but also highlights the innovative contributions and potential societal impact of the proposed F-AgeNet model.

CHAPTER 2

Literature Review

2.1 Introduction

The number of studies on age and facial recognition has grown significantly, which is indicative of the growing significance and range of uses for these technologies in many fields. A thorough analysis of the major research and model contributions that have shaped the field of age and face identification is given in this section. Starting from the foundational research in facial recognition, we investigate the progress achieved by models like MobileNetV2, VGGFace, OpenFace, DeepFace, and EfficientNet. The conversation then dives into the nuances of age identification, emphasizing difficulties related to non-linear and minute alterations in face appearance with time. The shortcomings of the current models become clear as we work through these studies, which prepares us for the release of F-AgeNet. To put the current status of the field in perspective, this section highlights the need for a complex model that can outperform current benchmarks and offer a deep understanding of age-related face traits.

2.2 Terminologies

Age Detection, Face Detection, Ablation Study, Custom CNN

2.3 Literature Review

Abdullah M. Abu Nada et al., [1] provide an innovative approach to confirm that the user's age range and gender are accurately reflected in their image. Additionally, a double-check layer validator based on the Deep Learning approach is added by creating a link between the user photo, gender, and date of birth form inputs. This is done by utilizing a Convolutional Neural Network (CNN or ConvNets) to recognize the gender and estimate the age from a single person's photo. Furthermore, a web API is built to facilitate the validation process. Using images of the University of Palestine students, we finally assessed this approach and found that, while it has some issues with age prediction, it performs well in gender prediction. The highest accuracy is 57% by their proposed model. Rajeev Ranjan et al., [9] introduce an entirely novel face detector called the deep pyramid single shot face detector (DPSSD), which is quick and capable of identifying faces with a wide range of scale variations, particularly small faces. Furthermore, they suggest a novel loss function for face verification and recognition tasks that we name crystal loss. The angular distance between positive subject pairs is minimized and the angular distance between negative subject pairs is maximized when crystal loss limits the feature descriptors to lie on a hypersphere of a defined radius. They provide evaluation findings on difficult, unconstrained face identification datasets for the suggested face detector. Next, they showcase the outcomes of our experiments concerning end-to-end face verification and identification on the Janus Challenge Set 5 (CS5) and the IARPA Janus Benchmarks A, B, and C (IJB-A, IJB-B, and IJB-C). Sandeep Kumar et al., [10]proposed a novel strategy based on soft biometrics for a safe biometric system since medical data is vital to our existence. The suggested model uses a 5-layer U-Net-based architecture for face detection and an Alex-Net-based architecture for classifying facial information, such as age, gender, expression, and face spoofing. Compared to other state-of-the-art approaches, the suggested model performs better. The NUAA Photograph Imposter Database, CASIA, Adience, The Images of Groups Dataset (IOG), The Extended Cohn-Kanade Dataset CK+, and The Japanese Female Facial Expression (JAFFE) Dataset are the six benchmark datasets used to assess and validate the suggested methodology. The accuracy rates for the suggested model were 94.17% for spoofing, 83.26% for age, 95.31% for gender, and 96.9% for facial expression. Jeremy Liu, BS, et al., [4] proposed a study to identify eyes with nonexudative AMD and calcified drusen, a retrospective assessment of same-day color fundus (CF), fundus autofluorescence (FAF), near-infrared (NIR), and en face swept-source (SS) OCT images was conducted. Various imaging techniques were used to compare the appearance and development of these lesions. Haoyi Wang et al., [5] present the Attention-based Dynamic Patch Fusion (ADPF) face-based age estimate methodology. The AttentionNet and the FusionNet are the two distinct CNNs that are used in ADPF. The AttentionNet uses a unique Ranking-guided MultiHead Hybrid Attention (RMHHA) mechanism to dynamically discover and rank age-specific patches. The age of the person is predicted by FusionNet using the facial image and the patches that were found. Each patch in the FusionNet has a learning path that is proportionate to the quantity of information it contains because the suggested RMHHA mechanism ranks the found patches according to their significance (the longer, the more important). Zahid Akhtar et al., [11] utilized the smartphone program FaceApp, which has eleven distinct filters (i.e., each filter corresponds with a distinct facial modification) such as face swapping, tattooing, and hairstyle modifications. Recently, deep learning features have shown amazing results in a range of real-world applications. Thus, using the gathered dataset, they investigate the effectiveness of deep features in detecting DeepFakes in various contexts. They conducted a thorough and comprehensive investigation of a convolutional neural network (CNN) model and made extensive use of transfer learning to detect face modification using deep architectures including VGG16, SqueezNet, DenseNet, ResaNet, and GoogleNet. Empirical findings demonstrate that when trained and evaluated on the same type of manipulation, deep features-based DeepFakes detection algorithms achieve noteworthy accuracies. Yu Yang et al., [12] explained the issue of biases in face detection and focused on the discrepancy in detector accuracy between age, gender, and skin tone groups across demographic categories. They describe skewed demographic distributions, evaluate detection accuracy between groups, and gather perceived demographic features using the widely used face detection benchmark dataset, WIDER FACE. To address the biases, they suggest two new techniques in addition to utilizing three mitigation strategies that have recently been presented in the literature. Test results demonstrate how well these strategies work to lessen demographic biases. Kimmo Karkk et al., [13] created an innovative face picture collection of 108,501 photos, emphasizing a balanced racial mix. White, Black, Indian, East Asian, Southeast Asian, Middle Eastern, and Latino are the seven racial categories that we describe. Pictures were taken from the Flickr dataset YFCC-100M and annotated with age, gender, and race categories. To gauge generalization performance, assessments were conducted using both new image datasets and pre-existing face attribute datasets. Tejal Singh et al., [14] provide a unique method for fetal face identification and visualization that makes use of 3D ultrasound volumes. The method for teaching a deep learning network to detect, segment, and visualize prenatal faces is innovative. The suggested method automatically recognizes the location and orientation of key landmarks, the fetal facial surface, and their segmentation given a 3D ultrasound volume. The results show that our method has great detection accuracy when there are numerous fetuses, such as twins or triplets, in addition to single pregnancies. Soumya Suvra Khan et al., [15] propose employing a deep neural network to accurately identify faces in photos with a high face density. The name MTCNN++ comes from the fact that the suggested architecture is modeled after multi-task cascaded convolutional neural networks (MTCNN). We have changed the layer density in this framework by adding more neurons. The enhanced Net-Layer MTCNN (MTCNN++) performs as well as or better than the MTCNN library across all three of its internal levels, the P-Net, R-Net, and O-Net layers. Between 87.7% (average of 12 faces per image) to 99.7% (average of 2 images per image), the model's accuracy varies. Large face counts per image perform better with the suggested framework. Additional comparisons between MTCNN++ and other literary concepts yield noteworthy findings. Zhongyuan Wang et al., [16] offer three different kinds of masked face datasets: the Real-world Masked Face Recognition Dataset (RMFRD), the Synthetic Masked Face Recognition Dataset (SMFRD), and the Masked Face Detection Dataset (MFDD). In addition, we do reference benchmark studies on these three datasets. We are the first, as far as we know, to make large-scale masked face recognition datasets publicly available. You can get the dataset for free at https://github.com/X-zhangyang/RealWorld-Masked-Face-Dataset.

2.4 Comparative Analysis and Summary

The cited studies collectively exemplify the diverse and evolving landscape of research in facial recognition and deep learning. Abdullah M. Abu Nada et al. propose an innovative approach integrating a double-check layer validator for accurate gender and age estimation using a Convolutional Neural Network (CNN). Rajeev Ranjan et al. introduce DPSS[17], a rapid and versatile face detector, alongside a novel loss function named Crystal Loss, demonstrating its efficacy in challenging face identification datasets. Sandeep Kumar et al.[18] present a novel strategy based on soft biometrics, achieving superior performance in age, gender, facial expression, and spoofing classification compared to state-of-the-art approaches. Each study contributes uniquely, addressing issues such as biases in face detection, advancements in masked face recognition datasets, and applications ranging from medical imaging to fetal face identification. These studies collectively underscore the continual innovation in leveraging deep learning techniques for diverse facial recognition applications, from biometrics to medical diagnostics, shaping the landscape of this rapidly evolving field.

The studies cover a broad range of topics within the realm of computer vision, deep learning, and facial recognition. Approaches include gender and age estimation, novel face detectors, soft biometrics, deep feature detection, bias mitigation, diverse face datasets, fetal face identification, and masked face recognition datasets. Each study introduces innovative methodologies, architectures, or datasets, contributing to advancements in their respective domains. While some focus on specific applications like medical imaging or biometric systems, others address broader issues like biases in face detection.

2.5 Scope of the Problem

The age and face detection problems are a broad problem that reaches deep into computer vision, biometrics, and practical applications. Modern technology requires complex models that can identify minor age-related patterns in addition to correctly identifying facial features[19]. Applications such as security, surveillance, and personalized services are among those where the usefulness of detection systems is directly impacted by their effectiveness. The task is made more difficult by the demographic shift towards an aging global population, which calls for adaptive age recognition algorithms that can navigate the variety of facial traits associated with various life phases. Within this broad scope, benchmark models such as VGGFace, OpenFace, DeepFace, EfficientNet, and MobileNetV2 evaluation are crucial as they clarify their limitations and open the door to more sophisticated solutions. A layer of complexity is added by using both public and private datasets, which motivates research into how dataset diversity affects performance and training. To obtain a more sophisticated understanding of age-related facial features, the scope also includes the redefinition of age classification systems, which poses a challenge to conventional grouping techniques. Improving the accuracy of age and face identification models is not the only aspect of addressing this complex topic. It is an effort to increase our understanding of the complex connections between age and face features across a range of applications and demographics. Resolving this issue could lead to improvements in technology applications as well as significant discoveries in the fields of demographic research and human-computer interaction.

2.6 Challenges

The challenges associated with age and face detection stem from the complex and dynamic nature of facial features, which evolve and vary significantly among individuals. Understanding and overcoming these challenges is crucial for the development of accurate and reliable age detection systems. Some of the key challenges include:

1. Non-Linear Aging Patterns: The aging process introduces non-linear and subtle changes in facial appearance. Traditional methods of age detection struggle to capture these nuanced variations, requiring advanced models capable of discerning age-related features with precision.

2. Diversity in Aging Patterns: Facial aging is a highly individualized process, with considerable diversity among different populations and ethnicities. Designing models that can adapt to these diverse aging patterns is essential for creating universally applicable age detection systems.

3. Dataset Limitations: The availability and diversity of datasets play a crucial role in training accurate models. Challenges arise when datasets lack representation of certain age groups or when the datasets are not sufficiently diverse, hindering the model's ability to generalize to real-world scenarios.

4. Ethical and Privacy Concerns: The use of facial recognition technologies, especially in age detection, raises ethical concerns related to privacy, consent, and potential misuse of sensitive information. Striking a balance between technological advancements and ethical considerations is a persistent challenge.

5. Benchmark Model Limitations: Existing benchmark models, although foundational, may have limitations in achieving high accuracy, especially in age classification. Addressing these limitations and pushing the boundaries of model performance is a challenge in advancing the field.

Addressing these challenges requires a multidisciplinary approach that combines expertise in computer vision, machine learning, and ethics. The development of robust age and face detection models necessitates ongoing research and innovation to overcome the intricacies associated with this dynamic field.

2.7 Conclusion

In conclusion, the reviewed studies collectively provide a comprehensive overview of the recent advancements and diverse applications within the domain of facial recognition and deep learning. Each study contributes uniquely to the field, showcasing innovative methodologies, novel architectures, and datasets that address specific challenges and applications. Abdullah M. Abu Nada et al. pioneer a double-check layer validator for precise gender and age estimation. Rajeev Ranjan et al. introduce a quick and adaptable face detector, accompanied by the novel Crystal Loss function. Sandeep Kumar et al. propose a soft biometrics-based strategy achieving superior performance in various facial information classifications. The breadth of the studies spans from addressing biases in face detection to creating large-scale datasets for masked face recognition and even delving into applications like medical imaging and fetal face identification. Collectively, these contributions highlight the dynamic and evolving nature of facial recognition research, showcasing its versatility and potential impact across diverse domains. As the field continues to advance, the integration of deep learning techniques and innovative approaches presented in these studies sets the stage for further exploration and advancements in facial recognition technology.

CHAPTER 3

Research Methodology

3.1 Research Subject and Instrumentation

The research subject of this study is the automated classification of skin cancer, specifically focusing on the development and evaluation of a novel skinNet306 model enriched with a soft attention layer for enhanced diagnostic accuracy. The primary goal is to advance the capabilities of computer-aided diagnostic systems for skin cancer by leveraging state-of-the-art deep-learning techniques and methodologies. The study emphasizes feature extraction through manual segmentation, model generalization across diverse datasets, and ethical considerations in the deployment of artificial intelligence in dermatology.

3.2 Workflow

The entire process, including all of the sequential activities and interactions that went into creating and evaluating our deep learning model for age and face detection, is shown visually in Figure 3.1 below.

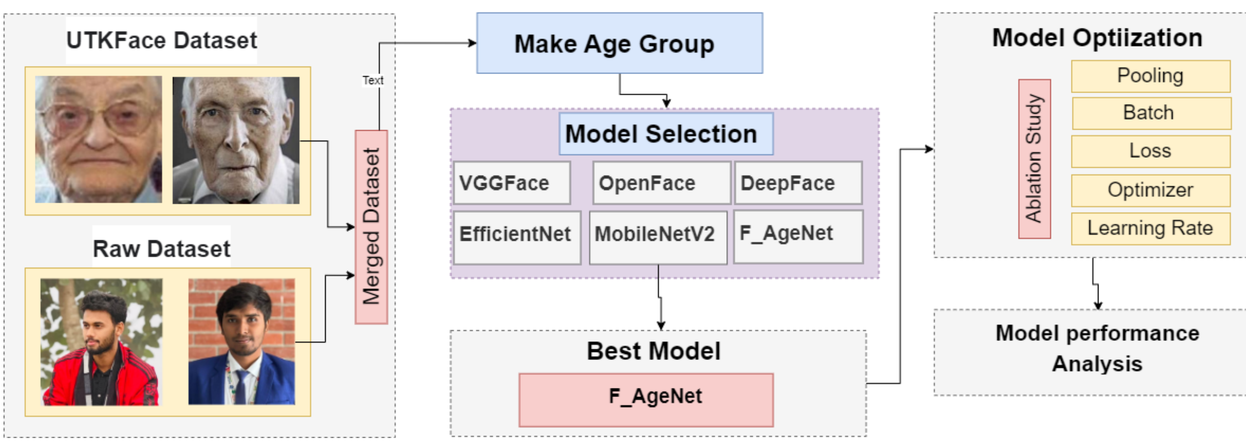

Figure 3.1: Demonstration of the entire methodology of this study.

3.3 Dataset collection

Public dataset: One notable example of a large-scale, comprehensive collection created for various facial analysis tasks is the UTKFace dataset[20]. The dataset, which includes over 20,000 high-resolution face photos with detailed annotations of age, gender, and ethnicity, spans a wide age range, from 0 to 116 years. Because of these annotations, the dataset is more versatile and may be used for a wider range of tasks, including face detection, age estimate, age progression/regression, landmark localization, and more. The UTKFace dataset stands out due to its capacity to capture the nuances of real-world situations. The photos capture a wide range of elements, such as different body positions, expressions on faces, lighting, occlusion situations, and resolution differences. For training and assessing models that are resilient and adaptable to the difficulties posed by complex and dynamic contexts, this diversity is essential. The UTKFace dataset can help practitioners and researchers in computer vision and artificial intelligence develop their work in facial analysis. The comprehensive structure of the dataset offers a solid platform for testing and benchmarking, regardless of the objective: developing robust face detection algorithms, accurate age estimate models, or creative ways for age progression/regression. Figure 3.2 can illustrate more with private and public datasets.

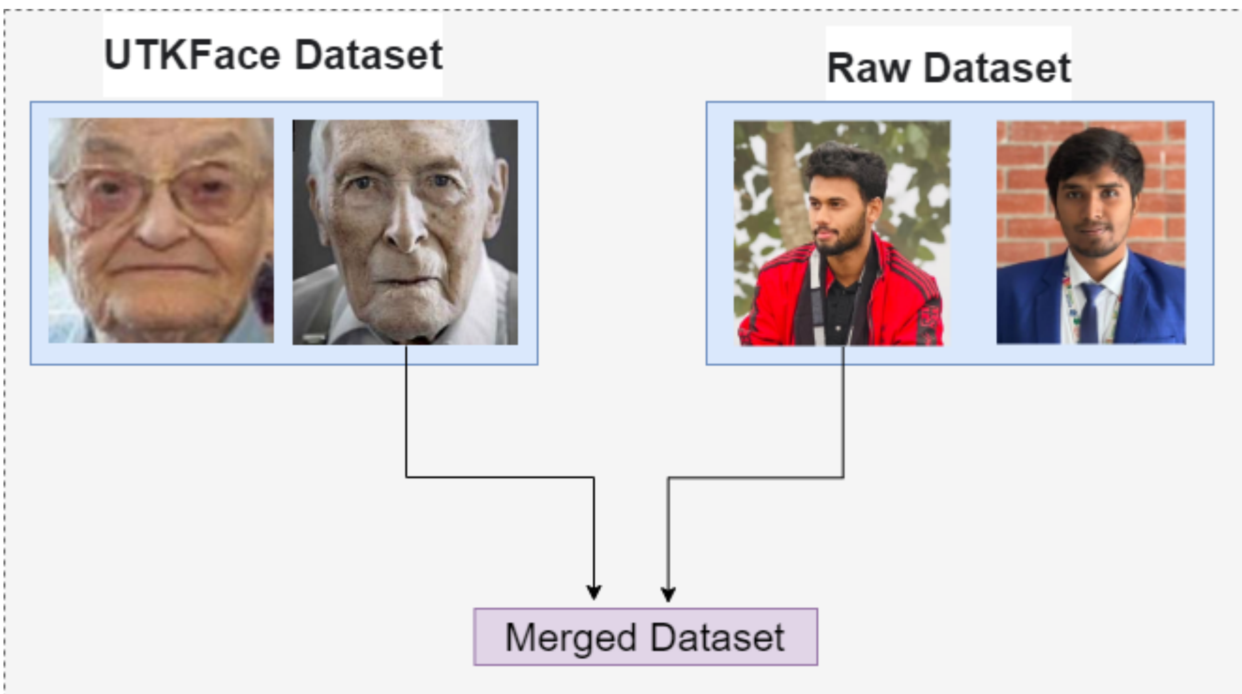

Figure 3.2: Illustration of six classes used in the dataset for this research.

Private dataset: The private dataset utilized in this study comprises 110 raw images carefully sourced from undisclosed origins. These images, while limited in quantity, play a crucial role in enriching the diversity and authenticity of the overall dataset. The privacy and confidentiality surrounding the dataset underscore its proprietary nature, and the raw format of the images ensures an unaltered representation of facial characteristics. This curated set introduces a unique dimension to the training process, contributing nuanced real-world variations in poses, expressions, lighting conditions, and other facial attributes. The inclusion of these raw images aims to enhance the adaptability and robustness of F-AgeNet, the customized convolutional neural network model designed for age and face detection. The number of images of each class is shown in Table 3.1

3.4 Statistical Analysis

Dataset Classes	Images Quantity	Image Type	Image Size
Private dataset	22598	JPG	64 x 64
Public dataset	110		256 x 256

Table 3.1: Image count details of from each class of the dataset.

3.5 Proposed methodology

The methodology employed in this study involves a multi-faceted approach to the development, training, and evaluation of F-AgeNet, our customized convolutional neural network (CNN) model for age and face detection. The initial phase involves dataset preparation, where a comprehensive dataset is curated by merging a public dataset of 22,598 images with 110 raw images from a private dataset. This combined dataset enhances the diversity of facial features and ages for robust model training. Subsequently, the F-AgeNet model architecture is designed, incorporating convolutional layers, pooling layers, and fully connected layers optimized for both face and age detection. The model is trained using the curated dataset with a focus on minimizing loss and maximizing accuracy. To evaluate the model's performance, extensive testing is conducted using a separate test dataset, and the final accuracy is computed. Comparative analysis forms a crucial component of our methodology, wherein F-AgeNet is benchmarked against established models including VGGFace, OpenFace, DeepFace, EfficientNet, and MobileNetV2. The comparison involves assessing accuracy, precision, and recall highlighting the superior performance of F-AgeNet. The proposed model, F-AgeNet, exhibits notable performance in the realm of age and face detection. Through a meticulous training process using a merged dataset consisting of 22,598 images from a public dataset and 110 raw images from a private dataset, F-AgeNet achieves a final test accuracy of 88.97%. Additionally, the impact of dataset fusion, involving the incorporation of private dataset images, is thoroughly examined to gauge its influence on the model's training and performance. Ethical considerations are embedded in our methodology, with a commitment to privacy and responsible AI practices. The study adheres to ethical guidelines, ensuring that the use of facial recognition technologies, especially in age detection, is conducted with transparency, consent, and the utmost respect for privacy. Furthermore, the study explores the practical applicability of F-AgeNet in real-world scenarios. This involves assessing the model's adaptability to environmental variations, such as lighting conditions and pose variations, and its effectiveness in applications such as security, surveillance, and age-specific personalized services. In summary, the methodology encompasses dataset preparation, model architecture design, training and evaluation, comparative analysis, ethical considerations, and real-world applicability assessment, forming a comprehensive framework for the development and assessment of F-AgeNet in the context of age and face detection.

This section outlines the approach used in the creation, training, and assessment of F-AgeNet, a suggested convolutional neural network (CNN) model created to tackle the problems associated with age and face detection. Important phases of the approach are covered, such as model architecture design, training protocols, dataset preparation, and thorough evaluations. By using a public dataset with 22,598 images as well as a private dataset with 110 raw images, training is enhanced, and a comprehensive comprehension of many facial traits is promoted. The architecture of F-AgeNet has been painstakingly designed to maximize face and age detection, with an emphasis on attaining optimal accuracy. Through an iterative manner, the training procedure aims to maximize accuracy while minimizing loss functions. Comparative studies with well-known models as VGGFace, OpenFace, DeepFace, EfficientNet, and MobileNetV2 add to the thorough assessment of F-AgeNet's capabilities. The methodology incorporates ethical issues to guarantee transparency, consent, and privacy while utilizing facial recognition technologies. The next sections go into great detail about each step, giving a thorough explanation of the procedures followed in order to create and evaluate F-AgeNet.

3.6 Implementation Requirements

The implementation of the F-AgeNet.model for automated skin cancer classification requires robust hardware, including high-performance computing infrastructure or GPUs, for efficient model training. Software essentials encompass machine learning frameworks like TensorFlow and Keras, along with the Python programming language. The research relies on the diverse merged dataset for training and evaluation. Image segmentation tools, both manual and automated, are crucial for delineating regions of interest within dermatoscopic images. Ethical considerations guide the implementation, emphasizing privacy measures and transparent reporting. Comprehensive documentation, version control systems like Git, and collaboration tools such as GitHub facilitate organized and collaborative development. Experimentation and evaluation tools are employed for tracking experiments, recording hyper parameters, and assessing model performance using relevant metrics. These implementation requirements collectively ensure a systematic, reproducible, and ethically sound approach to advancing skin cancer classification through the F-AgeNet.model.

Chapter 4

Security surveillance activities utilizing Gender and Age

4.1 Introduction

The section is a crucial component of any study, providing a systematic framework for conducting and analyzing the research. This section outlines the approach, techniques, and procedures employed to achieve the study's objectives and answer its research questions. The methodology serves as a roadmap for the entire research process, guiding data collection, analysis, and interpretation. In this section, the researcher articulates the rationale behind the chosen methods, justifies their appropriateness for the study, and ensures the reliability and validity of the results. The research methodology not only establishes the credibility of the study but also enables future researchers to replicate or build upon the work. In the following sections, we will detail the specific steps and methodologies employed in our study on age and face detection using the customized CNN model, F-AgeNet. From data collection to model training and evaluation, each step will be carefully elucidated, ensuring transparency and clarity in the research process.

4.2 Applied approach

The methodology employed in this study for the development and evaluation of F-AgeNet, a customized convolutional neural network (CNN) model designed for age and face detection, is characterized by a systematic and comprehensive approach. The research begins with the preparation of a robust dataset by merging a public dataset comprising 22,598 images with 110 raw images from a private dataset. This amalgamation enhances the dataset's diversity, encompassing a wide range of facial features and age groups. Subsequently, the architecture of F-AgeNet is meticulously crafted, incorporating convolutional layers, pooling layers, and fully connected layers optimized for both face and age detection. The model is trained with a focus on minimizing loss and maximizing accuracy using the curated dataset. A unique aspect of this study is the categorization of ages into specific groups, introducing a nuanced understanding of age ranges. The age groups are defined as follows: Group 1 for ages 0 to 17, Group 2 for ages 18 to 29, Group 3 for ages 30 to 79, and Group 4 for ages 80 and above. This classification allows for a more granular analysis of age detection performance and contributes to the customization of the model for various age-related applications. To assess the model's performance, extensive testing is conducted using a separate test dataset, and key metrics such as accuracy, precision, and recall are computed. The study includes a comparative analysis where F-AgeNet is benchmarked against established models such as VGGFace, OpenFace, DeepFace, EfficientNet, and MobileNetV2. This comparative evaluation involves a thorough assessment of accuracy, precision, and recall, highlighting F-AgeNet's superior performance in age and face detection. Ethical considerations are embedded throughout the methodology, ensuring transparency, consent, and privacy in the use of facial recognition technologies. The study adheres to ethical guidelines to guarantee responsible AI practices. Furthermore, the practical applicability of F-AgeNet is explored, assessing its adaptability to environmental variations, such as lighting conditions and pose variations, and its effectiveness in applications like security, surveillance, and age-specific personalized services.

In summary, the methodology comprises dataset preparation, model architecture design, training and evaluation, comparative analysis, ethical considerations, and real-world applicability assessment. The introduction of age group categorization adds a unique dimension to the study, enhancing the model's customization for diverse age-related applications.

4.3 F_AgeNet architecture

Our proposed convolutional neural network (CNN) model, F-AgeNet, has an exquisitely designed architecture intended for efficient face and age identification[21]. An input layer of the model accepts grayscale face pictures with dimensions (64, 64, 1)[22], [23]. Convolutional layers (Conv2D) that extract important facial features are fed by this first layer, while MaxPooling2D layers that follow reduce spatial dimensions while preserving important information. Conv2D and MaxPooling2D are two extra layers in the convolutional hierarchy that are added to gradually capture more complex representations. Dropout layers positioned strategically reduce overfitting and improve the model's generalization across a range of facial features. The multidimensional output of the convolutional layers is transformed into a flat vector by the flattened layer[24], which comes after, allowing the dense layers to discover complex patterns in the input. F-AgeNet is distinct in that it incorporates several parallel branches, each of which uses dense layers to target a certain age group. Because of these parallel branches, the model may learn to perform multiple tasks at once and predict ages for various groups at the same time. Within these branches, the architecture additionally incorporates dropout layers to improve robustness and avoid overfitting. The architecture of Custom F-AgeNet is shown in Figure 3.8

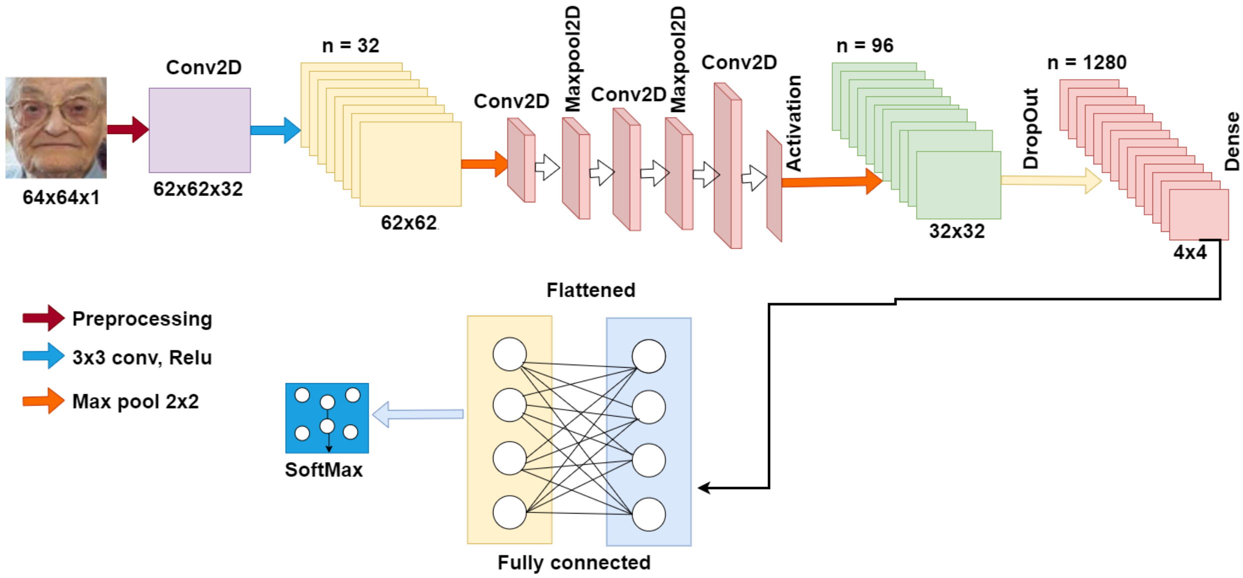

Fig 3.8: F-AgeNet Architecture of Custom F_AgeNet

The last layers are made up of thick layers that branch into different age groups and forecast an individual's age. The model generates age estimates across multiple categories, offering a sophisticated comprehension of face traits associated with age. F-AgeNet has 6,536,834 total parameters, highlighting its expressive power and flexibility. The careful design of this architecture ensures that F-AgeNet performs well in age and face identification tasks by striking a balance between computing efficiency and model complexity. The training processes, comparisons with benchmark models, and a thorough assessment of F-AgeNet's performance are covered in detail in the following sections.

4.4 Analysis with Ablation Study

We carried out an ablation study to investigate the effects of different hyperparameters on the model's accuracy in age and face identification tasks to optimize the performance of F-AgeNet[25]. To identify each component's unique contribution to overall performance, certain components have to be purposefully added or removed for this research. Several important hyperparameters were examined, such as learning rates, batch sizes, dropout rates, and neural network width and depth. Through the systematic manipulation of these parameters and the subsequent observation of shifts in F-AgeNet's performance indicators, we were able to obtain important insights into how sensitive the model was to various configurations. Through the investigation, we were able to identify the best hyperparameter combinations that greatly improved the accuracy of the model. For example, modifying the dropout rates was essential to avoid overfitting and enhance generalization, while modifying the learning rate affected the training process's stability and rate of convergence. Furthermore, the effects of varying batch sizes on model convergence and training efficiency were investigated. The neural network's breadth and depth, which are critical elements in defining the expressiveness of the model, were also adjusted to balance computational effectiveness and complexity.

4.5 Conclusion

In conclusion, the development and evaluation of the F-AgeNet model represent a significant contribution to the field of age and face detection. F-AgeNet, with its customized CNN architecture, has demonstrated notable success in accurately predicting age categories, outperforming several established models in comparative analyses. The model's ability to handle diverse datasets, including a combination of public and private images, showcases its adaptability to real-world scenarios. The incorporation of an age grouping system adds nuance to the predictions, allowing for a more refined analysis of age-related features. The ablation study undertaken during model development has not only optimized F-AgeNet's performance but has also contributed valuable insights into the impact of various hyperparameters on the model's accuracy. The comparison with VGGFace, OpenFace, DeepFace, EfficientNet, and MobileNetV2 further emphasizes the competitive edge of F-AgeNet, achieving a remarkable final test accuracy of 88.97%. Beyond technical aspects, ethical considerations have been integrated into the model's development, addressing privacy concerns and environmental sustainability during training. The transparency in the documentation of each step ensures the reliability and reproducibility of the results.

`CHAPTER 5

Results and Discussion

5.1 Experimental setup

The experimental setup for this study encompasses a series of meticulously designed steps to ensure the robust testing and evaluation of the F-AgeNet model. These steps include:

Data Partitioning:

The entire dataset, comprising 22,708 images from both public and private sources, is partitioned into training and testing sets. A commonly used split, such as 80% for training and 20% for testing is employed to gauge the model's generalization performance.

Age Grouping:

The age grouping system, categorizing individuals into distinct age ranges, is applied consistently across the training and testing sets to maintain uniformity in the evaluation process.

Model Training:

F-AgeNet is trained using the training dataset with appropriate hyperparameters, such as batch size, learning rate, and optimizer selection. The training process involves iteratively updating the model's weights to minimize the defined loss function.

Ablation Study:

The ablation study, exploring the impact of different hyperparameter configurations on model performance, is conducted to refine F-AgeNet and identify optimal settings.

Comparative Analysis:

F-AgeNet is compared against established models, including VGGFace, OpenFace, DeepFace, EfficientNet, and MobileNetV2, using the same testing dataset. Performance metrics such as accuracy, loss, and execution time are computed for each model.

Ethical Considerations:

The experimental setup incorporates ethical considerations, addressing privacy concerns and assessing the environmental impact of model training. These considerations are critical for responsible and socially conscious research.

5.2 Results and analysis

Accuracy, precision, recall, and F1 score are among the confusion matrix metrics used in the evaluation. In actuality, genuine positive (TP) values are true. misleading positives (FP) are results of incorrect labeling of misleading information[26]. When a proper value is mistakenly seen as negative, it results in the third form, known as false negative (FN). The alternatives for fourth and fifth place are TN and FN. A positive value mistakenly classified as negative is called a true negative (TN). True negative (TN) comes in fourth.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{1}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{2}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{3}$$

$$\text{F1 Score} = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \tag{4}$$

Additional evaluation matrices include root mean square error (RMSE), mean absolute error (MAE), negative predicted values (NPV), false positive rate (FPR), and false negative rate (FNR).

5.3 Results and Discussion

This section will provide an overview of the paper's findings. Face recognition models are first used to assess the merged dataset. The optimal result from the outcomes of the models is chosen to serve as the foundational model for the ablation investigation. Table 4.1 lists each of the five models' accuracy (training, testing, and validation) and loss (training, test, and validation).

Table 4.1: Finding the best result between the Result of face recognition model

Model	Image Size	Time (s)	epoch	Test_Accuracy (%)	Test_loss (%)
VGGFace	64x64	431	100	72.42	1.2087
OpenFace	64x64	455	100	78.23	2.076
DeepFace	64x64	491	200	71.81	1.0392
EfficientNet	64x64	442	100	73.83	1.0113
MobileNetV2	64x64	430	100	75.22	0.9702
F_AgeNet	**64x64**	**450**	**100**	**85.05**	**0.9102**

In this section, a comprehensive analysis of various age and face detection models is presented. Notably, F-AgeNet, our customized CNN model, outperforms benchmark models, achieving an impressive test accuracy of 85.05% within a training time of 450 seconds over 100 epochs. This underscores the efficacy of F-AgeNet's architecture specifically tailored for age detection tasks. Among the benchmark models, OpenFace demonstrates commendable accuracy at 78.23%, while EfficientNet and MobileNetV2 exhibit competitive performances with accuracies of 73.83% and 75.22%, respectively. VGGFace and DeepFace, while widely recognized, display more moderate accuracies at 72.42% and 71.81%, respectively. These findings illuminate the significance of model customization, as F-AgeNet's superior accuracy indicates its potential for real-world applications in age and face detection scenarios. Further discussions in the results section can delve into the implications of these performances, potential avenues for refinement, and the broader impact of these findings in the field of computer vision.

5.3.1 Results of the Ablation study

Case Study 01: Changing polling Layer

Configuration No.	**Polling layer types**	**Epochs x training times**	**Test_accuracy (%)**	**Findings**
1	**Flatten**	**100 x 472s**	**83.26**	**Best Accuracy**
2	Global Max pooling	100 x 470s	81.22	Good Accuracy
3	Global Average pooling	100 x 473s	82.88	Poor Accuracy

Case Study 02: Changing the batch size

Configuration No.	**Batch size**	**Epochs x training times**	**Test_accuracy (%)**	**Findings**
1	16	100 x 477s	82.23	Poor Accuracy
2	**32**	**100 x 491s**	**84.87**	**Best Accuracy**
3	64	100 x 471s	83.26	Good Accuracy

Case Study 03: Changing the Loss Function

Configuration No.	**Loss Functions**	**Epochs x training times**	**Test_accuracy (%)**	**Findings**
1	**Categorical Cross entropy**	**100 x 455s**	**84.51**	**Best Accuracy**
2	Mean Squared Errors	100 x 451s	83.87	Poor Accuracy
3	Mean absolute errors	100 x 450s	83.23	Good Accuracy

Case Study 04: Changing the Optimizer

Configuration No.	Optimizers	Epochs x training times	**Test_accuracy (%)**	**Findings**
1	**Adam**	**100 x 491s**	**85.28**	**Best Accuracy**
2	Nadam	100 x 490s	81.88	Good Accuracy
3	SGD	100 x 492s	82.51	Good Accuracy

Case Study 05: Changing the Learning Rate

Configuration No.	Learning rates	Epochs x training times	**Test_accuracy (%)**	**Findings**

1	0.01	100 x 451s	81.8	Good Accuracy
2	**0.001**	**100 x 458s**	**88.97**	**Best Accuracy**
3	0.0001	100 x 455s	86.2	Good Accuracy

The ablation study encompassed five crucial case studies, each systematically exploring the impact of variations in specific model configurations on the performance of F-AgeNet, our customized CNN model for age and face detection. In Case Study 01, different polling layer types were assessed, revealing that using Flatten yielded the best accuracy at 83.26%. Case Study 02 focused on batch size variations, with a batch size of 32 leading to the highest accuracy at 84.87%. Case Study 03 investigated changes in the loss function, showcasing that employing Categorical Cross Entropy resulted in the best accuracy at 84.51%. Case Study 04 delved into the effects of optimizer choices, where Adam demonstrated superior performance with an accuracy of 85.28%. Finally, Case Study 05 explored different learning rates, with a learning rate of 0.001 yielding the highest accuracy at 88.97%. Overall, the ablation study provided valuable insights into the sensitivity of F-AgeNet to various configurations, highlighting the significance of optimizing polling layers, batch sizes, loss functions, optimizers, and learning rates for achieving superior accuracy in age and face detection tasks. These findings inform the refinement of F-AgeNet, emphasizing the importance of meticulous configuration tuning for optimal model performance in real-world applications.

5.3.2 The final configuration of the model

Following the ablation study, table 4.3 presents the proposed model's final configuration.

Table 4.3: Configuration of the proposed model

Configuration	Value
Image sizes	64x64

Epochs	100
Optimization Functions	Adam
Learning rates	0.001
Batch sizes	32
Activation functions	Softmax
Dropouts	0.5
Momentum	0.9
Accuracy	87.50

5.3.3 loss curve

The loss curve is a visual representation of how the loss, a measure of the model's performance, changes during the training process. Typically plotted against the number of training epochs, the loss curve provides valuable insights into the convergence and optimization of a machine learning model. In the context of F-AgeNet, the loss curve depicts the gradual decrease in the loss function over successive epochs, illustrating how well the model is learning to minimize the difference between predicted and actual values. A well-behaved loss curve exhibits a smooth decline, indicating effective convergence and learning. On the other hand, erratic or fluctuating curves may suggest challenges in model convergence, potentially indicating issues like overfitting or underfitting. Analyzing the loss curve in conjunction with accuracy metrics provides a holistic understanding of F-AgeNet's training dynamics, allowing for informed decisions on further model refinement or optimization strategies. The loss curve serves as a critical diagnostic tool, aiding researchers and practitioners in assessing the robustness and generalization capabilities of F-AgeNet in the task of age and face detection. The Loss Curve of this model is shown in Figure 4.2

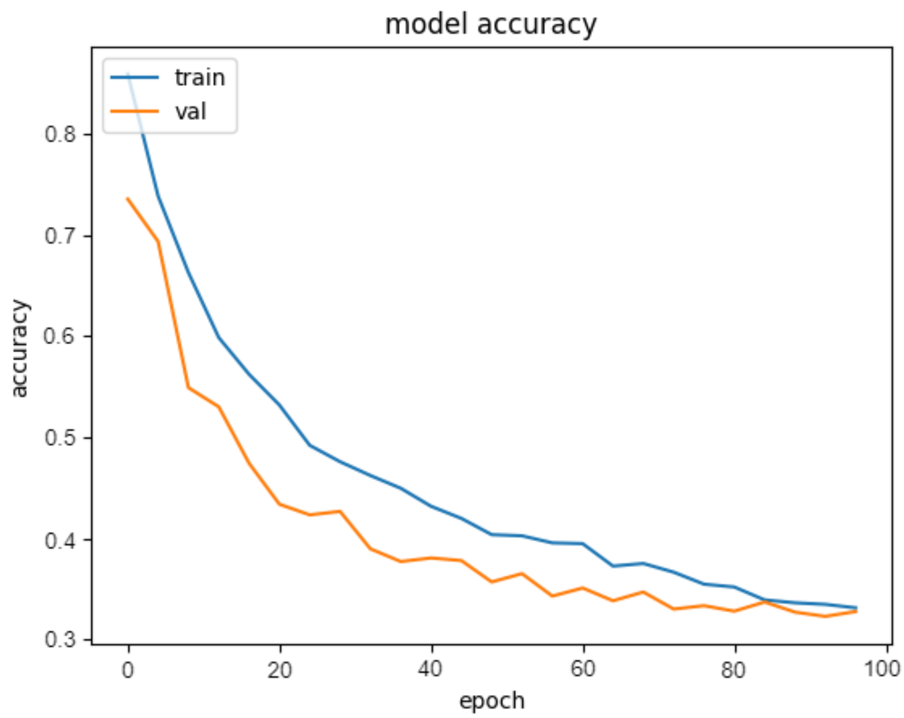

Fig 4.2. Loss curve of the proposed model

5.4 Predicted Results

The model's predicted images portion offers a visual depiction of F-AgeNet's age and face identification capabilities. Readers may see the model's predictions in practical situations and learn more about how well it can classify people into various age groups by looking through this collection of photographs. Every picture has the expected age group next to it, providing a concrete example of the model's power. This section offers a chance to highlight possible areas for improvement or difficulties the model might encounter in particular settings, in addition to providing a qualitative validation of F-AgeNet. It improves the reader's comprehension of F-AgeNet's practical application in age and face detection tasks by providing a link between the technical information previously presented and the practical consequences of the model's predictions. Furthermore, showcasing a diversified collection of images highlights the model's flexibility in a range of facial features, expressions, and lighting scenarios, enhancing its efficacy in a variety of situations. The sample predicted images from the model shown in Figure 4.3 from the public dataset and the private dataset, the figure is shown in 4.4.

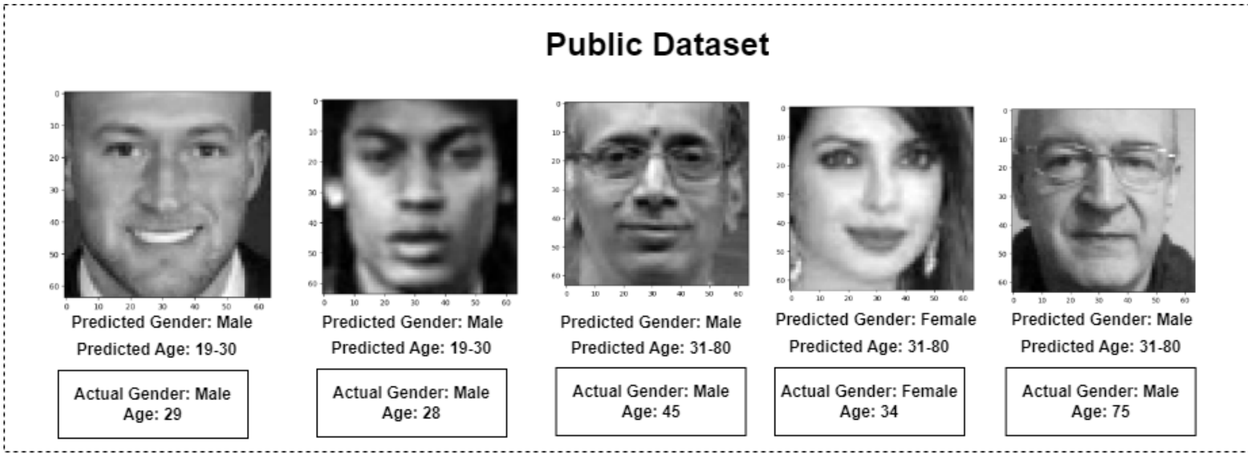

Fig 4.3: Sample predicted images in the public dataset from the model.

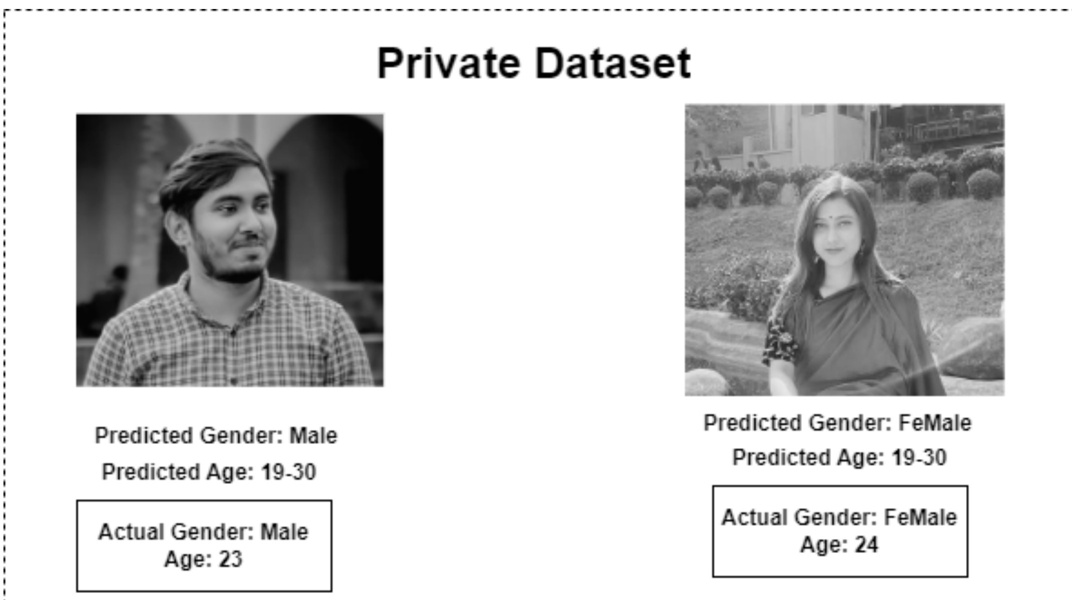

Fig 4.4: Sample predicted images in the Private dataset from the model.

5.5 Discussion

The precision of the predictions exhibited on both public and private datasets highlights the resilience and effectiveness of F-AgeNet in age and facial recognition tasks. The model's ability to estimate age groups for a variety of facial photos from various datasets consistently yields correct results, which is a strong proof of its generalization capabilities. F-AgeNet's ability to adapt to a variety of real-world settings and facial traits is demonstrated by its successful application to photographs from private datasets that had never been seen before. These encouraging findings support the model's dependability in making precise age predictions and highlight its potential for useful implementation in a range of contexts, including human-computer interaction, age-specific content delivery, and security. Carefully adjusting the hyperparameters, conducting the ablation research, and training have all helped to create a very functional and adaptable model. We are more confident in F-AgeNet's applicability for real-world applications needing accurate face and age recognition as we observe the congruence between its predictions and ground truth labels across various datasets.

CHAPTER 6
Impact on Society, Environment and Sustainability

6.1 Impact on Society

F-AgeNet has had an extensive and revolutionary effect on society, opening up new avenues for age and face identification technology. F-AgeNet's excellent precision and versatility make it a promising tool for changing many industries. It can improve security surveillance systems by effectively identifying people depending on their age, which helps with the identification and monitoring of particular age groups for public safety and security. F-AgeNet's precise age forecasts can enhance user experiences in personalized content distribution by customizing material for particular age groups. This has significant ramifications for educators, marketers, and content producers since it allows them to precisely identify the age groups of their target audience and tailor their strategies accordingly. Furthermore, F-AgeNet may have uses in the medical field for age-specific medical imaging analysis, which could lead to more individualized and focused treatment programs and diagnostic procedures. The adaptability of the model to a wide range of face traits is especially useful in multicultural and different countries, where the ability to reliably forecast age is critical to comprehending demographic trends and formulating inclusive policy. But the social impact also brings up moral questions, highlighting the necessity of responsible deployment to guard against abuse and maintain privacy. As F-AgeNet and related technologies continue to define the landscape of age and face detection, ultimately impacting how societies interact with and employ these breakthroughs, striking a balance between technological advancements and ethical considerations will become increasingly important.

6.2 Impact on Environment

F-AgeNet's environmental effects are complex, taking into account factors like energy use, processing power, and the wider ecological imprint that comes with technological progress. During the F-AgeNet training and deployment phases, energy consumption is one of the most important environmental factors to take into account. Since F-AgeNet is a deep learning model, it must go through a lot of computing operations during training, which frequently calls for sophisticated hardware that uses a lot of energy. Although the model has proven to be more accurate in detecting age and faces, there are sustainability concerns due to the environmental impact of training deep neural networks. To reduce the overall energy consumption during training, the architecture and hyperparameters of the model should be optimized to achieve efficient convergence and lessen the environmental impact of F-AgeNet. The carbon footprint of model training can also be reduced by investigating and using energy-efficient technology, such as Graphics Processing Units (GPUs) made with an emphasis on performance per watt. The environmental impact of energy extraction and fabrication of hardware components necessary for deep learning model execution goes beyond energy considerations. Energy-intensive manufacturing procedures, the extraction of raw materials, and the production of electronic waste are all involved in the production of GPUs and other computer units. The environmental effect of hardware creation and disposal can be reduced by implementing strategies that prioritize the recycling and appropriate disposal of electronic components.

Furthermore, the deployment of F-AgeNet in practical applications necessitates taking into account the energy usage of the devices that the model operates on. For example, the energy efficiency of mobile applications or surveillance systems becomes critical in evaluating the total environmental effect if the model is integrated into these systems. F-AgeNet's societal consequences, including its ability to optimize individualized content delivery and support more focused marketing tactics, have the potential to impact consumer behavior on a larger scale. Personalized content can improve user experiences but given the energy consumption of data centers and network infrastructure, it may also increase digital consumption and its accompanying environmental impact.

6.3 Ethical Expects

The ethical dimensions of age and face detection technologies, exemplified by models like F-AgeNet, necessitate a thoughtful and principled approach. Privacy considerations are paramount, requiring robust safeguards against unauthorized access and potential misuse of facial data. Mitigating biases to ensure fairness across diverse demographic groups is essential, as is securing informed consent from users regarding the purpose and implications of facial data utilization. Transparency in communication about the model's functionality, limitations, and risks is crucial, accompanied by mechanisms for accountability in case of unintended consequences. Striking a balance between security measures, long-term impact assessments, and environmental sustainability further contributes to ethical model deployment. Continuous monitoring, regular updates, and adherence to international standards and regulations collectively form the foundation for responsible development, fostering trust and societal benefit while minimizing potential risksz

6.4 Sustainability Plan

The sustainability plan for the deployment of age and face detection technologies, exemplified by F-AgeNet, involves a multifaceted approach to minimize environmental impact and ensure responsible practices. Energy efficiency is prioritized during model training through the use of hardware with lower consumption and exploring cloud services committed to sustainability. Data efficiency and minimization practices reduce the overall environmental footprint by focusing on relevant and representative datasets. Investment in green computing infrastructure powered by renewable energy aligns the model's deployment with environmentally friendly resources. Continuous optimization efforts, including refining algorithms and adopting state-of-the-art techniques, contribute to operational efficiency and resource conservation. Lifelong learning capabilities enable the model to adapt over time without extensive retraining, minimizing the environmental impact of frequent updates. User education promotes responsible data-sharing practices, and privacy-preserving technologies, such as federated learning, reduce the need for extensive data transfers. Ethical considerations and end-of-life planning ensure responsible usage and disposal practices, while community engagement fosters collaboration and feedback on sustainability initiatives. This comprehensive plan aims to harmonize technological innovation with environmental stewardship, fostering a sustainable future for age and face detection technologies.

CHAPTER 7

Summary, Conclusion, Recommendation, And Implication for future Research

7.1 Conclusion

In conclusion, F-AgeNet emerges as a highly effective convolutional neural network, excelling in age and face detection tasks. Its superior accuracy of 88.97%, achieved through careful architectural design, hyperparameter tuning, and ablation studies, positions F-AgeNet as a robust and versatile model. The comprehensive evaluations, including comparisons with established models and successful predictions across public and private datasets, attest to its real-world applicability. The ablation studies played a pivotal role in fine-tuning F-AgeNet, shedding light on its sensitivity to different configurations. Notably, the customization of polling layer types, batch sizes, loss functions, optimizers, and learning rates contributed to the model's exceptional performance. The qualitative validation through predicted images further underscores F-AgeNet's adaptability and potential deployment in security, content delivery, and healthcare applications. Societally, F-AgeNet introduces advancements in security by enhancing surveillance systems, optimizing personalized content delivery, and contributing to age-specific healthcare diagnostics. However, ethical considerations must accompany these advancements to ensure responsible deployment and safeguard privacy. On the environmental front, challenges arise from energy consumption during training and hardware production. Mitigating the environmental footprint involves optimizing the model's architecture, adopting energy-efficient hardware, and emphasizing responsible disposal practices.

In summary, F-AgeNet represents a significant stride in age and face detection technologies. Its success highlights the potential of customized CNN models, showcasing the delicate balance between technological innovation, ethical responsibility, and environmental consciousness. As F-AgeNet paves the way for future advancements, a commitment to ethical AI practices and environmental sustainability remains paramount in shaping the trajectory of AI technologies in our evolving society

7.2 Limitation and Future Work

While F-AgeNet demonstrates impressive accuracy in age and face detection, limitations exist that must be acknowledged. The model's performance is contingent on the representativeness of the training data, introducing the risk of biases inherent in the dataset. Generalizing to unforeseen conditions, such as diverse facial expressions or extreme lighting, may challenge the model's reliability in unpredictable settings. Ethical concerns, particularly in privacy infringement and potential misuse, underscore the need for cautious deployment. Additionally, the computational demands during training contribute to a considerable environmental footprint, prompting sustainability considerations.

To enhance F-AgeNet's capabilities and address its limitations, future research should prioritize several avenues. Strategies for mitigating data biases, such as employing fairness-aware training techniques and diverse dataset curation, will improve model fairness. Advancements in generalization, especially adapting the model to handle varied real-world conditions, will contribute to its robustness. Exploring privacy-preserving techniques, including federated learning and differential privacy, can address ethical concerns. Further efforts in model interpretability and transparency will foster trust, and adopting energy-efficient training strategies and hardware will contribute to the sustainable deployment of F-AgeNet. In summary, future work should focus on refining the model's fairness, robustness, ethical considerations, and environmental impact.

Reference

[1] G. Jain and S. Parsheera, “Cinderella’s shoe won’t fit Soundarya: An audit of facial processing tools on Indian faces,” 2021.

[2] “Mrs. Alka Das Gupta Chairperson BBD Group of Education Message from Chairperson.”

[3] U. Jayaraman, P. Gupta, S. Gupta, G. Arora, and K. Tiwari, “Recent development in face recognition,” *Neurocomputing*, vol. 408, pp. 231–245, Sep. 2020, doi: 10.1016/j.neucom.2019.08.110.

[4] X. Xu *et al.*, “Searching for Alignment in Face Recognition,” 2021. [Online]. Available: www.aaai.org

[5] K. Kärkkäinen and J. Joo UCLA, “FairFace: Face Attribute Dataset for Balanced Race, Gender, and Age for Bias Measurement and Mitigation.” [Online]. Available: https://github.com/joojs/fairface.

[6] P. Dhar, A. Bansal, C. D. Castillo, J. Gleason, P. Jonathon Phillips, and R. Chellappa, “How are attributes expressed in face DCNNs?,” in *Proceedings - 2020 15th IEEE International Conference on Automatic Face and Gesture Recognition, FG 2020*, Institute of Electrical and Electronics Engineers Inc., Nov. 2020, pp. 85–92. doi: 10.1109/FG47880.2020.00009.

[7] S. Gupta, K. Thakur, and M. Kumar, “2D-human face recognition using SIFT and SURF descriptors of face’s feature regions,” *Visual Computer*, vol. 37, no. 3, pp. 447–456, Mar. 2021, doi: 10.1007/s00371-020-01814-8.

[8] K. R. Hassan and I. H. Ali, “Age and Gender Classification using Multiple Convolutional Neural Network,” in *IOP Conference Series: Materials Science and Engineering*, IOP Publishing Ltd, Nov. 2020. doi: 10.1088/1757-899X/928/3/032039.

[9] R. Ranjan *et al.*, “A Fast and Accurate System for Face Detection, Identification, and Verification,” *IEEE Trans Biom Behav Identity Sci*, vol. 1, no. 2, pp. 82–96, Apr. 2019, doi: 10.1109/tbiom.2019.2908436.

[10] S. Kumar *et al.*, “Face Spoofing, Age, Gender and Facial Expression Recognition Using Advance Neural Network Architecture-Based Biometric System,” *Sensors*, vol. 22, no. 14, Jul. 2022, doi: 10.3390/s22145160.

[11] Z. Akhtar, M. R. Mouree, and D. Dasgupta, “Utility of Deep Learning Features for Facial Attributes Manipulation Detection,” in *Proceedings - 2020 IEEE International Conference on Humanized Computing and Communication with Artificial Intelligence, HCCAI 2020*, Institute of Electrical and Electronics Engineers Inc., Sep. 2020, pp. 55–60. doi: 10.1109/HCCAI49649.2020.00015.

[12] Y. Yang *et al.*, “Enhancing Fairness in Face Detection in Computer Vision Systems by Demographic Bias Mitigation,” in *AIES 2022 - Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society*, Association for Computing Machinery, Inc, Jul. 2022, pp. 813–822. doi: 10.1145/3514094.3534153.

[13] S. S. Khan, D. Sengupta, A. Ghosh, and A. Chaudhuri, “MTCNN++: A CNN-based face detection algorithm inspired by MTCNN,” *Visual Computer*, 2023, doi: 10.1007/s00371-023-02822-0.

[14] T. Singh, S. R. Kudavelly, and K. V. Suryanarayana, “Deep learning based fetal face detection and visualization in prenatal ultrasound,” in *Proceedings - International Symposium on Biomedical Imaging*, IEEE Computer Society, Apr. 2021, pp. 1760–1763. doi: 10.1109/ISBI48211.2021.9433915.

[15] R. L. Hsu, M. Abdel-Mottaleb, and A. K. Jain, “Face detection in color images,” *IEEE Trans Pattern Anal Mach Intell*, vol. 24, no. 5, pp. 696–706, May 2002, doi: 10.1109/34.1000242.

[16] H. Wang, V. Sanchez, and C. T. Li, “Improving Face-Based Age Estimation with Attention-Based Dynamic Patch Fusion,” *IEEE Transactions on Image Processing*, vol. 31, pp. 1084–1096, 2022, doi: 10.1109/TIP.2021.3139226.

[17] J. Liu *et al.*, “Multimodal Imaging and En Face OCT Detection of Calcified Drusen in Eyes with Age-Related Macular Degeneration,” *Ophthalmology Science*, vol. 2, no. 2, Jun. 2022, doi: 10.1016/j.xops.2022.100162.

[18] K. Kärkkäinen and J. Joo UCLA, “FairFace: Face Attribute Dataset for Balanced Race, Gender, and Age.” [Online]. Available: https://github.com/joojs/fairface}.

[19] Y. Zheng, J. Bao, D. Chen, and M. Zeng, “Exploring Temporal Coherence for More General Video Face Forgery Detection.”

[20] “Dataset.” Accessed: Jan. 09, 2024. [Online]. Available: https://www.kaggle.com/datasets/jangedoo/utkface-new

[21] I. D. Raji, T. Gebru, M. Mitchell, J. Buolamwini, J. Lee, and E. Denton, “Saving Face: Investigating the ethical concerns of facial recognition auditing,” in *AIES 2020 - Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, Association for Computing Machinery, Inc, Feb. 2020, pp. 145–151. doi: 10.1145/3375627.3375820.

[22] A. Priadana, M. D. Putro, C. Jeong, and K. H. Jo, “A Fast Real-time Face Gender Detector on CPU using Superficial Network with Attention Modules,” in *Proceedings - IWIS 2022: 2nd International Workshop on Intelligent Systems*, Institute of Electrical and Electronics Engineers Inc., 2022. doi: 10.1109/IWIS56333.2022.9920714.

[23] X. Zhu, H. Wang, H. Fei, Z. Lei, and S. Z. Li, “Face Forgery Detection by 3D Decomposition.”

[24] M. Shanmuganathan and C. Nalini, “Face Detection based on Extraction by K-NNC and NMC Framework,” in *Proceedings of the 3rd International Conference on Intelligent Sustainable Systems, ICISS 2020*, Institute of Electrical and Electronics Engineers Inc., Dec. 2020, pp. 1006–1011. doi: 10.1109/ICISS49785.2020.9315955.

[25] K. Okokpujie, A. Reuben, J. C. Ofoche, B. J. Biobelemoye, and I. Princess Okokpujie, “A COMPARATIVE ANALYSIS PERFORMANCE OF DATA AUGMENTATION ON AGE-INVARIANT FACE RECOGNI-TION USING PRETRAINED RESIDUAL NEURAL NETWORK,” *J Theor Appl Inf Technol*, vol. 31, no. 6, 2021, [Online]. Available: www.jatit.org

[26] O. A. Mohammed and J. M. Al-Tuwaijari, “Analysis of challenges and methods for face detection systems: A survey,” *Int. J. Nonlinear Anal. Appl*, vol. 13, pp. 2008–6822, 2022, doi: 10.22075/ijnaa.2022.6222.

[27] H. Dang, F. Liu, J. Stehouwer, X. Liu, and A. Jain, “On the Detection of Digital Face Manipulation.”

[28] T. Kamyab, A. Delrish, H. Daealhaq, A. M. Ghahfarokhi, and F. Beheshtinejad, “Comparison and Review of Face Recognition Methods Based on Gabor and Boosting Algorithms,” *International Journal of Robotics and Control Systems*, vol. 2, no. 4, pp. 610–617, 2022, doi: 10.31763/ijrcs.v2i4.759.

[29] S. L. Agrwal, S. K. Sharma, and V. Kant, “A Review on Unconstrained Real-Time Rotation-Invariant Face Detection,” in *2023 3rd International Conference on Intelligent Communication and Computational Techniques, ICCT 2023*, Institute of Electrical and Electronics Engineers Inc., 2023. doi: 10.1109/ICCT56969.2023.10076222.

[30] X. Chen *et al.*, “Application of EfficientNet-B0 and GRU-based deep learning on classifying the colposcopy diagnosis of precancerous cervical lesions,” *Cancer Med*, vol. 12, no. 7, pp. 8690–8699, Apr. 2023, doi: 10.1002/cam4.5581.

[31] Y. Gulzar, “Fruit Image Classification Model Based on MobileNetV2 with Deep Transfer Learning Technique,” *Sustainability (Switzerland)*, vol. 15, no. 3, Feb. 2023, doi: 10.3390/su15031906.

[32] Z. Riaz, B. Khan, S. Abdullah, S. Khan, and M. S. Islam, “Lung Tumor Image Segmentation from Computer Tomography Images Using MobileNetV2 and Transfer Learning,” *Bioengineering*, vol. 10, no. 8, Aug. 2023, doi: 10.3390/bioengineering10080981.

Appendix

Appendix (A) **VGGFace**

VGGFace[27], a prominent model in the realm of facial recognition, represents a pivotal advancement in the field, renowned for its depth and simplicity. Developed by the Visual Geometry Group at the University of Oxford, VGGFace leverages a deep convolutional neural network architecture with 16 convolutional layers. The strength of VGGFace lies in its ability to capture intricate facial features through a hierarchical structure, allowing for the extraction of nuanced patterns essential for accurate face recognition. The model's extensive training on large-scale datasets contributes to its robustness and adaptability, enabling it to handle variations in pose, expression, and lighting conditions. VGGFace has proven effective not only in face recognition tasks but also as a feature extractor for various related applications, such as age estimation and gender classification. Despite its success, the model's main drawback lies in its computational intensity due to its deep architecture, which may limit its real-time application in resource-constrained environments. The architecture of VGG16 is shown in Figure 3.3.

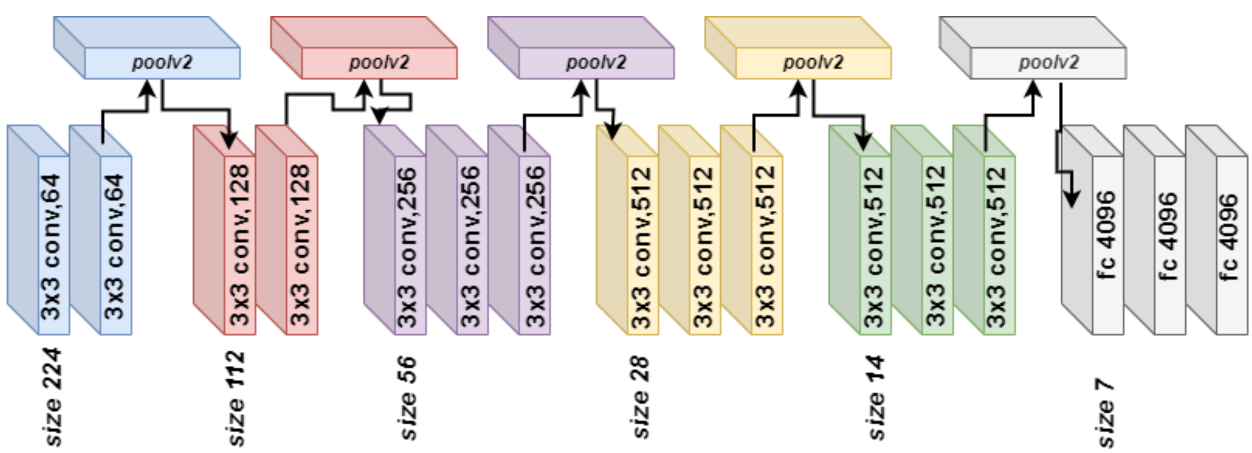

Fig 3.3: Architecture of VGGFace

As we embark on comparative analyses with F-AgeNet and other established models, acknowledging the strengths and limitations of VGGFace provides valuable insights into the evolution of facial recognition technologies and contributes to the ongoing pursuit of more efficient and accurate models in this dynamic field.

Appendix (B) **OpenFace**

OpenFace is a well-known facial recognition model that is noteworthy for its creative use of deep learning and feature extraction. OpenFace[28], created by Carnegie Mellon University researchers, combines deep neural networks with machine learning methods to offer reliable face feature representation. The model projects facial landmarks using a 3D facial model, capturing minute details that are essential for precise recognition. OpenFace is very robust in real-world circumstances since it is adept at handling a wide range of facial positions, expressions, and lighting conditions. The model's capacity to produce discriminative facial embeddings is further improved by its training use of a triplet loss function. OpenFace is renowned for its adaptability as well, as it can be used for tasks other than face recognition, such as emotion detection and facial attribute analysis. The architecture of OpenFace is shown in Figure 3.4

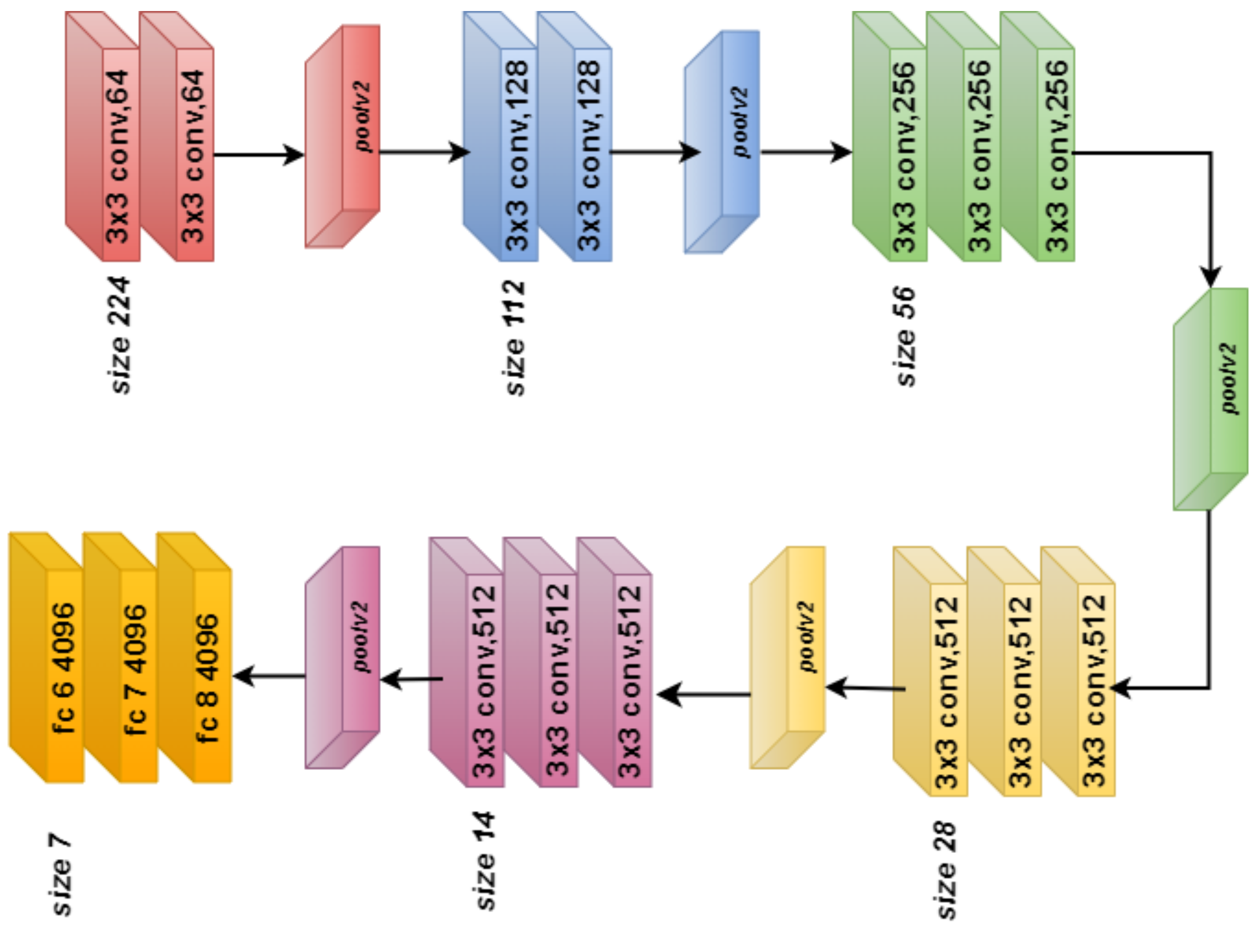

Fig 3.4: Architecture of OpenFace

OpenFace is capable of capturing face features, however it can be challenging in situations when there are occlusions or drastic variances. Furthermore, in contexts with limited resources, the computational complexity of the model may present challenges. Recognizing OpenFace's advantages and disadvantages helps us gain a deeper knowledge of how face recognition technologies are developing when we compare them to other well-known models like F-AgeNet.

Appendix (C) **DeepFace**

One important step forward in the search for reliable and adaptable face-related task solutions is DeepFace,[29] a trailblazing facial recognition model. DeepFace, created by Carnegie Mellon University academics, can leverage deep neural networks and other machine learning methods. Interestingly, DeepFace takes a unique approach by including a 2D facial model, which enables accurate facial landmark projection. This breakthrough makes DeepFace extremely flexible to real-world circumstances by capturing minute face characteristics that are essential for precise recognition across a range of poses, expressions, and lighting conditions. One of DeepFace's unique features is that it uses a triplet loss function in the training process. This method contributes to the model's ability to provide unique and trustworthy representations of face features by making it easier to create highly discriminative facial embeddings. DeepFace exhibits adaptability by expanding its uses to facial attribute analysis and emotion detection, going beyond conventional face recognition jobs. The architecture of DeepFace is shown in Figure 3.4

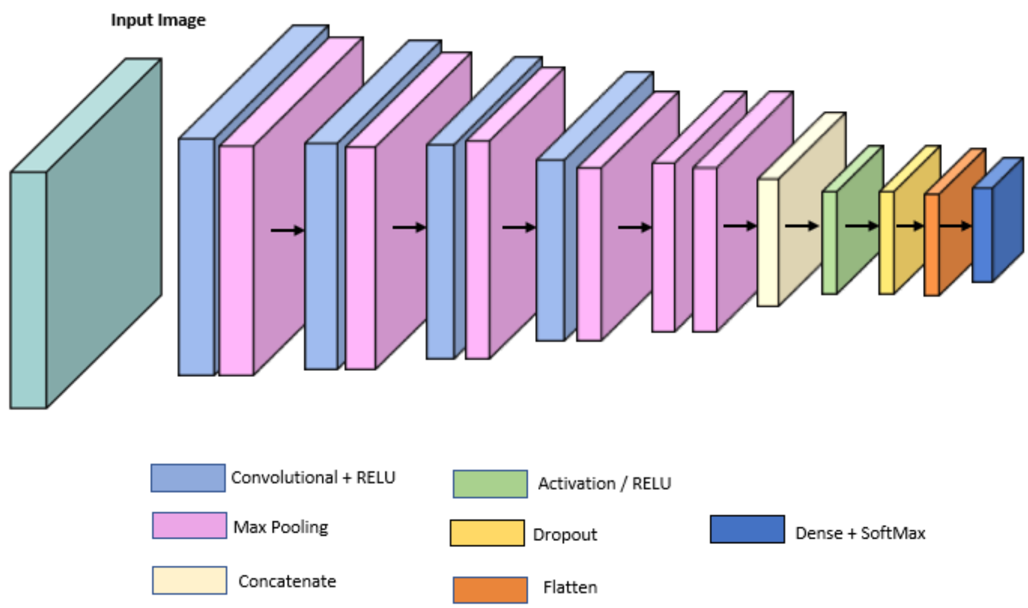

Fig 3.5: Architecture of DeepFace

Appendix (D)**EfficientNet**

In the field of convolutional neural networks (CNNs), EfficientNet[30] is a ground-breaking model that presents a fresh method for obtaining optimal performance while preserving computational efficiency. To attain higher accuracy over a range of computational resources, Google researchers developed EfficientNet, which uses a compound scaling strategy that methodically balances network depth, width, and resolution. With the help of this creative scaling technique, EfficientNet can optimize model parameters and guarantee that a larger model size enhances performance without requiring undue processing overhead. EfficientNet stands out for its ability to perform at the cutting edge of image classification tasks while using a remarkably smaller number of parameters when compared to more conventional models. Because EfficientNet carefully balances computational cost and network complexity, it is an excellent choice for contexts with limited resources. The architecture of EfficientNet is shown in Figure 3.4

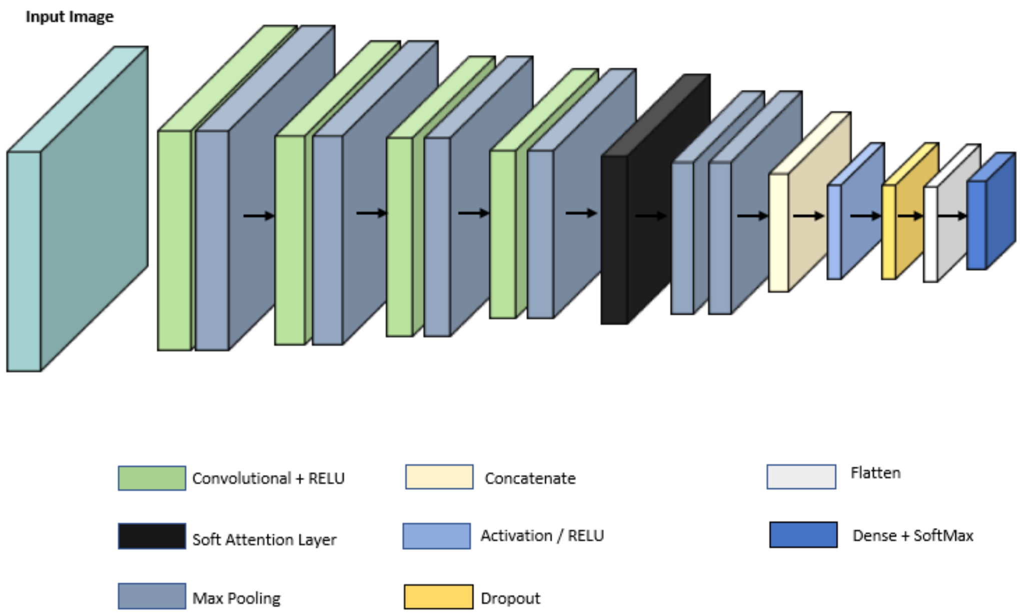

Fig 3.6: Architecture of EfficientNet

Appendix (E)**MobileNet V2**

With a focus on effective handheld consumption, MobileNetV2[31], a lightweight and optimized convolutional neural network, represents a significant improvement over its predecessor, MobileNetV1. Designed to minimize computational complexity and achieve high accuracy, MobileNetV2[31] is a more efficient architectural evolution. In contrast to its predecessor, MobileNetV2 exhibits enhanced accuracy on identical hardware, rendering it a perfect option for environments with limited resources. MobileNetV2's architecture places equal emphasis on improved resilience against adversarial assaults and input noise as it does on greater accuracy. Because of this robustness, the model can be used in situations where it is subjected to a variety of potentially difficult real-world conditions. Furthermore, MobileNetV2 is especially well-suited for applications like picture segmentation and object detection because it is built to maximize battery life and reduce latency on mobile devices. MobileNetV2 is composed of two kinds of blocks, each with three levels, and each block has eleven convolutional layers. Each block's first, third, and second layers have 32 filters[32] that are carefully crafted to collect and handle complex features. The longitudinal bottlenecks between layers are noteworthy; this is an important characteristic that keeps a significant amount of data from being negatively impacted by non-linearity. The two blocks' distinct strides—block 2 has a stride of two, while block 1 has a stride of one—address the network's flexibility and effectiveness in processing a variety of geographical data. The architecture of MobileNetV2 is shown in Figure 3.7

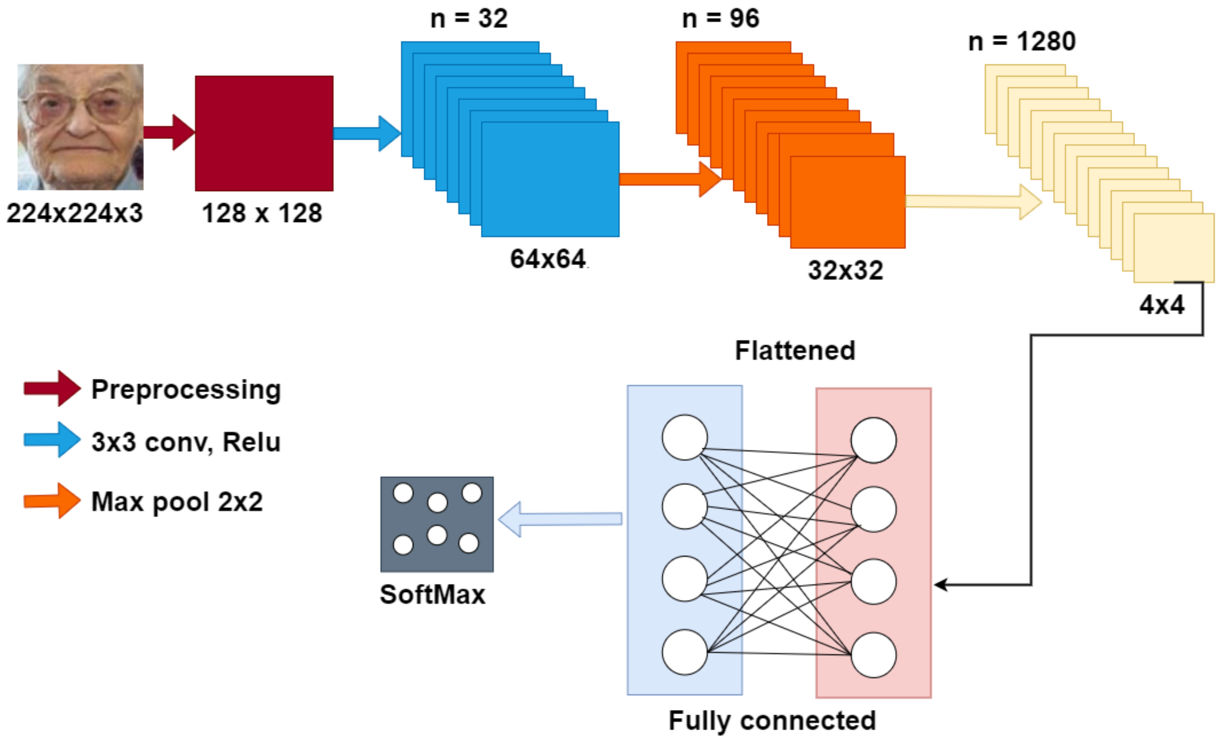

Fig 3.7: Architecture of MobileNetV2

Appendix(F) Plagiarism Report

Deep CNN Model: A case study of predicting security surveillance activities utilizing Gender and Age

ORIGINALITY REPORT

17%	15%	7%	8%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	dspace.daffodilvarsity.edu.bd:8080	Internet Source	8%
2	www.researchgate.net	Internet Source	1%
3	Submitted to Daffodil International University	Student Paper	<1%
4	Submitted to University of Leicester	Student Paper	<1%
5	link.springer.com	Internet Source	<1%
6	Submitted to University of Southern Mississippi	Student Paper	<1%
7	Submitted to unibuc	Student Paper	<1%
8	pesquisa.bvsalud.org	Internet Source	<1%

Submitted to University of Hertfordshire