NEURAL NETWORK WITH AGNOSTIC META-LEARNING MODEL FOR FACE-AGING RECOGNITION

RASHA RAGHEB ATTAALLAH

FACULTY OF COMPUTER SCIENCE AND
INFORMATION TECHNOLOGY
UNIVERSITI MALAYA
KUALA LUMPUR

2022

NEURAL NETWORK WITH AGNOSTIC META-LEARNING MODEL FOR FACE-AGING RECOGNITION

RASHA RAGHEB ATTAALLAH

THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

FACULTY OF COMPUTER SCIENCE AND INFORMATION TECHNOLOGY
UNIVERSITI MALAYA
KUALA LUMPUR

2022

UNIVERSITI MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: RASHA RAGHEB ATTAALLAH

Matric No: WHA150072

Name of Degree: Doctor of Philosophy

Title of Project Paper/Research Report/Dissertation/Thesis ("this Work"):

AGE INVARIANT FACE RECOGNITION USING META-COGNITIVE NEURAL NETWORK

Field of Study: Human Computer Interaction

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature Date: 1st April 2022

Subscribed and solemnly declared before,

Witness's Signature Date: 1st April 2022

Name:

Designation:

NEURAL NETWORK WITH AGNOSTIC META-LEARNING MODEL FOR FACE-AGING RECOGNITION

ABSTRACT

Face recognition is one of the most popular and accessible verification techniques. It is also accepted by users as it is non-invasive. Nevertheless, the aging process may change the face shape and texture. Therefore, aging is considered one factor that affects the accuracy of face recognition applications. Existing techniques and methods for face aging recognition are degraded in performance due to many factors, such as the uncontrolled nature of aging processes of the human face. Thus, there is a need to further investigate face aging recognition techniques, particularly the one which can be used to compare two images of faces for the same person at a different age. This study aims to design and model a framework to recognize face aging based on artificial neural networks and Model-agnostic meta-learning (MAML), using parameters obtained from identical tasks with certain updates on these parameters. The thesis presents the main parts of the research methodology. The literature review and problem extraction are reviewed in phase 1. Phase 2 explains the research objectives. Phase 3 shows the proposed model design and implementation. Finally, phase 4 shows the analysis of the results based on the proposed framework.

The methodology starts with phase one, which includes the literature review and problem extraction. During this phase, a review of existing research was done, followed by the extraction of all face recognition challenges. This helps in defining the problem statement. The second phase includes the determination of the research objectives and explaining them. The third phase proposes the framework, shows the design and the implementation. The fourth phase explains the results, demonstrates the analyses and the evalutes the data.

The framework is evaluated using three datasets i.e., CALFW, AT&T and MEDS. It shows a good result as compared to the previous work. The accuracy for CALFW dataset is 90.40%, AT&T dataset is 96.20%, and for MEDs dataset is 86.5%. The confusion matrices for the three datasets are calculated and the results show that for the first dataset CALFW the specificity is 0.9333, precision is 0.935 and the false negative rate is 0.12280. For the second dataset AT&T the specificity is 0.960, precision is 0.9620 and the false negative rate is 0.03797, and for the third dataset MEDs the specificity is 0.9090, precision is 0.83 and the false negative rate is 0.144. This study has demonstrated that the proposed framework can be used to detect face aging. The novelty of the research is to detect the same face at different age, which means more than 10 years' difference between the first image and the second image for the same person with high accuracy and accepted results.

Keywords: Face Aging, Face Recognition, Artificial Neural Network, Meta Learning, CALFW.

RANGKAIAN SARAF DENGAN MODEL META-PEMBELAJARAN AGNOSTIK UNTUK PENGIKTIRAFAN PENUAAN WAJAH

ABSTRAK

Pengecaman wajah adalah salah satu teknik pengesahan yang paling popular dan mudah diakses. Ia juga diterima oleh pengguna, kerana ianya tidak invasif. Walaupun begitu, proses penuaan boleh mengubah bentuk dan tekstur wajah. Oleh itu, penuaan dianggap sebagai salah satu faktor yang mempengaruhi ketepatan aplikasi pengecaman wajah. Prestasi bagi teknik dan kaedah yang sedia ada untuk pengesanan penuaan wajah akan merosot kerana pelbagai faktor seperti sifat proses penuaan yang tidak terkawal yang berlaku pada wajah manusia. Oleh itu, terdapat keperluan untuk menyelidik teknik pengenalan penuaan wajah dengan lebih mendalam, terutama teknik yang dapat digunakan untuk membandingkan dua imej wajah orang yang sama pada tahap usia yang berbeza. Kajian ini bertujuan untuk membina kerangka untuk mengenali penuaan wajah berdasarkan jaringan saraf buatan dan meta-pembelajaran Model-agnostik (Model-agnostic meta-learning (MAML) yang menggunakan parameter yang diperoleh dari tugas yang sama dengan pengemaskinian yang tertentu pada parameter tersebut. Tesis ini mengemukakan bahagian-bahagian utama metodologi penyelidikan. Fasa 1 menghuraikan kajian literatur dan pengekstrakan masalah. Fasa 2 menerangkan mengenai objektif kajian. Fasa 3 pula, mempamerkan reka bentuk dan pelaksanaan model yang dicadangkan. Hasil analisis terhadap kerangka yang dicadangkan diterangkan di dalam Fasa 4. Kerangka ini dinilai menggunakan tiga set data iaitu CALFW, AT&T dan MEDS. Hasilnya, ia menunjukkan dapatan yang baik berbanding yang sebelumnya. Ketepatan untuk set data CALFW adalah 90.40%, set data AT&T adalah 96.20% dan untuk set data MED adalah 86.5%. Matriks kekeliruan untuk ketiga set data dikira dan hasilnya menunjukkan bahawa untuk set data pertama CALFW pengkhususannya adalah 0.9333, ketepatan 0.935 dan kadar negatif palsu adalah

0.12280, manakala untuk set data kedua AT&T pengkhususannya adalah 0.960, ketepatan adalah 0.9620 dan negatif palsu pula kadarnya ialah 0,03797, dan bagi data ketiga iaitu MED, pengkhususannya adalah 0,9090, ketepatan adalah 0,83 dan kadar negatif palsu adalah 0,144. Kajian ini telah membuktikan bahawa kerangka yang dicadangkan dapat digunakan untuk mengesan penuaan wajah. Pembaharuan dalam penyelidikan ini adalah untuk mengesan wajah yang sama pada usia yang berbeza, yang bermaksud perbezaan lebih dari 10 tahun antara gambar pertama dan gambar kedua untuk orang yang sama dengan ketepatan tinggi dan hasil yang boleh diterima.

Kata kunci: Penuaan Muka, Pengecaman Wajah, Rangkaian Neural Buatan, Pembelajaran Meta, CALFW.

ACKNOWLEDGEMENTS

First and foremost, I would like to praise and thank Allah, the almighty, who has granted countless blessings, knowledge, and opportunity to me to finally accomplish the thesis.

Throughout the writing of this dissertation, I have received a great deal of support and assistance.

I would like to thank my supervisor, Associate Professor Dr MAIZATUL AKMAR BINTI ISMAIL. Associate Professor Dr AMIRRUDIN BIN KAMSIN. for guidance and contribution of time, ideas and energy in making my PhD experience productive and stimulating.

I would like to thank the Faculty of Computer Science and Information Technology, the University of Malaya, for providing me with an excellent academic environment and also offering me a well-equipped laboratory.

First and foremost, I have to thank my father, and I will be forever grateful for my mother's assistance and I am sorry that my father and my mother have not lived to see me graduate for their love and support throughout my life.

Thank you both for giving me the strength to reach for the stars and chase my dreams. My sisters (Rana, Haneen, Randa), my brother(Hossam), uncles (Raed , Manar , Adel) and my aunt (Amal , Adela , Ratiba) for their precious support, patience and assurance throughout my education at the University of Malaya.

Special thanks to all my friends, especially Dr Ahmad Sami Al.Shamayleh and Dr Sherin Abdelrahman and all of those who supported me during the completion of my research.

Also, I would like to express my deepest gratitude and appreciation to Yayasan Khazanah, Malaysia, for the financial support of my study during previous years.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

ANN	:	Artificial Neural Network
MAML	:	Model-agnostic meta-learning
MFR	:	Meta Face Recognition
RNN	:	Recurrent Neural Network
CALFW	:	Cross-Age LFW
FPR	:	False Positive Rate
FNR	:	Fales Negative Rate
SLR	:	Systematic Literature Review
WoS	:	Web of Science
SVM	:	Supporting vector machines
CNN	:	Convolutional Neural Networks
MEDs	:	Middle East Data Set MEDS

CHAPTER1: INTRODUCTION

1.1 Introduction

The human face is regarded as a valuable source of information since it reveals personal traits such as expression, gender, age, race, and identity.Face-based Human-Computer Interaction (HCI) has been widely utilized to replicate the facial information displayed in human-to-human communication (Geng et al., 2006).Also, one of the most significant human unique biometrics is the facial feature; this unique characteristic plays a critical part in determining human identity and emotions. It is regarded as a hotbed of activity in the field of facial technology. The nose, eyes, lips, and wrinkles are among these characteristics, as are advanced qualities like gender and mood.Facial biometric is still the gold standard in biometrics. This is because it is simple to deploy and implement. There is no direct physical contact with the user. Furthermore, the verification/identification processes for face recognition and face match are quick. As a result, face recognition systems have become one of the most effective biometric techniques that have gotten much attention so far (Jain et al., 2016). It is utilized in various real-world applications, including both public and private .Finding missing children, monitoring criminals, and face aging recognition are just a few examples (Rothe et al., 2018). In addition, the 2020 Olympic Games in Tokyo have a major focus on facial recognition. This technology is used to recognize and grant access to authorized individuals automatically. Face recognition is also essential in identification card de-duplication, which prevents a person from acquiring numerous ID cards under various names, such as driver's licenses and passports (Gelb & Metz, 2018). Face recognition may also be used in human-computer interaction, surveillance systems, secure access control, video conferencing, financial transaction systems, forensic applications, pedestrian detection, driver alertness monitoring systems, and picture database management systems (Clements & Smith, 2018).

Face recognition in pictures is a challenge . Face recognition is significantly more difficult for photos with variations in the backdrop, stance, orientation, lighting conditions, and facial emotions; although several face recognition algorithms existed before 2001. Therefore, finding a technique to extract fast-computing, discriminating, and consistent features is the most significant element of the problem in identifying face patterns (Chan et al., 2015).One of the problems of face recognition is the aging of the face. It is the gradual accumulation of changes through time and how age differs from one person to another. Our genes, environmental factors, and lifestyle all have a role in the form and texture of our skin as we age (Akhtar et al., 2015).In addition, prior research has revealed that face recognition models are vulnerable to various facial postures. The face posture shifts when a person's head moves or the viewing angle shifts (Gao et al., 2020).

1.2 Research Motivation

There are different challenges involved in face recognition, including enhancing the performance, process time, various poses of face, wearing glasses and aging. These challenges affect the accuracy, performance, and false rate of the face recognition models. In addition, aging is one of the challenges that affect the recognition rate because the natural process could not be controlled, and this aging process differs from one person to another.

Many researchers used different efforts to improve face recognition accuracy and performance. So, recognition under aging is still a challenge because people mightlook older or younger than the actual age (Elmahmudi & Ugail, 2019). As a result, when people grow up, the appearance of their faces can be very different, making the recognition process more difficult not only for machines but even for human (Azmeen & Borah, 2021).

Different models are based on various techniques used for face aging recognition. However, these models provide low accuracy and performance. A model based on Artificial Neural Networks was used for aging recognition with 65% accuracy (Owayjan, 2016). In another model based on Viola-Jones, the accuracy was 82%. (Murphy, 2016). An extra ANN model provided 38.177% accuracy (Panjaitan et al., 2018). Moreover, another Neural Network face aging recognition model gave 82% accuracy (Hussein et al., 2019). One more model based on convolution neural networks tested for face aging recognition with an accuracy of 57.18%. The face recognition model is based on deep learning neural network and meta-learning techniques. The model is tested with a small sample dataset from 6 faces. The recognition rate of the face can reach 92.6% (Peng & Zheng, 2021). The Meta Face Recognition (MFR) framework to solve generalized face recognition was built based on meta-learning; the framework was used for face recognition under LFW with an accuracy of 93%. Nevertheless, this framework did not test for the aging face (Faraki et al., 2021).

The previous studies show that Model-Agnostic Meta-Learning (MAML) was used for face recognition but was not applied for face aging recognition. And even there are applied models used for face aging recognition, the accuracy is still low. Even Though it is satisfactory that many works have been done on building different models for face aging recognition, aging is still an issue with many gaps. In this proposed study, a practical face recognition model was built based on ANN and enhanced MAML techniques to detect the face of the same person but in different time gaps. This time gap is more than ten years, detecting the aging face.

1.3 Open Issues

One of the friendly authentication methods that users accept is the face recognition. Also, face aging is recognized as age progression, which is one of the face recognition challenges. Previous research has revealed that aging has a significant impact on facial biometrics due to changes in shape and texture. Therefore, face recognition is confronted with the face aging problem. This challenge can be attributed to the uncontrolled nature of the aging process. However, this runaway process leads to different facial changes from one to another, despite being of the same age (Antipov et al., 2017a; Zhang et al., 2017). Moreover, the aging process differs from person to person; thus, it attracts researchers' attention. As a result, there are different methods to solve the aging problems. Also, the aging process is hard to model because it depends on genes and the environment (Wang et al., 2016). However, many approaches discard personalized information, and all people share the same aging pattern (Wang et al., 2016).

Many neural network algorithms and techniques have been effectively used for human face recognition. These techniques have their strengths and their limitations. Vanishing gradients is one of the Recurrent Neural Network (RNN) limitations. The gradients have essential data that can be used in the RNN parameter; any loss in these gradients means losing features (Xu et al., 2015). These lost features can effectively train the models, especially face recognition models.

This thesis focuses on Artificial Neural Network models and how the proposed approach can handle different issues of architecture and settings. The architectures and settings problems include classification, regression, and policy gradient reinforcement learning.

The artificial neural network (Vedel et al.) approach applies minimal modification. The implemented meta-learning approach aims to train ANN meta-learning to extract local and global features; this increases the probability of recognizing the person at different ages. Another challenge faced the face recognition techniques is that it has been found that the methods fail to recognize people's faces (Oloyede et al., 2020). In addition, plastic surgeries performed on human faces change skin texture between photographs of the same person, making facial identification difficult (Sabharwal & Gupta, 2019). Another challenge facing face recognition is low-resolution images, especially in recognizing faces. Various cameras can be used to obtain low-resolution pictures, and it is well known that extracting information from low-resolution pictures is difficult (Zangeneh et al., 2020). Face recognition is also hampered by the fact that the same individual might have a variety of facial expressions (Peña et al., 2020). Occlusion is extremely difficult to face identification, and it is one of the essential difficulties in face recognition models (Zeng et al., 2020). Face recognition has encountered a variety of obstacles, and it is still progressing. One of the most hotly debated topics is the subject of aging. As stated in the research motivation section, the goal of this study is to aid in solving this problem.

1.4 Problem Statement

The face recognition models have several gaps, as mentioned in the previous section. Furthermore, it clarifies from the open issues that it is essential to enhance the face images in the preparation stage for the image segmentation process. Additionally, the automated face recognition model performance for peoples' faces while aging is an open issue. Therefore, it is seen from the previous studies that there is a need to improve the automated face aging recognition performance.

The aging problem in the face recognition model is the apparent variation in human faces. When a significant difference in time occurs between the targeted facial picture of the same person, the variance arises (Oloyede et al., 2020). Low aging accuracy has been seen in the literature when there is a substantial age difference between the query and target images. Individuals' general face shape can be significantly influenced by their age (Bouguila & Khochtali, 2020). Furthermore, as people become older, their distinctive face features might alter (Merler et al., 2019). Also, extracting the facial characteristics is required to improve face aging recognition accuracy. These problems are regarded as open research issues in image processing. These were the issues discussed in the thesis. In other words, the issue is the lack of high performance for faces aging recognition, especially for more than ten years gap between ages.

1.5 Research Questions

The research questions are:

1. What are the limitations of existing face aging recognition techniques?
2. How to develop an efficient model for face aging recognition that can detect the same face but in different time gaps?
3. How to evaluate the performance of the proposed model?

1.6 Research Aim and Objectives

The research aims to build a model for face aging recognition with acceptable accuracy, which can detect the same person's face with a differance of more than ten years.

The objectives are as follows.

1. To identify the limitations of existing frameworks and techniques used in face aging recognition.
2. To develop a unified model for face aging recognition using an artificial neural network with (Vedel et al.) and enhancement MAML.

3. To evaluate the proposed face aging recognition framework by calculating the accuracy and confusion matrix.

1.7 Research Scope

The scope of the proposed model was limited to the enhancement of face aging recognition and facial feature extraction. These two problems are the main contributors to achieving better accuracy. Hence, the scope of the proposed model was limited to the two problems.

The proposed model includes the following scopes:

1. The model is based on ANN as a classifier. ANN is used because it achieved high recognition accuracy.
2. The scope is limited to enhance the recognition aging rate.
3. Solve face aging recognition.
4. The model uses only image extension PGM.
5. The model is built by Matlab 2014a edition to validate the model and present the result. It is used to create a face recognition model that can recognize the face in real-time.

On the other hand, the proposed model does not include the following criteria:

1. The minimum age for the persons in the dataset is 18, not less than this as the facial feature before 18 years is changeable and only becomes invariable after 18 years.
2. The model cannot read any images unless it is image extension PGM.

1.8 Research Methodology

The research technique used in this thesis is described in this section. Figure 1.1 depicts the steps of the study technique. The literature review and problem extraction until formulating the problem statement are reviewed in Stage 1, identified as ‘Problem Formulation’. Stage 2 explains the proposed model. Stage 3 demonstrates the Experiments and Results. Finally, Stage 4 shows the Evaluation; presents the analyses and evaluation.

Stage 1: Literature Review and Problem Statement Formulation

This thesis focused on developing a model for face aging recognition. Consequently, the undertaken literature review focused on the following related works:

1. Face recognition and Databases: The thesis review 23 databases used to assess various face recognition framework. Also, it shows the age gaps in the different datasets.
2. Artificial neural networks models: In the thesis, various ANN Face Recognition frameworks have been discussed because it is understood that several other frameworks are used for face recognition with different techniques.
3. Review for current aging systems: during thesis writing, studying the various systems used for human face aging and age progression.
4. After that, extract the challenges and open issues for face recognition to be identified from the literature review.
5. Finally, the problem statement is formulated; the research questions and the research objectives are determined. The primary goal of this thesis is to create an effective face-aging recognition model. The objectives determined are methodically followed until the main goal is achieved.

Stage 2 Proposed Model

This stage presents the whole design and implementation of the proposed model based on ANN combined with modified MAML. The proposed model has four phases;

pre-processing, feature selection using wavelet transforms, training based on modified MAML technique, and finally, the classifier phase based on ANN. The detail of the proposed model is discussed in Chapter 4 of this thesis. Also, in the same chapter, the model implementation was discussed. Various tools used to implement the concept were covered in the same chapter.

Stage 3: Experiments and Result

The third stage presents the experiments and the results. The model was experimented under three datasets Cross-Age LFW, CALFW and AT&T, own collected dataset.

Stage 4: Evaluation

In Chapter 5, the results of the tests are presented and analyzed to assess the suggested model's performance. The various models are evaluated using standard evaluation metrics from the literature, including Accuracy, Precision, Recall, F-measure, and False Positive Rate (FPR).

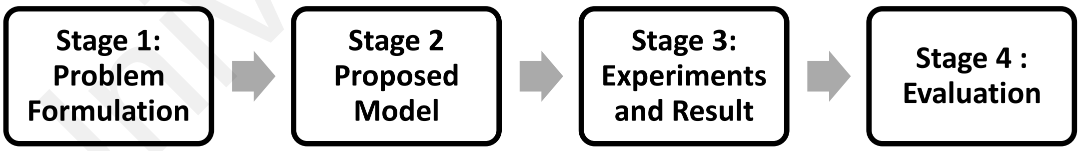

Figure 1.1: The research methodology stages. It has four stages

1.9 Research Contributions

The following are the thesis' key contributions:

1. The effective face-aging recognition framework adopts an ANN combined with model-agnostic meta-learning (MAML). Based on the essential human face

characteristics, the classifier extracts effective features to improve the accuracy of the network.

2. The framework effectively extracts the same features in the different datasets and obtains robust dependent features.

3. The framework assumes multiple features that correspond to the same person at different ages.

The significance of the ANN-MAML framework is to detect the same face at a different age. In other words, it detects the same face with more than ten years difference between the first image and the second image of the same person, and to achieve this goal, the study improves Artificial Neural Network (Vedel et al.) by adding Adaptation Model-Agnostic Meta-Learning (MAML) for face-aging recognition.

1.10 Thesis Outline

Figure 1.2 shows the thesis structure, highlighting the essential activities involved and their respective chapters. The backdrop for the investigation and the significance of the study are presented in **Chapter 1**. It also covers open issues such as study motivation, problem statement, research questions, goals, scope, technique, and contributions.

Chapter 2 provides a brief literature review of the face images dataset, a review of neural network models in the previous works, and the different models for aging process recognition. Moreover, this chapter explains some terms such as image recognition, face recognition and face features. Then it presents the current approach that faces techniques problems like availability of the images, running time and aging process. Then it presents the current approaches for Artificial neural networks (Gelb & Metz), recognition models and their architecture. These models have different phases: data acquisition, pre-processing and segmentation, feature extraction, and classification. Also, it shows the Model Agnostic Meta-Learning (MAML) models and the different aging process techniques, and at the end, it gives a problems summary of current systems reviewed.

Chapter 3 – This chapter describes the research methodology flow, consisting of five phases. The first phase presents the literature review; the second phase explains the requirements; the third phase shows the development phase, the fourth is the analysis phase, and finally, the last phase is the evaluation phase. This includes a discussion of the suitability of the data collection and analysis methods used and other significant factors.

Chapter 4 - presents the framework design and implementation in detail. The ANN-MAML framework has three phases; the first step is the image pre-processing phase. The pre-processing stage has three steps: Grey-level features reshaped to a vertical and Histogram equalization. The second step is feature extraction. Finally, the third step is the recognition phase.

Chapter 5 - provides the evaluation results and discussion; it presents the results of the framework of CALFW, AT&T and Middle east dataset. Also, this chapter shows the different evaluation metrics used to evaluate the framework, such as Specificity (True Negative Rate), Sensitivity (recall or actual positive rate), False Positive Rate (FPR), Precision, F1- Score (F- measure) and Fales Negative Rate (FNR). Moreover, this chapter shows the comparison with previous works.

Chapter 6 – summarizes the thesis and its significant contributions and its flaws. The chapter also examines the thesis' research implications and makes recommendations for further research.

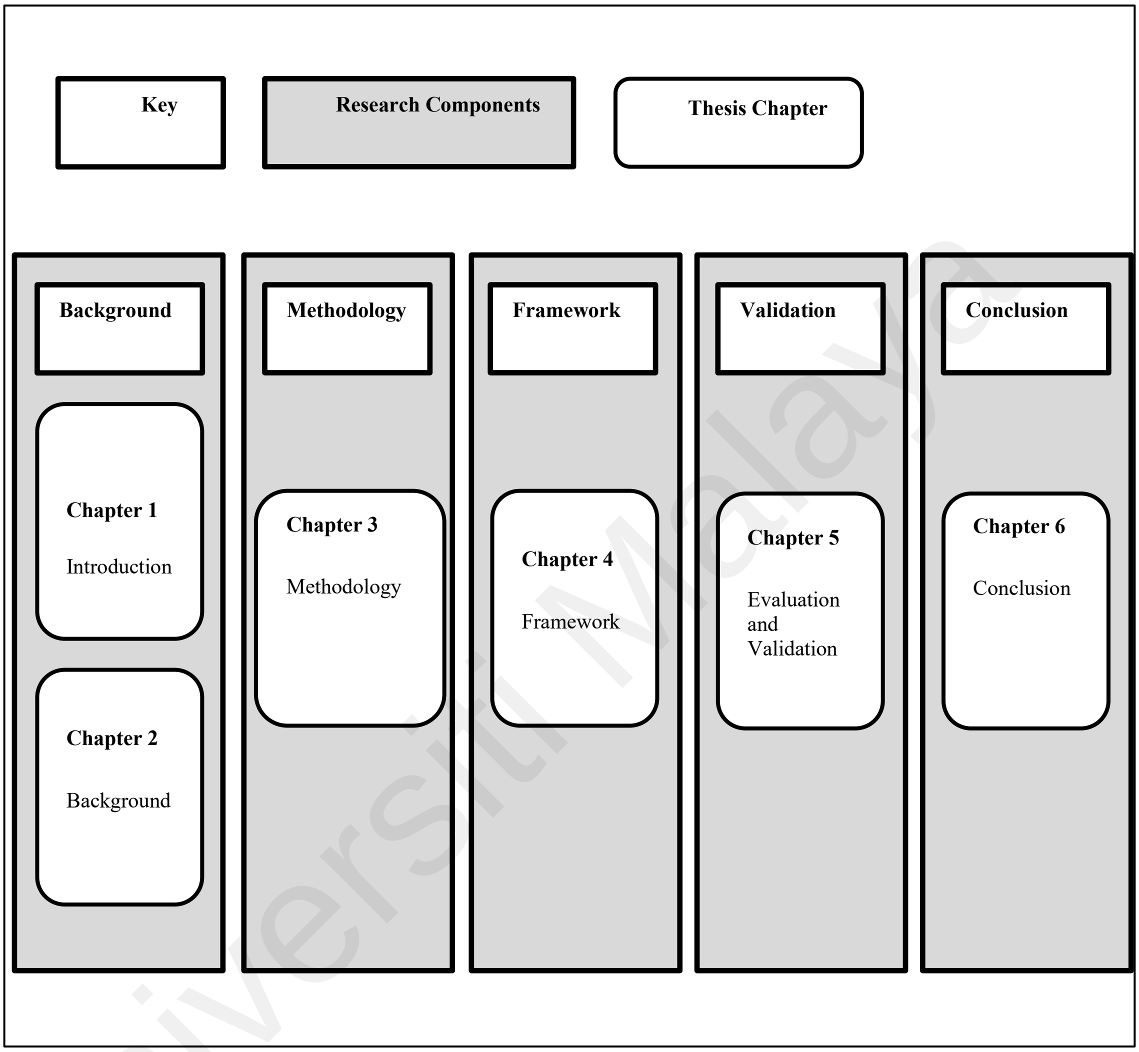

Figure 1.2: Thesis structure

1.11 Summary

Face recognition is one of the deeply studied topics in computer recognition. Face recognition is based on objective analysis and has been proposed as one of the essential biometrics. It is considered a rich source of information used in recognition systems.

This chapter presents the importance of increasing the accuracy and performance of face aging recognition systems either by modifying the algorithms or adding additional helpful algorithms at pre-processing or training stages. Also, this chapter provides the open issues of face recognition. Based on those issues, the proposed model is defined to solve the face aging issue—problem statement, research questions, objectives, and research methodology. This chapter also covers the steps of the research technique, which connect the research topic to research objectives and the study's contribution. As a result, the research becomes more apparent. It is seen as means of ensuring the success of the study. It is also used to persuade readers to agree with the findings of this research investigation. In addition, this chapter serves as a thesis outline.

CHAPTER2: LITERATURE REVIEW

2.1 Introduction

The relevance of employing face characteristics as a biometric recognition application was discussed in the previous chapter. Additionally, it presents the problem statement for this thesis which implies how to recognize the face for the same person but at different times with high performance. In addition, the questions and research goals and the methods defined in the previous chapter would aid in solving this problem. Additionally, it presents the different open issues in this field, such as face aging, plastic surgery, and low-resolution images.

In chapter 2, the challenges in face recognition models were discussed in more detail, such as the processing time, availability of images, and aging because the aim is to improve the accuracy of recognition aging models. However, the listed problems should be justified by reviewing state-of-the-art systems. As a result, this chapter evaluates existing algorithms for the open issues that have yet to be solved. This chapter presents an overview of the systematic method used to choose previous studies and papers. Theoretical concepts, after that, explains the different face recognition databases and continue to present the reviewing of face recognition models. The following section presents artificial neural network models with an explanation for the model agnostic meta-learning (MAML). Then, it summarizes the recent development in aging process techniques in face recognition. At the end of this section, a review of the current models was established. The third section shows the problems and the open issues in face recognition.

2.2 Systematic Literature Review (SLR)

1. Review of Existing Research:

This section discusses several facial recognition and recognition methods. This section also goes in-depth into the issues with facial recognition and how to overcome them. It also goes into facial recognition algorithms' methodologies, architecture, benefits, and drawbacks. Face recognition and age estimation are the essential terms in the scope of this research. Any system that does not employ the facial face in its procedures is excluded. This research is likewise limited to relevant English articles. On the other hand, the research takes into account all extracted characteristics as well as aging regions.

A. Information sources

For the review article search, five digital databases were considered, including 1) Springer database, which provides access to computer science, biomedical, and mathematical journal articles; 2) ScienceDirect database, which provides access to computer science and mathematical articles; 3) IEEE Xplore library of technical literature in engineering and technology; 4) Web of Science (WoS) service, which provides indexing of cross-disciplinary research in computer science, mathematics, and social sciences; and 5) ACM Digital Library (DL), which provides access to computer science, biomedical, and mathematical journal articles.

B. Paper Selection

The selection method consisted of looking for relevant scientific publications in the five databases listed. Iteration to remove papers and filtering were the two steps of the procedure. Irrelevant and duplicated papers were eliminated throughout the extraction process. After reading the entire material, the filtering step began.

C. Eligibility criteria

Each scholarly article that satisfies the criteria outlined in Figure3.1 was chosen. The selection and filtering criteria covered the period from 2010 to 2021. Face characteristics, face recognition, and age estimation are essential topics in the selected papers, which must be written in English.

D. Results

The initial search was conducted on five digital database websites, with the query containing one of the following keywords: face recognition, age estimation, aging, facial extract characteristic, or facial expression. The result for the first phase to select the relatable papers according to the keywords is :

1. Springer database: 100
2. ScienceDirect database: 20
3. IEEE Xplore library:280
4. Web of Science (WoS) service:350
5. ACM Digital Library (DL) :250

The total founded papers was 100+20+280+ 350 + 2250 = 1000 papers.

The second phase was to remove the duplicated papers. It was found that there were 300 duplicated papers. The remaining papers are 1000 – 300 = 700 papers.

After reading the title and the abstract, the third phase filters the papers based on the chosen criteria. The two criteria for filtering the 700 are:

1. The article is in the English language.
2. Facial characteristics, face recognition, and age estimation are some of the essential concepts.

The result for the third phase was 700 – 350 = 350. So, the relevant papers after following the criteria were 350 papers.

The fourth phase was to filter the remaining papers after reading the full text. It was found that 239 were not suitable for the literature review. So, the final papers that were used in the literature review are 111 papers.

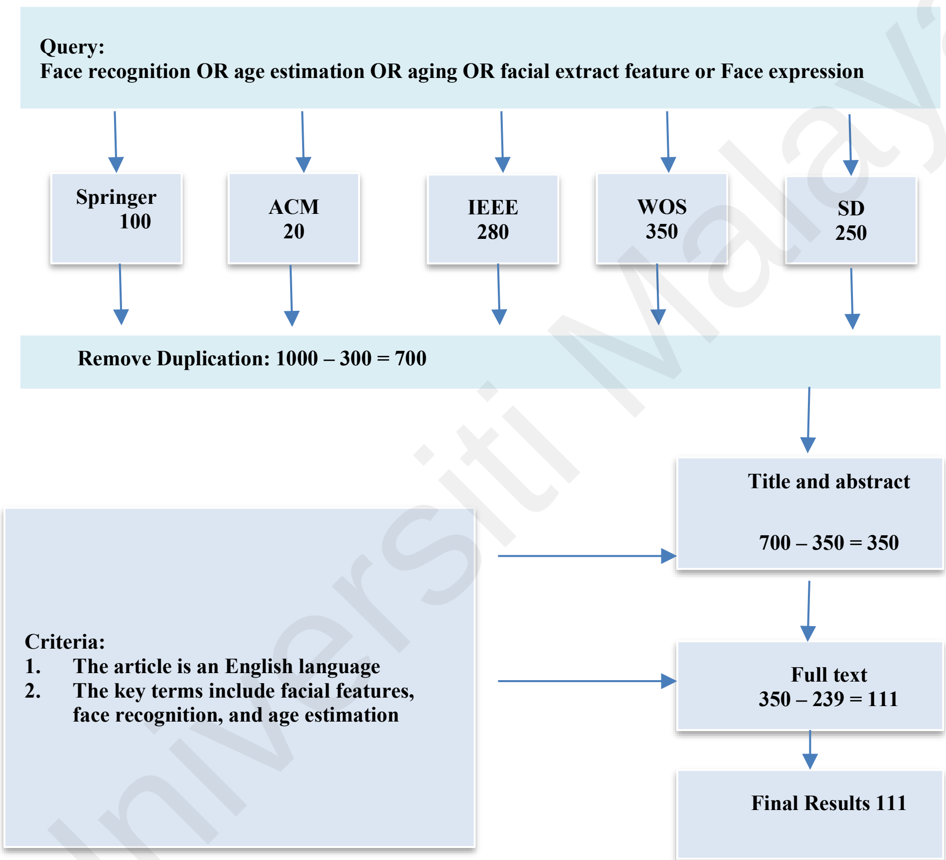

Figure 2.1: Flowchart of the study selection

The summary for using these papers at the literature review is as the following:

- A summarized table from 22 datasets.
- A summarized table for ANN techniques.

- A summarized table for aging process techniques.

2. Challenge's extraction and definition of the problem statement:

The challenges and the open issues of the face recognition systems were extracted from reviewing the 111 papers. The extracted challenges were nine open issues: availability of images, running time, low resolution, face expression, plastic surgery, occlusion, pose, the aging process, and age. From the open issues, the problem statement was defined. The problem statement was explicitly defined. As mentioned, face aging recognition was one of these challenges that the study attempts to solve. Unclear problem statements lead to unsuccessful proposals and vague, unmanageable documents. Three elements are frequently seen in problem statements:

1. The problem itself is expressed correctly with sufficient context to show why it is significant.
2. The procedure for resolving the issue is defined.
3. The thesis's purpose, declaration of purpose, and breadth are stated.

After determining the problem statement, the research questions, aim and objectives, scope, methodology, and contributions were written.

Phase2: Explanation of the research objectives

Defined objectives lead to successful research and help the researchers stay focused on the expectations of their research progress. In phase 2, the research objectives are stated clearly and in detail. The thesis has three main research objectives after determining the problem statement. Thus, the thesis has three research questions that reflect the thesis title and help achieve the objectives. Then the research scope was determined to explain the boundaries of the research.

2.3 Theoretical Concepts

This section presents the theoretical concepts for the proposed model.

2.3.1 Image Recognition

Image recognition is the identification of a selected object in an image. An object can be seen in various forms, such as color, texture, and shape (Koch et al., 2015). Recognition refers to many kinds of visual abilities, which include identification, categorization, and discrimination. Typically, object recognition can categorize an object into an instance of a particular object class (Gokturk et al., 2016). An object is also recognised from different viewpoints, such as front, side, or back, in various places and sizes (C. Liu et al., 2016). Therefore, image recognition probably starts with image processing techniques such as noise removal and feature extraction (Zhou et al., 2017). Image processing is a technique that applies a specification operation to a selected region or image (Lang et al., 2015).

2.3.2 Face Recognition

Facial recognition is identifying an individual using the face. Three elements are frequently seen in problem statements: The problem is expressed correctly and with sufficient context to demonstrate its significance. It is fast developing into a part of peoples' everyday lives. It was founded for a long time, either in the academic or industrial fields. Several major industries have gained from Facial Recognition technology over the past 60 years, including finance, law enforcement, mobile technology, and border control.

Face recognition models are divided into verification and identification (Pandya et al., 2013). Face verification is the process of comparing the entered face with a template face picture whose identity is being asserted, whereas face identification is the process of comparing the inputted face with all the photos in the database (Abate et al., 2007).

Almost all face recognition algorithms adhere to the Figure 2.1 pipeline (D. Wang et al., 2017). As illustrated in Figure 2.1, the process begins with face recognition, followed by face alignment, appearance normalization, feature description, feature extraction, and matching.

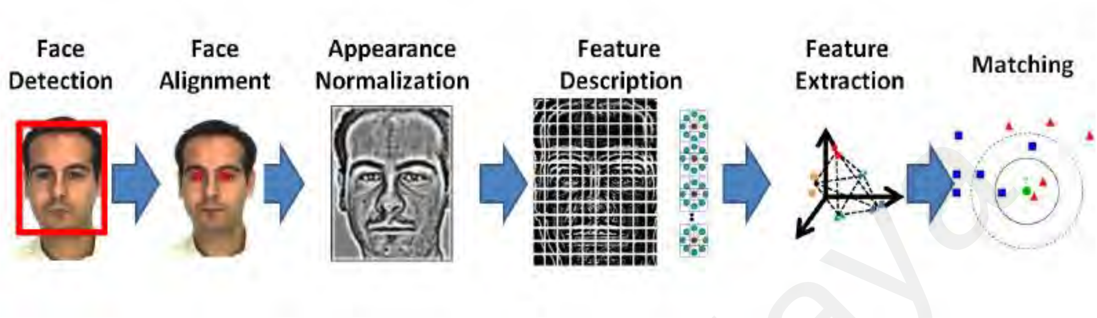

Figure 2.2: Face recognition process (Dayong et al., 2017)

Modeling facial aging has some limitations, including the age from face to face, craniofacial growth, and skin aging. The directional age primitive pattern is used to build a model for the local face descriptor (Iqbal et al., 2017). The model is based on three functions for DAPP coding, which begin with Age-Primitives, that encode aging based on necessary texture primitives. Then, apply the latent secondary direction and finally employ global adaptive thresholding for more differentiation in the textured region. Based on the ADIENCE and GALLAGHER datasets, the model applied to various age sets recognition tasks (Iqbal et al., 2017). Many models and methods with varying performance and limitations are used for face recognition.

Verification and identification are considered basic processes in face recognition models (Pandya et al., 2013). Face identification differs from face verification in that the input face in verification is a template, whereas, in face identification, the input face is compared to the pattern face (Abate et al., 2007). The three face recognition steps are depicted in Figure 2.2: face recognition, feature extraction, and face recognition. Currently, two methods are used: two-dimensional (2D) and three-dimensional (3D).

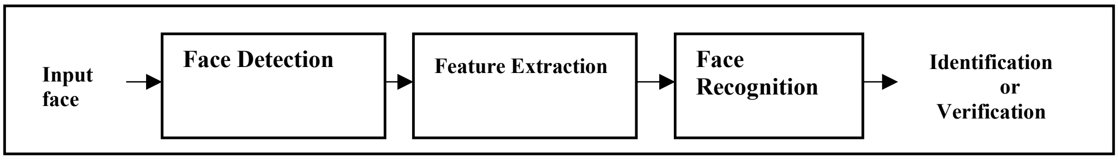

Figure 2.3: Flow of Recognition

2.3.3 Two-Dimension Face Recognition Models

This section presents various models based on two-dimension face recognition models (2D), which focus on image intensity:

Supporting vector machines (SVM) is a supervised learning model aided by data-analysis learning techniques. It is used to identify 2D characteristics (Kao et al., 2010).

Neural Network-based face recognition models have been created. Four hundred pictures from the ORL database were used to test the Quantum Neural Network (QNN) model. The model resulted in a recognition rate of 97.8% (Y. Xu et al., 2011). Face recognition is also assisted by a classifier based on Artificial Neural Networks. The technology was tested with 395 pictures from the University of Essex's database. The method had a 95% success rate (P. Zhang & Guo, 2012).

Convolutional Neural Networks (CNN) are used to calculate age and gender with a 96.2% accuracy (Dehghan et al., 2017)). Also, the Convolutional Neural Networks (CNN) model was supported by a joint Bayesian and fusion network used for face recognition with an 88.70% recognition rate (G. Hu et al., 2015).

The Genetic Algorithm is one of the search heuristic algorithms that shows natural selection activity. (Geng et al.,2006) improved the recognition rate by 96.507% (Sukhija, Behal, & Singh, 2016).

The K-Nearest Neighbors Algorithm is one categorization approach (K-NN). K NN classifies objects in the feature space using training datasets. K-NN is used to extract features from pictures using high-rank classes. On average, the recognition rate was 92.7% (H. Li & Suen, 2016).

3. One of the statistical approaches is Eigenface. It pulls the fundamental characteristics of a face from a variety of pictures. For face identification and detection, the Eigenfaces technique is utilized. Eigenfaces extract features and have a recognition rate of 26% (Ghorbel et al., 2016).

2.3.4 Three Dimension Face Recognition Approach

The rapid development of three-dimensional (3D) faces recognition sensors shows a new face recognition route that may overcome the fundamental limits of 2D technology. The geometric information in 3D facial data might significantly enhance identification accuracy in difficult-to-recognize situations (Adjabi et al., 2020). Many academics have shifted their attention to 3D face recognition, resulting in a new research trend. The acquisition of 3D face samples necessitates specialized gear, classified as active acquisition systems (Dargan el at., 2020). Because the data contains human faces and distracting characteristics such as hair, ear, neck, eyeglasses, and jewelry, acquired 3D face data cannot be directly utilized as input to feature extraction algorithms (Zhou el at., 2018). Valid, when we humans recognize each other, these characteristics may be helpful (Abate et al., 2021). However, computers, at least for the time being, are not as intelligent as humans. Hair, spectacles, and jewelry are examples of features that might be altered at any time. Moreover, ear and neck characteristics are not consistently recognized for various head postures. The location and orientation of the human face are detected as the initial stage in pre-processing. The human face is "turned" straight against the camera axis via geometric modifications (Ahmed et al., 2019). The pre-processing then isolates the human face area from distracting characteristics with the aid of clearly recognizable facial components such as the nose; ‘Segmentation’ is the term for this process. 3D models, unlike 2D face pictures, contain geometry information and are unaffected by changes in posture or illumination.

In recent years, many research institutes have set up various types of 3D face databases to test and assess their 3D face recognition techniques. For obtaining 3D face models, there are two types of acquisition techniques: active acquisition technologies and passive acquisition technologies (Adjabi et al., 2020). A stereo camera is the most common passive acquisition device. Triangulation technology is utilized in active acquisition techniques, such as the Minolta Vivid scanners. The scanner shines laser light on the face and then records the image of the light spot with the camera. Automatic facial expression identification, age-invariant face recognition, and transfer learning are only a few of the unsolved challenges in 3D face recognition (Mahmood et al., 2017). Three open challenges in their infancy include 3D facial expression analysis, recognition under age variations, and transfer learning. 3D sensing and visualization would benefit from future 3D technologies. 3D sensing is a depth-sensing technique that enhances camera capabilities for facial and object identification in augmented reality, gaming, self-driving cars, and other applications (Chen et al., 2020). As Apple, Google, and Samsung compete to put 3D sensors into their next generation of smartphones, 3D sensing technology is set to go ubiquitous. 3D visualization is the most recent widespread technology that allows users to create high-quality digital goods by creating items in three-dimensional space using 3D software. Because this is the same sphere for creating and improving 3D visualization production, it would be utilized in games, cartoons, films, and motion comics (Xu et al., 2021). In addition, 3D facial recognition technology has been used in a variety of applications, including access control and automated driving. Face ID is a feature of the iPhone X that allows you to unlock your phone by scanning your face with infrared and visible light. It is highly secure and operates in a range of situations. Face recognition models have two main challenges; the first one is face recognition in static images, and the second one is face recognition in real-time. Most face recognition frameworks try to extract parts of the entire face by removing the background and extra parts of the person's head, such as hair, that are not important for the face recognition task. Most face detection systems use an example-based learning approach to decide whether it is a face or not. There is another technique for determining whether there is a face inside the face detection system's window - using Template Matching.

Real-time face detection involves detecting a face from a series of frames from a video capturing device. While the hardware requirements for such systems are far more Stringent from a computer vision standpoint, real-time face detection is actually a far more straightforward process than detecting a face in a static image. This is because people are continually moving, unlike most of our surrounding environment. We walk around, blink, fidget and wave our hands. Since in real-time face detection, the system is presented with a series of frames that are used to detect a face, using Spatio-temporal filtering (finding the difference between subsequent frames), the area of the frame that has changed can be identified, and the individual detected.

2.3.5 Facial Features

Facial features, such as eyes, lips, and nose, are distinguishing features of the face. Recently, computer applications have begun to use these features to identify a person's face, and they can be used to estimate age, either the exact age or a range of ages. Lip, ear, eye distance, eye width, forehead length, nose length, the distance between the minor axes, and distance between the minor axes are used to find variations in physical personality characteristics. Identifying a person is typically accomplished by using obtained facial features extracted by software from various computer applications. This can be used in different applications, such as finding lost people, fraud recognition, recruiting procedures, the criminal judging system, and military selection.

Facial features are used for biometric validation. They can also be used for other applications. For example, facial characteristics such as the ears, nose, forehead, and the distance between the eyes can be used to assess a person's personality (Hackett et al., 2020). The shape of the face is one of the facial characteristics utilized for face recognition. The facial shape is calculated by dividing the minor axis of the face and the axis drawn parallel to the eye level by the minor axis of the face and the axis drawn parallel to the lip level, as shown in Figure 2.4.

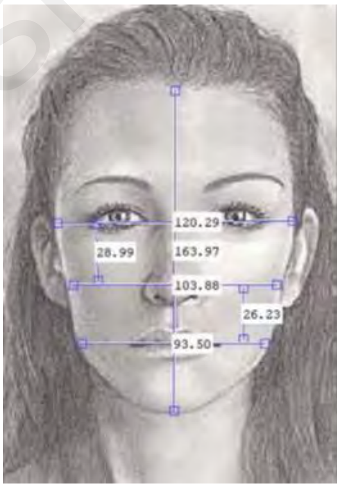

Figure 2.4: Image for detecting Face Shape (Kini et al., 2014).

One of the critical features is the nose and ear length measurements, as shown in Figure 2.5. The nose length is the distance between the center of the brow and the tip of the nose. The distance between the highest and bottommost points of the ear can also be described as ear length. Ear length has a significant impact on shaping a person's personality.

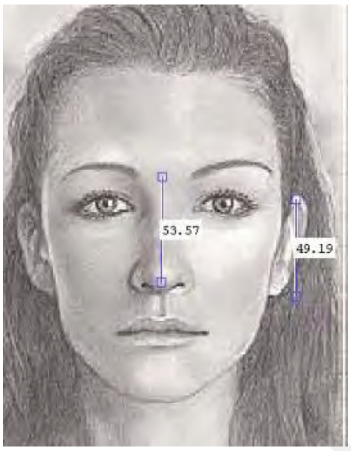

Figure 2.5: Detecting ear and nose length (Kini et al., 2014).

The width and length of the forehead, which provide an image of the face, is another essential facial feature. The forehead length is the distance between the midpoint of the brows and the midpoint of the top line of the forehead (Merler et al., 2019). The breadth of the forehead is perpendicular to the length of the forehead and goes across the middle. The length and width of the forehead are measured in Figure 2.6.

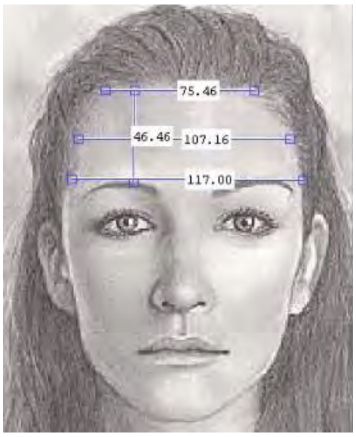

Figure 2.6: Image for determining the length and width of the forehead (Kini et al., 2014).

The width of the eyes adds another facial trait that may be used to identify a person's personality, and the difference is the distance between the eyes. The eye's breadth is the distance between both ends of the eye. The distance between the eyes' endpoints is known as the eye distance. The process for determining eye width and distance is depicted in Figure 2.7.

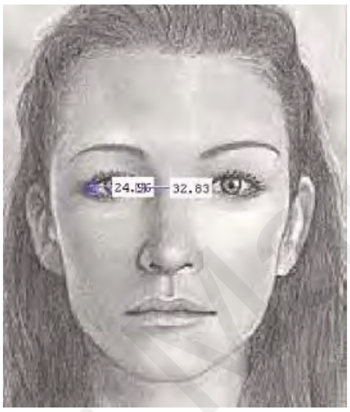

Figure 2.7: Image for determining the width and distance of the eye (Kini et al., 2014).

2.4 Image processing techniques in face recognition

Image processing techniques are used for two reasons; to enhance the image or reduce the noise in the image. These techniques are used according to the different usage purposes. For example, image enhancement techniques are generally used to focus on the important information in images, such as the edge-preserving self-snake model-based techniques. One of these techniques is image denoising, which is mostly used to remove the useless information in the image. Also, image denoising techniques were categorized into spatial domain and transform domain.

2.5 Image File Formats

A PGM image file is a grayscale image saved in the portable gray map (PGM) format. It is a file encoded with one (8 bits) or two bytes (16 bits) for each pixel, and has header data and a network of numbers that define various gray-level scales from the black color (0) to the white color (up to 65,536) and the gray-level scale between them. PGM image files have a binary presentation and are naturally saved in ASCII text format (Du et al., 2021).

2.6 Face Image Representation

Comparatively huge multidimensional images of faces are produced by many software applications. The common feature of images is the compound importance of several factors and approaches related to lighting environments, image structure, imaging and viewpoint, illumination and contain diverse facial expressions and head poses. The multidimensional objects are formally termed (tensorial data/objects). As it is well known, when data consists of one dimension, it is defined as a vector, and when it consists of two dimensions, it is defined as a matrix, whereas when it consists of more than two dimensions, it is known as a tensor. Tensor elements are described with indices that define their order, and each index refers to one of its modes. Naturally, different objects have some specific structures that take the form of two-, three- or even higher-order tensors. For instance, a grey-level face image is a two-dimensional tensor object, while a color image and also the sequence data from grey images in a video are in the form of third-order tensor objects. Thus, algorithms are required to extract useful information from such large data as these data objects in a high-dimensional space.

2.7 Review of the existing models for face recognition Previous Studies

2.7.1 Face Recognition and Database

Table 2.1 indicates the assessment techniques for various approaches and algorithms in the selected review papers. This table shows 23 databases used to evaluate the different face recognition models. The evaluation results are different from one database to another. The minor database in this survey section is BANCA, which contains 52 images, and the most extensive dataset is CAS-PEAL-R1, which contains 30,900 images.

Table 2.1: Various Databases Applied for Face Recognition

Database Name	Total of Images	Model	Accuracy	Analysis
Olivetti Research Laboratory (ORL)	400	Eigenvalues (Gaidhane et al., 2014)	99.40	To compare the similarities, the model was utilized to extract the characteristics and categorize the faces. They employ various illuminations and poses. It is a positive thing because the time is short.
		Quantum Neural networks (QNN) (Xu et al., 2011)	97.8	It is a QNN-based model intended for face recognition. The model has a high level of robustness.
		Vector Projection Length (Hu et al., 2014)	96.88	The model is based on the vector projection length used to solve the pattern of face recognition. The models pose variations that need to be improved.
		Sparse Boosting Representation Based Classification (T. Liu et al., 2016)	89.50	The model can deal with gesture changes, illumination, and corruption. It provides a low time cost.
		Face Recognition System Using Genetic Algorithm (Geng et al.)	95.895	A Genetic Algorithm based approach is being proposed for face recognition.
		SVM (Biswas & Sil, 2017)	98.4	In a reduced transformed domain, a face recognition technique employs statistical measurements and histogram analysis.
Yale database B	5760	Eigenvalues (Gaidhane et al., 2014)	97.50	The model was used to extract the features and categorize the faces to compare their similarity. They use different illumination and pose. The time is less this is good.
		Vector Projection Length (Hu et al., 2014)	73.16	The model is based on the vector projection length used to solve the pattern face recognition. The models pose variations that need to be improved.
		Sparse Boosting Representation Based Classification (T. Liu et al., 2016)	79.785	The model can deal with gesture changes, illumination, and corruption. It provides a low time cost.
FERET	14,126	Eigenvalues (Gaidhane et al., 2014)	98.00	To compare the similarities, the model was utilized to extract the characteristics and categorize the faces. They employ various

				illuminations and poses. It is a positive thing because the time is short.
		Vector Projection Length (Hu et al., 2014)	67.35	The model is based on the vector projection length used to solve the pattern face recognition. The models pose variations that need to be improved.
		Kernel collaborative representation (KCR) (Wang et al., 2015)	88.3	The face recognition model is based on kernel collaborative representation. The model has to improve the mix of characteristics to enhance recognition.
		Sparse Boosting Representation Based Classification (T. Liu et al., 2016)	65.40	The model can deal with gesture changes, illumination, and corruption. It provides a low time cost.
		SVM (Biswas & Sil, 2017)	97.945	In a reduced transformed domain, the face recognition technique employs statistical measurements and histogram analysis.
		Principal Local Binary Patterns (Pujol & García, 2012)	94	The Local Binary Pattern was designed for face recognition, but the system incorrectly classifies images into different poses.
		Partial Least Squares (PLS) (Schwartz et al., 2010)	88.825	A reduction model based on Partial Least Square (PLS) was built to obtain an effective feature combination for better learning.
AR	4000	L2-Norm Regularisation (Liu et al., 2014)	95.3	The face recognition model is based on the l2-norm.
		Kernel Collaborative Representation (KCR) (Wang et al., 2015)	99.3	The model is based on a kernel of collaborative representation for face recognition. To improve recognition, the model feature combination must be improved.
		Sparse Representation Based Classifier (SRC) (Yang et al., 2015)	98.5	The model is based on a Sparse Representation Based Classifier (SRC) used for face recognition.
		Sparse Boosting Representation Based Classification (T. Liu et al., 2016)	87.3	The model can deal with gesture changes, illumination, and corruption. It provides a low-cost time.
		Linear Discriminant Approach (Gao et al., 2013)	69.88	It is a face recognition model based on the Linear Discriminant Approach.
		Principle Component Analysis (PCA) & Locality Preserving	86.23	The model could find the face's hidden submanifold structure.

		Projections (LPPs) - (Zang et al., 2012)		
CMU PIE	41,368	l2-norm regularization(Liu et al., 2014)	94.4	The model is based on the l2-norm for face recognition.
Sheffield(previously UMIST))	564	L2-Norm Regularisation (Liu et al., 2014)	89.3	The model is based on the l2-norm for face recognition.
CAS-PEAL- R1	30,900	l2-norm regularization(Liu et al., 2014)	76.0	The model is based on the l2-norm for face recognition.
Yale	165	Vector Projection Length (Hu et al., 2014)	96.67	The model is based on vector projection length was used to solve the pattern face recognition. The models pose variations that need to be improved.
		Linear discriminant approach (Gao et al., 2013)	100	It is a face recognition model based on the Linear Discriminant Approach.
UMIST	757	Vector Projection Length (Hu et al., 2014)	100	The model is based on the vector projection length used to solve the pattern of face recognition. The models pose variations that need to be improved.
		Genetic Algorithm (Geng et al.)	96.386	A Genetic Algorithm based approach is being proposed for face recognition.
		Linear discriminant approach (Gao et al., 2013)	89	It is a model for face recognition based on the Linear Discriminant Approach.
Extended Yale B	2414	Kernel collaborative representation (KCR) (Wang et al., 2015)	99.8	The model is based on a kernel of collaborative representation for face recognition. The model needs to improve the combination of features to improve its recognition.
		Nearest-farthest subspace (NFS) (Mi et al., 2013)	82.47	A model for face recognition is based on linear regression. The nearest-farthest subspace (NFS) is used for face recognition, but the model still does not cover all types of images; they intend to improve the model to address the incremental issue.
		Grayscale Arranging Pairs (GAP) (Zhao et al., 2013)	99.85	The model for face recognition is based on Grayscale Arranging Pairs (GAP), which compares the intensity relationship of pixel point pairs in pictures.
		Principle component analysis (PCA) & Locality preserving projections (LPPs) - (Zang et al., 2012)	91.66	The model cannot discover the face's hidden submanifold structure.

		Kernel Discriminant Transformation (KDT) (Chu et al., 2011)	97.9	The Kernel Discriminant Transformation (KDT) is a face recognition model.
CMU Multi-PIE	7750000	Sparse Representation based classifier (SRC) (Yang et al., 2015)	85.75	Face recognition was performed using a model based on the Sparse Representation Based Classifier (SRC).
		Linear discriminant approach (Gao et al., 2013)	88.85	It is a model for face recognition based on the Linear Discriminant Approach.
		Principle component analysis (PCA) & Locality preserving projections (LPPs) - (Zang et al., 2012)	86.46	The model cannot discover the face's hidden submanifold structure.
LFW	13000	Sparse Representation based classifier (SRC) (Yang et al., 2015)	76.75	The face recognition model is based on the Sparse Representation Based Classifier (SRC).
		Sparse boosting representation based classification (T. Liu et al., 2016)	51.46	The model can deal with gesture changes, illumination, and corruption. It provides me with a low-cost time. However, the precision is poor.
		Kernel Discriminant Transformation (KDT) (Chu et al., 2011)	65.3	Kernel Discriminant Transformation (KDT) is a model for face recognition; its accuracy is low.
		Deep neural network (Taigman et al., 2014)	97.35	Images are labeled inside the database using a deep neural network model for facial identification.
Libor Spacek's	7240	k-class (Alqudah & Al-Zoubi, 2015)	97	The K-class is used for face recognition.
IRIS Thermal/Visible Face	4228	Dictionary construction & sparse representation (Bi et al., 2016)	91.5%	The model for face recognition uses a local binary pattern called the Gabor jet descriptor. If they use weight features, the performance will improve with a low number of images.
Indase	150	Genetic Algorithm (Geng et al.)	97.44	A Genetic Algorithm (Geng et al.) based approach is being proposed for face recognition.
JAFFFE	230	SVM (Biswas & Sil, 2017)	97.145	In a reduced transformed domain, a face recognition technology is based on statistical measurements and histogram analysis.
		Principle component analysis (PCA) & Locality	86.42	The model cannot discover the face's hidden submanifold structure.

		preserving projections (LPPs) - (Zang et al., 2012)		
Georgiatech (GT)	750	Nearest-farthest subspace (NFS) (Mi et al., 2013)	92.29	A model for face recognition is based on linear regression. The Nearest-farthest subspace (NFS) is applied for face recognition, but the model still did not cover all kinds of images. They want to improve the model to solve the incremental issue.
AT&T	400	Nearest-farthest subspace (NFS) (Mi et al., 2013)	97.125	A model for face recognition is based on linear regression. The nearest-farthest subspace (NFS) is used for face recognition, but the model still does not cover all types of images; they intend to improve the model to address the incremental issue.
	400	CNN, RELU (Aiman et al., 2017)	95.21%	The limitation of this approach is that it uses supervised learning with human-annotated data
	400	CNN (Pranav et al., 2020)	98%	A system based on CNN used for face recognition was evaluated with AT&T, but the system did not evaluate for aging
XM2VTS	200	Principle component analysis (PCA) & Locality preserving projections (LPPs) - (Zang et al., 2012)	32.63	The model is unable to reveal the face's hidden submanifold structure solely.
BANCA	52	Principle component analysis (PCA) & Locality preserving projections (LPPs) - (Zang et al., 2012)	41.98	The model cannot discover the hidden submanifold structure in the face.
FRGC	86,634	Partial Least Squares (PLS) (Schwartz et al., 2010)	91.7	A reduction model based on Partial Least Square (PLS) was built to obtain an effective feature combination for better learning.
		Shearlet transformation and LBP. (Chen et al., 2017)	68.35	A model proposed used for face recognition for noise images but did not try on face aging
		CNN (Lu et al., 2018)	93.42	A model for face recognition based on CNN, but they did not mention facial aging.
IIT(BHU)	2100	Deep Convolutional Neural Network(	89.58	The model is based on CNN for face recognition, but it has not been tested for illumination or poses any challenges.

		CNN) (Singh & Om, 2017)		
FG-NET	1002	Decorrelated Adversarial Learning (Hu et al.)	94.5	The model reduces the relationship between identity decomposed features and their age. It gives good results.
MORPH Album 2	10000	Decorrelated Adversarial Learning (Ni et al., 2019)	98.93	The model lowers the connection between identity deconstructed characteristics and age, and it produces good results.
CACD-VS	4000	Decorrelated Adversarial Learning algorithm (Ni et al., 2019)	99.40	The model aims to minimize the age-related link between identity deconstructed characteristics. It yields positive results.
FGNET	1002	CNN (Mustafa et al, 2020)	81.5	The model is built using CNN and the VGG-Face. This model is used for face recognition. The models detect the face without a big age gap (did not detect the face aging)
	1002	PCA (DHAMIJA et al., 2020)	98	A model based on PCA used for face recognition, the accuracy of 98.87
	1002	CNN(Nimbarte et al., 2020)	91.41	A model used for face recognition based on CNN with an accuracy of 91.41
	1002	CNN	76.6% (Nimbarte et al, 2018).	Limited size for the images 32*32 to get best images
MORPH	5 5134	CNN	92.5% (Nimbarte et al., 2018).	Deep Learning with Convolutional Neural Networks provides us with a combination of feature extraction and classification
	55,134	CNN (Mustafa et al, 2020)	96.5	The model is built using CNN and the VGG-Face. This model is used for face recognition. The models detect the face without a large age gap (did not detect the face aging)
	55,000	CNN (Nimbarte et al , 2020)	98.40	A model used for face recognition based on CNN with an accuracy of 98.4
Texas3D	1149	Gupta et al. 2010	90	A model based on Euclidean distance was used for detects the face expression and possess.

As shown in Table 2.1, different databases are applied for face recognition with different techniques and results.

The model based on Eigenvalues is used to extract characteristics and classify faces to compare similarities. The model used an Olivetti Research Laboratory (ORL) dataset, which contained 400 face images. They use various lighting and poses. Time is of the essence, which is a good thing (Gaidhane et al., 2014).

Quantum Neural Networks (QNN) is used in another model. It is intended for use with face identifiers, giving good robustness. The model performed well on ORL, with an accuracy of 97.8% (Y. Xu et al., 2011). One of the models is based on the Vector projection length technique. Pattern recognition problems are solved using the model. The ORL dataset gives it a score of 96.88%. However, the model's pose variations need to be improved (C. Hu et al., 2014). The Sparse Boosting Representation Based Classification model also uses the ORL dataset. The model gives 89.50% with low-cost time (T. Liu et al., 2016). The Support vector machine (SVM) is also evaluated under the ORL and received a score of 98.4%. The model is a face recognition method applying statistical measures and histogram analysis in a reduced transformed domain (Biswas & Sil, 2017).

Another model based on Eigenvalues was created for face recognition. This model employed the Yale database B, which had a 97.50% accuracy (Gaidhane et al., 2014). A model based on the Vector Projection technique was also developed, with an accuracy of 73.16% (C. Hu et al., 2014). Then a Sparse Boosting Representation Based Classification model used Yale data-based B with 79.785% accuracy (T. Liu et al., 2016). FERET is another dataset. This dataset contains 14,126 face images.

For facial recognition, a model based on Partial Least Squares (PLS) was employed in 2010 with an accuracy of 88.825% (Schwartz et al., 2010). In 2021, a system based on Principal Local Binary Patterns would provide a 94% accuracy (Pujol & García, 2012). A model built on Eigenvalues was used with 98 % accuracy (Gaidhane et al., 2014). Another model with a 67.35% accuracy is the Vector Projection (C. Hu et al., 2014). Another model, with an accuracy of 88.3%, is based on Kernel collaborative representation (KCR) (Dong Wang et al., 2015). In 2016, a model based on Sparse Boosting Representation Based Classification provided 65.40% accuracy (T. Liu et al., 2016).

There are 4000 face images in the AR dataset. In 2012, a model of Principle Component Analysis (PCA) & Locality Preserving Projections (LPPs) for face recognition. This model gives 86.23% (Zang et al., 2012). A Linear Discriminant Approach for face recognition yielded 69.88% in 2013 (Q. Gao et al., 2013). L2-Norm Regularisation was used for face recognition in 2014, and it received a 95.3% rating (H.-D. Liu et al., 2014). A model based on Kernel Collaborative Representation (KCR) received 95.3% (Dong Wang et al., 2015). A facial recognition model based on the Sparse Representation Based Classifier (SRC) got a 98.5% accuracy (M. Yang et al., 2015). For face identification, a Sparse Boosting Representation Based Classification model was developed in 2016. The accuracy of this model was 87.3 % (T. Liu et al., 2016).

CMU PIE is one of the current datasets. The model is based on the l2-norm for face recognition. CMU PIE gave the model a 94.4% accuracy rating. However, when tested against the Sheffield dataset, the model had an accuracy of 89.3%. Furthermore, the model was evaluated with 76% accuracy on the CAS-PEAL-R1 dataset (H.-D. Liu et al., 2014). Yale was used to evaluate face recognition models. In 2013, a model with high accuracy of 100% based on a linear discriminant approach was developed (Q. Gao et al., 2013). However, in 2014, a model based on Vector Projection length was discovered with an accuracy of 96.67% for face recognition (C. Hu et al., 2014).

UMIST is also used for face recognition model assessment. The Linear Discriminant Approach was used to create a face recognition model in 2013. It scored 89% (Q. Gao et al., 2013). In 2014, another model based on vector projection length was used to solve the pattern of face recognition. Even if the accuracy is perfect, the models must improve pose variations (C. Hu et al., 2014). In 2016, a Genetic Algorithm (Geng et al.) based approach was proposed for face recognition with accuracy of 96.386% (Sukhija et al., 2016).Furthermore, Extended Yale serves as a dataset for comparing various face models. The Extended Yale B dataset was used to test the Kernel Discriminant Transformation (KDT) model for face recognition in 2011, and it was found to be 97.9% accurate (Chu et al., 2011). Extended Yale B evaluated the Principle component analysis (PCA) model for face recognition in 2012. The model accuracy was 91.66% (Zang et al., 2012).

In 2013, a face recognition model based on linear regression was developed. Face recognition is performed using the nearest-farthest subspace (NFS). However, the model does not cover all sorts of pictures; they wanted to enhance the model to solve the incremental issue. This model was evaluated on the Extended Yale B with an accuracy of 82.47% (Mi et al., 2013). Another methodology for face recognition uses Grayscale Arranging Pairs (GAP) to analyze the intensity relationship between point pairs in pictures (Zhao et al., 2013). In 2015, a face recognition technique based on kernel collaborative representation was developed. The model needs to improve the combination of features to improve its recognition (Dong Wang et al., 2015).

CMU Multi-PIE is another face dataset that evaluates a model based on the Sparse Representation Based Classifier (SRC) used for face recognition. This model has an accuracy of 85.75% (M. Yang et al., 2015). LFW is a face dataset with various models used to train different model. Kernel Discriminant Transformation (KDT) is a model for face recognition; with 65.3% accuracy (Chu et al., 2011). Another model for face recognition is based on a deep neural network where images are labeled inside the database; the accuracy is 97.35 % (Taigman et al., 2014). A new model based on the

Sparse Representation Based Classifier (SRC) is used for face recognition. The LFW dataset yielded an accuracy of 76.75% for this model (M. Yang et al., 2015). In 2015, a model based on K-class was used for the face recognition model evaluated by Libor Spacek's dataset with an accuracy of 97% (Alqudah & Al-Zoubi, 2015). In 2016, a classification model based on sparse boosting representation was discovered. This model can recognize gesture changes, illumination, and corruption. At 51.46%, it provides low-cost time (T. Liu et al., 2016). In 2017, a system for face recognition was built based on CNN and RELU. The system was evaluated with AT&T, which has 400 images and an accuracy of 95.21%. The limitation of this approach is that it uses supervised learning with human-annotated data. Also, the system did not train for aging (Aiman et al., 2017). A model based on Shearlet transformation combined with LBP used for face recognition was evaluated on the FRGC dataset after adding noise for thee images, but the model did not examine fae aging. The model accuracy was 68.35% (Chen et al., 2017).

In 2018, A model based on deep learning with convolutional neural networks provided a combination of feature extraction and classification with an accuracy of 76.6% on FGNET and 92.5% on MORPH. However, the image size should be 32*32 to give the best result (Nimbarte et al., 2018). A model for face recognition based on CNN did not mention face aging. The accuracy is 93.42% (Lu et al., 2018). In 2019, a model based on the correlated Adversarial Learning (Hu et al. 2019) algorithm was used to reduce the relationship between identity decomposition and age. It performs well with the MORPH Album 2 dataset. The figure is 98.93%. It also received a 94.5 from the FG-NET when it was evaluated. The highest accuracy for this model was with the CACD-VS Dataset, as it gives 99.40% (Ni et al., 2019). In 2020, different datasets with different models were used for face recognition. One of the models was built using CNN and the VGG-Face and is used for face recognition. MORPH evaluated the model with an accuracy of 96.5%. However, the model gives 81.5 % accuracy with the FGNET dataset. Although the model detects the face without a significant age gap, it did not detect the face aging (Mustafa, 2020). Also, a QSVM-PCA model was built based on PCA, and the model was evaluated using 280 images from FGNET. The model process is done in 80 seconds to obtain the result, which is a good performance with an accuracy of 98.87 % (DHAMIJA et al., 2020). A model based on CNN was used for face recognition and evaluated with two databases first database is MORPH, which has 55,000 images , and the second database is FGNET with 1002 images (Nimbarte et al., 2020). The model accuracy for MORPH is 98.40% and for FGNET is 91.14%. A system design of a real-time based on CNN was used for face recognition evaluated with AT&T, but the system did not evaluate aging. The system accuracy was 98% (Pranav et al., 2020).

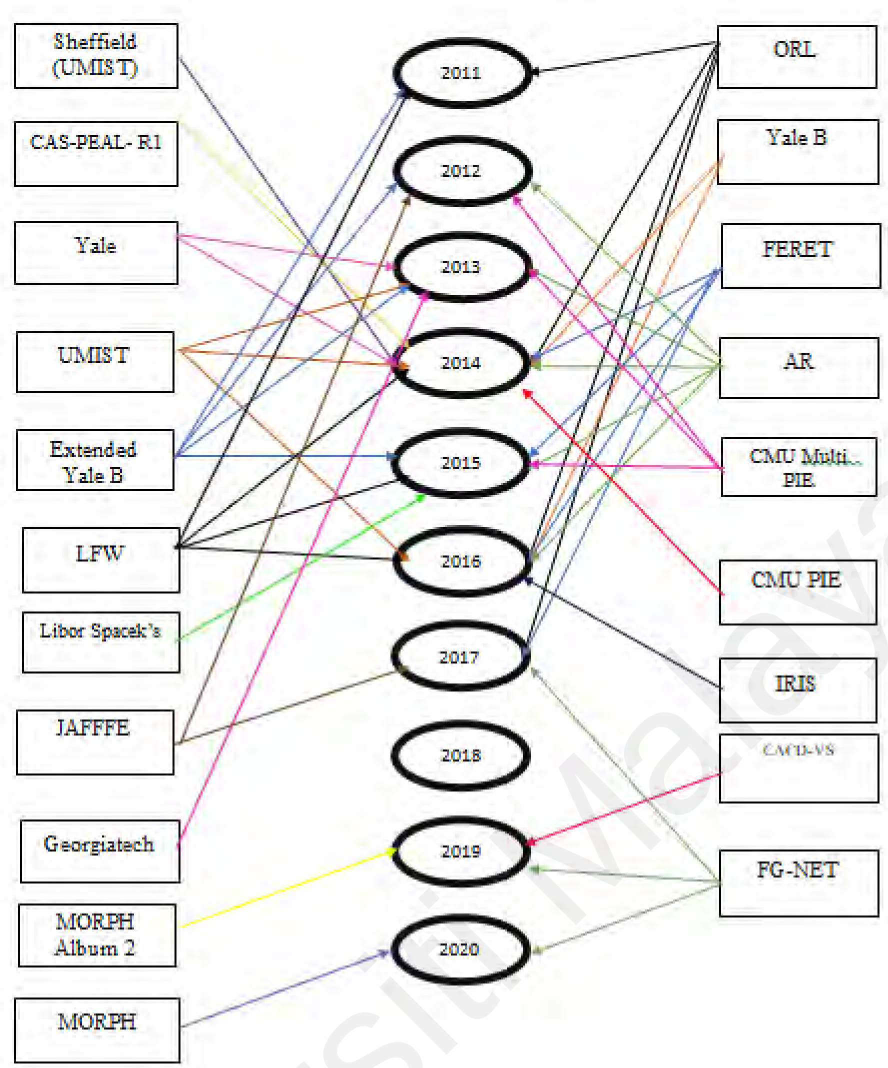

Figure 2.8: Different datasets based on Years

Table 2.2: Different datasets with different gap age

Dataset Name	References	Years – gap	Database size
AR	(T. Liu et al., 2016)	14 days	4000
BANCA	(Zang et al., 2012)	3 months	52
XM2VTS	(Zang et al., 2012)	4 months	200
CMU Multi-PIE	(M. Yang et al., 2015)	5 moths	775000
CAS-PEAL- R1	(Liu et al., 2014)	6-month	30,900
Georgiatech(GT)	(NFS) (Mi et al., 2013)	10 months	750
Yale database B	(T. Liu et al., 2016)	Less than one year	5760
Yale	(Hu et al., 2014)	Less than one year	165
UMIST	(Sukhija et al., 2016)	Less than one year	757
Extended Yale B	(Wang et al., 2015)	Less than one year	2414
LFW	(T. Liu et al., 2016)	Less than one year	13000
IRIS	(Bi et al., 2016)	Less than one year	4228

Thermal/Visible Face			
JAFFFE	(Biswas & Sil, 2017)	Less than one year	230
FRGC	(Schwartz et al., 2010)	Less than one year	86634
FERET	(Schwartz et al., 2010)	3 years	14,126
Libor Spacek's	(Alqudah & Al-Zoubi, 2015)	3 years	7240
AT&T	(NFS) (Mi et al., 2013)	3 years	400
Olivetti Research Laboratory(ORL)	(Gaidhane et al., 2014)	3 years	400
MORPH	CNN (Mustafa, 2020)	5 years	55,134
MORPH Album 2	(Ni et al., 2019)	5 years	10000
CACD-VS	(Ni et al., 2019)	0 -10	4000
FGNET	(Ni et al., 2019)	0 -45	1002

Table 2.2 shows different datasets with the gap in years between the images. The dataset age gap has been categorized into five categories, as shown in Figure 2.8. These categories are one year and more petite, three years and less, five years and less, ten years and less, 45 years and less.

AR is the first dataset; it contains 4000 face images. These images have various facial expressions, occlusions, and illumination settings. However, the age gap between the same person images is 14 days (T. Liu et al., 2016). BANCA is a face image dataset with 52 images. BANCA images have an age gap between same person images not more than three months (Zang et al., 2012). XM2VTS, which contains 200 images, is also considered a face image dataset. The age gap is not more than four months (Zang et al., 2012). Another face image dataset is CMU Multi-PIE, which has 7750000 images with an age gap of not more than five months (M. Yang et al., 2015). The CAS-PEAL-R1 dataset for face images has 30900 images, and the age gap between the two pictures of the same individual is less than six months (H.-D. Liu et al., 2014). Georgia Tech (GT) has a dataset containing 750 images. The age gap is not more than ten months (Mi et al., 2013). Yale database B has 5760 face images in the dataset with an age gap of not more than one year (T. Liu et al., 2016). Yale is one of the face image datasets that contains 165 images. The age gap between the same images for a person is less than one year (C. Hu et al., 2014). UMIST contains 757 images, and the time difference is less than a year (Sukhija et al., 2016). Extended Yale B is one of the face image datasets. This dataset has 2414 images and an age gap of not more than one year (Dong Wang et al., 2015). LFW contains 13000 face images with an age difference between the images of the same person less than a year (T. Liu et al., 2016). IRIS Thermal/Visible Face is also one of the face datasets. This dataset has 4228 images. The age gap is not more than one year (Bi et al., 2016). JAFFE is likewise one of the face datasets, with 230 images. The age gap is not more than one year (Biswas & Sil, 2017). Similarly, FRGC has 86634 face images with an age gap of less than one year (Schwartz et al., 2010). FERET contains 14126 images. The age gap for the same person is not more than three years (Schwartz et al., 2010). Libor Spacek's face image dataset is one of the best (Alqudah & Al-Zoubi, 2015). The age gap is not more than three years. One of the current datasets is AT & T. There are 400 images in the AT & T dataset. The age gap is not more than three years (Mi et al., 2013). The Olivetti Research Laboratory (ORL) also contributed to the data. This dataset has 400 images and an age gap not more than three years (Gaidhane et al., 2014).

MORPH has a dataset as well. This dataset contains 555134 images. The age gap is not more than five years (Mustafa, 2020). MORPH Album 2 contains a total of 10,000 face images. The age gap is not more than five years. CACD-VS has 4000 face images. The age gap is not more than ten years (Ni et al., 2019). FGNET is the last dataset on the table. It has 1002 images. The age gap is not more than 45 years (Ni et al., 2019).

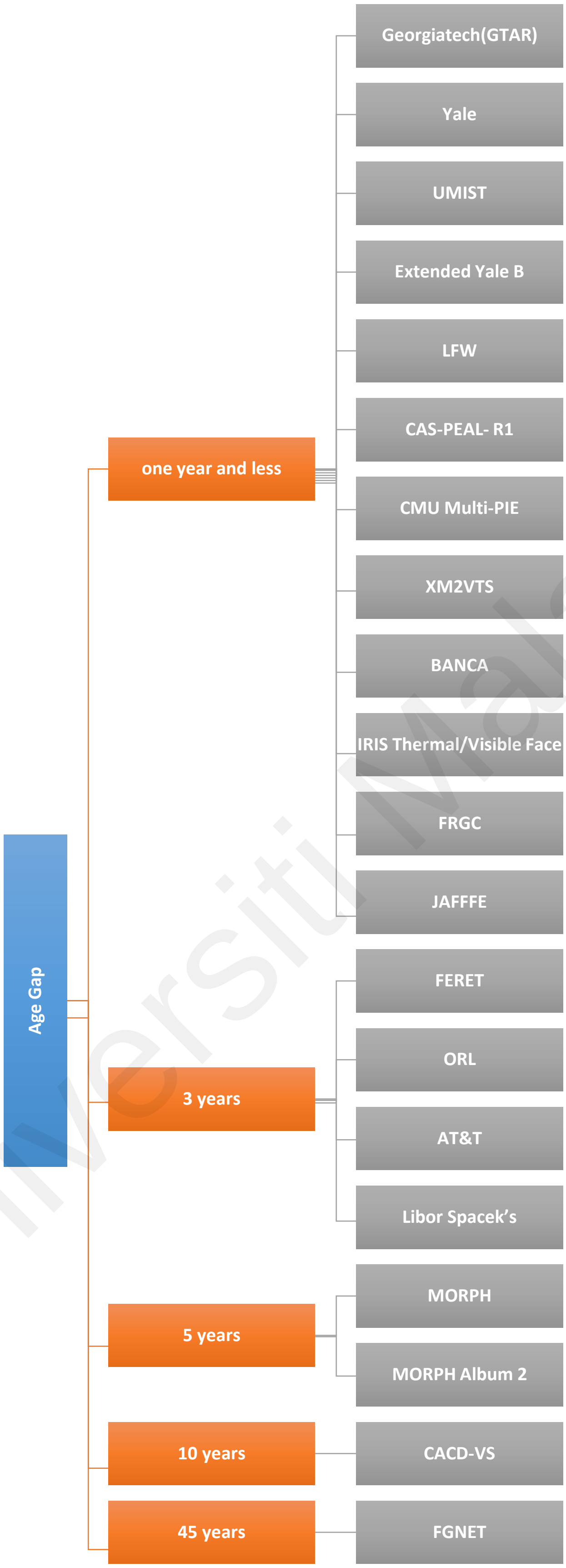

Figure 2.9: The five categories for the age gap

2.7.1.1 Summary of Face Aging Datasets

The chronology for face datasets in this thesis started from 2011 to 2020, as shown in Figure 2.8, which shows the datasets based on Years and Table 2.1 that explain the various databases applied for face recognition. As a result of reviewing the previous studies and datasets at Table 2.1, Table 2.2 was founded to show the age gaps between the images at the datasets, and Figure 2.9 shows the five categories for age gap at the databases images. Finally, it is found that not all the databases can be used for the face aging recognition models according to the natural images for the datasets. Figure 2.10 shows the taxonomy of datasets that can be used for face aging recognition models to utilize them for face aging recognition models. The most crucial database that can be used for evaluating the face aging recognition models are CALFW, MORPH Album 2, MORPH, FGNET, CACD-VS.

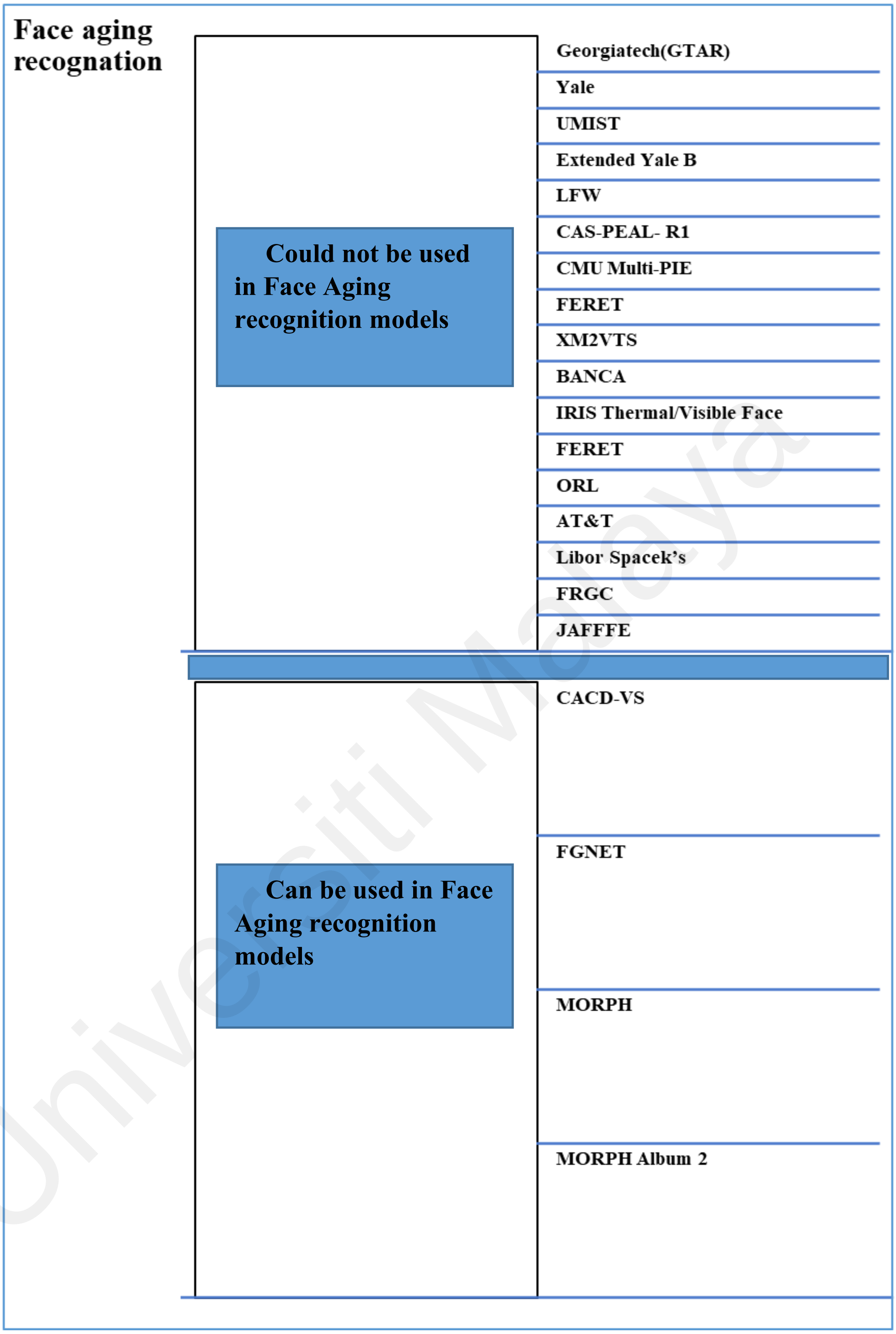

Figure 2.10: The taxonomy of datasets that can /cannot be used in face aging recognition system.

2.7.2 Review of Face Recognition Models

This section aims to study the different models for human face aging and age progression, commonly called age synthesis. Generative adversarial networks (GANs) are used to assess automatic face aging, a pair of neural networks.

A system for detecting the face while wearing a mask is used to conduct the analysis. A system based on Convolutional Neural Network was trained using different datasets, such as LFW , CPLFW ,CALFW and RFW. The system was trained with more than 7000 images (Wang et al., 2021). Another system based on CNN with GoogleNet is used for face verification in an uncontrolled environment. The system trained on two datasets; LFW and YTF . The system gives 99% with LFW datasets which have 13,233 face images. For YTF, which has 3425 face videos, the accuracy was 96% (Ben Fredj et al., 2021). Also, a multi-face recognition system based on Convolutional Neural Network (CNN) was tested with 12000 images collected by the authors. The accuracy was 87% (Diyasa et al., 2021).

Another system based on Gabor filter was used to detect the face with makeup, Pose, Illumination, and Expression (PIE). The system was trained under YMU dataset with 5000 images. The system accuracy was 82%. A face recognition system based on Convolutional neural networks was used to detect the faces. The system accuracy was 8.56% but ignored the impact of illumination in recognition of the human face (Tabassum et al., 2020). A Deep convolutional neural networks system is used to detect the presence of a face in the image. The system evaluated two databases; Extended Yale B Face and CMU PIE, with an accuracy of 99.89 % and 99.44% (Bendjillali et al., 2020). Also, a system based on Neural Network was used to detect the different poses for the face. The system was tested with Olivetti and Oracle Research Laboratory(ORL) with an accuracy of 97.25% (Gooya et al., 2020). A framework was used to detect facial expressions based on Discrete Wavelet Transform (DWT), which is combined with four different algorithms: error vector of principal component analysis (PCA), eigen vector of PCA, eigen vector of Linear Discriminant Analysis (ALDayel & Magdy) and Convolutional Neural Network (CNN). The framework was tested by collected datasets by the authors with 89.56% accuracy (Tabassum et al., 2020).

Convolution Neural Network (CNN) is a common technique used to detect the pose invariant for the face. The system was tested under four datasets MS-Celeb1M, UMD Faces, VGGFace2 and WebFace, with accuracy of 95.5%, 98.2%,98%.9% and 98.9%, respectively (Wang & Guo, 2019). Besides this, a system based on Generative Adversarial Network was used to detect the pose invariant for the face. The system was evaluated under CMU-PIE with an accuracy of 99.30 % (Huang et al., 2018). Similarly, a model based on Multicolumn Networks is used to decide if two images belong to the same person. The model performance is 88.7% (Xie & Zisserman, 2018). Also, a system was trained on 110k images. 10k images were applied to the evaluation of the IMDB-WIKI clean dataset. The system performance was 82.9% (Antipov et al., 2017). A statistical approach uses wrinkle information to estimate age from a facial picture was categorized using a fuzzy c-means clustering algorithm. The precision has increased to 87.5% (Jana & Basu, 2017). Using a Siamese network to decrease the aging factors in face recognition, a statistical approach uses wrinkle information to estimate age from a facial picture, categorized using a fuzzy c-means clustering algorithm with MAE of 5.13 (X. Wang et al., 2017). A model based on local pattern selection (LPS) matches a person's older face with his (or her) younger one. The MORPH and FGNET databases were used to evaluate the model. The accuracy is 94.87% (Li et al., 2016). The RNN model provides smooth face aging between each RNN-network neighboring group (Wang et al., 2016).

Table 2.3: The current techniques of face recognition

Paper	Ref.	Objective	Dataset	No. Images	Problem	Method	Result	Limitation
Multi-face Recognition for the recognition of Prisoners in Jail using a Modified Cascade Classifier and CNN	(Diyasa et al., 2021)	To build a multi-face recognition system	Collected images by the authors	10000	Recognition the face	Convolutional Neural Network (CNN)	87%	The dataset is only for 11 person
FaceX-Zoo: A PyTorch Toolbox for Face Recognition	(Wang et al., 2021)	T detect a face wearing mask face	CPLFW ,CALFW, RFW	Total of 7000	Masked face recognition	Convolutional neural network	61.56 % , 76.52 % , 73.27	The accuracy is not high
Face recognition in unconstrained environment with CNN	(Ben Fredj et al., 2021)	face verification on an uncontrolled environment	LFW and YTF	13,233 images , 3425 videos	Face Verification	CNN model based on GoogLeNet	99 %, 96 %	The model did not recognize face aging
Illumination normalization techniques for makeup-invariant face recognition	(Saeed et al., 2021)	Detect the face with makeup, Pose, Illumination, and Expression (PIE)	YMU	5000	Detect the face with pose and makeup.	Gabor Filtering	82%	The model did not recognize face aging
Human face recognition with combination of DWT and	(Tabassum et al., 2020)	Build a system to detect the facial expression	Collected images by the	150	Detect facial expression	Discrete Wavelet Transform (DWT) is	89.56 %	The model did not recognize face aging and the number

machine learning			authors			combined with four different algorithms: error vector of principal component analysis (PCA), eigen vector of PCA, eigen vector of Linear Discriminant Analysis (ALDayel & Magdy) and Convolutional Neural Network (CNN		of images was low
Human face recognition with combination of DWT and machine learning	(Tabassum et al., 2020)	Recognize the face on the image	Collected images	30	Recognize the face with ignorance the impact of illumination in recognition of human face	Convolutional neural networks	89.56 %	It is used for facial expression
Illumination-robust face recognition based on deep convolutional neural networks	(Bendjillali et al., 2020)	Detect if there is a face or no at the image	Extended Yale B Face , CMU PIE	16128 images , 41,368 images	Detect the face	Deep convolutional neural networks	99.89 , 99.44	It is used for face detection

architectures								
Robust and discriminating face recognition system based on a neural network and correlation techniques	(Gooya et al., 2020)	Detect the different pose for the face	Olivetti and Oracle Research Laboratory(ORL)	Total images 100 , 10 different images of each of 40 distinct Subjects.	Detect the different pose for the face	Neural Network	97.25 %	Used for recognizing different poses for the same person
Benchmarking Deep Learning Techniques for Face Recognition	(Wang & Guo, 2019)	Detect the pose invariant for face	UMDFaces, WebFace, VGGFace2and MS-Celeb1M	3,8386 54,236 856,31 1890,3 93700	Detect the different pose for the face	Convolution Neural Network (CNN)	95.5 % , 98..2 %,98 % .9%, 98.9 %	Used for recognizing different poses for the same person
Towards Pose Invariant Face Recognition in the Wild	(Huang et al., 2018)	Detect the pose invariant for face	CMU-PIE	754204	Detect the different pose	Generative Adversarial Network	99.30 %	Used to recognize different poses for the same person
Multicolumn Networks for Face Recognition	(Xie & Zisserman, 2018)	Set-based face recognition	VGGFace2	9131	Decide if two sets of images of a face are of the same person or not.	Multicolumn Networks	88.7 %	Used images at the same age for the same person
A new approach for automatic face emotion recognition and classifica	(Salunke & Patil, 2017)	To build a framework for detecting the facial expression	FERC-2013	2000	Recognition of facial expression for unknown	Convolutional Neural Networ	80%	Used to recognize different facial expressions for the same person

tion based on deep networks.						k (CNN)		
Face aging with conditional generative adversarial networks	(Antipov et al., 2017b)	Gan-based Method for automatic face aging	IMDb-wiki cleaned	12k	Synthetic aging of human faces	Generative adversarial networks	82.9 %	The accuracy is still not high
Automatic age estimation from face's image	(Jana & Basu, 2017)	Estimates the age of a person by analysis of wrinkles on their face images.	Collected images	140	Age progression influences skin texture.	Fuzzy means	87.5 %.	It is used for estimating the age by using wrinkles
Unleash the black magic in age: a multi-task deep neural network Approach for cross-age face verification	(X. Wang et al., 2017)	To improve cross-age face verification performance with age information	Morph	15000	Face recognition is affected by aging.	Siamese network	Mae 5.13	need to improve the error rate
Aging face recognition: a hierarchical learning model based on local patterns selection	(Li et al., 2016)	Matching a person's older face to his (or her) younger one.	Morph, fgnet	22692	The facial appearance of humans changes dramatically as they age.	Local pattern selection (lps)	94.87 %	Limited images for aging
Recurrent face aging	(Wang et al., 2016)		Lfw , morp ,cad	10635	The absence of labeled face data on the same person captures a wide age range.	Recurrent neural network	98.40	The model is based on pattern and still need to add new pattern for aging

2.7.3 Artificial neural networks Models

There are different models based on Artificial neural networks (Vedel et al.) used for face recognition. These techniques are reviewed in Table 2.3, which presents the different ANN models used for face recognition.

Table 2.4: ANN Face Recognition models

Model	Citation	Objective	Dataset	No. Images	Problem	Method	Result	Criticize
Face recognition with Expression Recognition using Artificial Neural Networks	(OWAYJAN, 2016)	It describes the face recognition model with expression. Recognition Using Artificial Neural Networks.	Cohn-Kanade	60	Identify the face, and then recognize the facial expression.	An artificial neural network manages multi-layer-Perception with a backpropagation algorithm for features Extraction and classification.	65%	The number of images is reduced, with an average result of 65%.
Hybrid network combined with Viola-Jones used for face recognition	(MURPHY, 2016)	Enhanced the results in face recognition	CMU/ VASC	354	frontal face recognition	A neural network employs the Viola-Jones cascade algorithm.	82	Ten nodes are used for ten features in the model, but the ten hidden nodes always produce good results. Also, the model minimized the errors by using the backpropagation algorithm for

								network weights
Artificial neural network with multi-layer combines with Gabor filter used for recognizing the face expression	(VERMA, 2017)	Artificial neural network with multi-layers combines with Gabor filter to classify the facial expressions.	JAFFE	213 face images of 10 models for apanese ady. Every subject is considered an emotion.	Recognition the face features and then classify the facial expressions	Artificial Neural Network Combine with Gabor filter (Gelb & Metz)	85.7	The model assessed Japanese females but did not compare people from different countries.
A model for recognizing the face based on Videos	(Ding & Tao, 2017)	CNN is used to recognize blur-insensitive images and extract features automatically. Also, to improve the strength of CNN features to occlusion and pose variations.	COX Face, PaSC, and YouTube Faces.	6227 videos	The problem is to detect the images, which are a blur, dramatic pose variations, and occlusion.	Convolutional Neural Networks	95	The time cost increased slightly
Artificial Neural Network (Vedel et al.) recognizes the person three times per day.	(PANJAITAN ET AL., 2018)	A model based on ANN to increase accuracy compares faces at different times (Morning, Afternoon, Evening).	Collected dataset by the authors	50	Recognize the faces to prevent the fake identification	Artificial Neural Network (Vedel et al.) with K-Nearest	38.177 %	The accuracy is too low
A model is used for recognizing the face using a single hidden layer analytic	(OH ET AL., 2018)	To build a model to minimize the error rate by improving feature extracting proves, then classificati	FERET b-series, COX, CMU-PIE, Multi-PIE, LFW-a, FERET f-series	120000	The percentage of recognition can the model detect	Gabor Feedforward Network	96.7	The model improves the time cost even though the model is

Gabor feedforward network.		on.			with pose variation and angle, the robustness of the model for pose sensitivity			complex by training a single sample per person.
Advanced artificial neural network for face recognition	(HUSSEIN, 2019)	A model BP-ANN has been established to enhance the accuracy of face recognition	Collected dataset by the authors	100	Face recognition	Backpropagation artificial neural network	82	The neural network has two hidden layers, each with 100 neurons, indicating that the model is more complex and time-consuming.
The model used Principal Component Analysis (PCA) algorithm or extract features and Artificial Neural Network to classify the output image.	(MUKHAIYAR & SAFITRI, 2019)	The goal is to differentiate the different features from one person to another person and recognize their features.	Images collected by authors	100 images for five persons	The face change in actual time with different expressions and variations of distance.	Artificial Neural Network	80 %	The model contains 100 neurons and seven hidden layers. That means finding the time to work on it will be difficult.
Face Emotion Classification	(Devi, 2020 #6)	facial emotion classification	JAFFE	213 face images of 10	observe the originality	facial emotion classification using	99%	The model assessed

using AMSER with Artificial Neural Networks				models for apanese ady. Every subject is considered an emotion.	of the human face	AMSER algorithm along with ANN		Japanese females but did not distinguish between different people.
Neural networks	(SEIBOLD, 2020 #7)	neural networks for face morphing	Gathering images from different databases	1900	Face morphing attacks	A biometric facial recognition system	98	To increase the robustness of the neural network models, the model requires pre-training.
A model for face recognition using deep neural networks	(ZHAO, 2020 #8)	Extract facial features to improve the face recognition	CAS-PEAL	99,450	Different face presentation attacks can attack current face, recognition models.	Convolutional neural a network model for face recognition combine with PCA algorithm to extract the features	98.52	Convolution combined with pooling is a powerful combination.
Face Recognition model based on Hybrid Algorithm for Features Extraction	(ALHASHMI, 2020 #9)	face recognition	ollected Dataset	00	Variation in face form, the illumination effect, and the combined background of the image	Artificial Neural Networks (Gelb & Metz) using Elman Neural Network (Chu et al.)	91	If the model provides dynamic learning rate values, it will provide the best accuracy.

					.			
A model for recognizing the face before and after plastic surgery. The model based on Neural Network.	(Sable, 2021)	To build a model for obtaining an effective facial recognition of plastic surgery	Collected Dataset	150 faces	The skin and feature face exchanged according to plastic surgery.	Neural Neton	performance gradient is 3.0156e 08 mean	The model evaluated 150 faces and required additional images for training.
A model was proposed to identify age-related features and face recognition.	(Huang et al., 2021)	Face recognition is affected by age variation.	n CACD-VS, FG-NET, and CALFW	64446	Face recognition is affected by age variation.	convolutik	99.55 5.62 7.18	It is a complex model. It takes time away.

As shown in Table 2.4, many other models are used for face recognition with different techniques. A model for recognizing the face and its expression based on Artificial Neural Networks was founded in 2016. The model scored 65% of the 60 images evaluated. However, the number of images is only 60 pictures with low accuracy (Owayjan, 2016). A model based on a hybrid network combined with a Viola-Jones-based system is used for face recognition. The model enhanced the accuracy of the results. The model was tested on 354 images from the CMU/VASC dataset and yielded 82% accuracy. Ten nodes are used for ten features in the model, but the ten hidden nodes always produce good results. Also, the model minimized the errors by using the backpropagation algorithm for network weights. (Murphy, 2016).

The artificial neural network with multi-layers was combined with a Gabor filter to classify facial expressions in 2017. Then, the accuracy increased to 85.7%. This model is used to detect and categorize facial expressions. The model was tested using the JAFFE dataset, containing only Japanese females (Verma, 2017). The model assessed Japanese females but did not assess people from various backgrounds. A model also recognizes the face based on the video.CNN is used to identify blur-free images and automatically extract features. Also, to improve the strength of CNN features for occlusion and pose variations. The model was evaluated by three databases: COX Face, PaSC, and YouTube Faces. These databases have more than 6277 videos. The average accuracy is 95%, but the time cost has slightly increased (Ding & Tao, 2017).

In 2018, an ANN-based model improved comparing faces at different times (morning, afternoon, and evening). However, the average accuracy was 38.177%, which is too low (Panjaitan et al., 2018). Moreover, another model for recognizing the face uses a single hidden layer analytic with a Gabor feedforward network. The goal is to reduce the error rate by improving the feature extraction process and the classification process. The model was evaluated on six databases: FERET b-series, COX, CMU-PIE, Multi-PIE, LFW-a, and FERET f-series. The model improves the time cost even though the model is complex by training a single sample per person. The average accuracy is 96.7% (Oh et al., 2018).

A model based on an enhanced artificial neural network was already be established by 2019 to improve face recognition accuracy. The neural network includes two hidden layers, each with 100 neurons, suggesting a more complicated and time-consuming model. A dataset collected by the authors was used to evaluate the model. HUSSEIN (2019) gave the model an 82%.

Another model extracted features using the Principal Component Analysis (PCA) technique and classified the resulting pictures using an Artificial Neural Network. Another model extracted features using the Principal Component Analysis (PCA) technique and classified the resulting picture using an Artificial Neural Network. This model aims to distinguish different characteristics from one person to another and recognize one's characteristics. The model was found to solve the issues of changing the face features at the actual time and detecting the face with different expressions and variations of distance. This model has 100 neurons with seven hidden layers. That means finding the time to work on it is difficult. The authors evaluated the model using a database collection. These images belong to five people, each of whom has 30 images. For 100 images, the highest accuracy is 80% (Mukhaiyar & Safitri, 2019).

In 2020, Face Emotion Classification using AMSER with Artificial Neural Networks was used for facial emotion classification. The model evaluated the Japanese female dataset known as JAFFE, but it did not evaluate different people from different locations. The model gave a score of 99% (Yoganand et al., 2020). Also, another model is used for biometric facial recognition based on neural networks. The model evaluated gathering images from different databases with 98% accuracy. The model requires pre-training to increase the robustness of the neural network models (Seibold, 2020). Furthermore, a face recognition model based on convolution neural network convolution mixed with pooling provides a solid face combination. The model uses PCA to extract the features. This convolution model combined with pooling is a powerful tool. CAS-PEAL evaluated the model as a dataset for face images. The model had a 98.52% accuracy rate (ZHAO, 2020). Additionally, the Elman Neural Network (Chu et al.) model were designed to improve the accuracy of Artificial Neural Networks (Gelb & Metz). This model is based on a hybrid feature extraction method and is used for face recognition. The accuracy is 91%. If dynamic learning rate values are assigned to the model, its accuracy would improve (Alhashmi, 2020).

In 2021, a model could tell the difference between a face before and after plastic surgery. The model is based on a Neural Network that has been optimized with particle swarms. As a result, this model aims to create a model for efficient facial recognition in plastic surgery. A collected dataset evaluated the model with 150 faces in a dataset collected by the authors. The model performance gradient is 3.0156E-08 (Sable, 2021).

A convolution neural network-based model was suggested to identify age-related characteristics and face recognition. As a result, developing this model aims to reduce the impact of aging on face recognition. Various datasets are used to assess the model, like CACD-VS, FG-NET, and CALFW. The model accuracy is different from one database to another. FG-NET gives 57.18%, CACD-VS gives 99.55%, and CALFW offers 95.62%. The model is complex, so it costs more time (Huang et al., 2021).

2.7.4 Model Agnostic Meta-Learning (MAML)

The Model Agnostic Meta-Learning (MAML) algorithm was used in different fields, and it was recently used for face recognition.

Table 2.5: Face Recognition models uses MAML

Model	Citation	Objective	Dataset	No. Images	Method	Result	Limitation
A model for face recognition	(Peng et al., 2021)	To build a face recognition model that increases the recognition accuracy	Collected by author	6	Deep learning neural network with MAML	92.6%	Small sample
A model to detect if the person is wearing a mask or not	(Zehtabet al,2020)	To build a model for Determining whether the incoming users are wearing a mask or not	Celeba	5000	A meta-learning-based classifier combines with neural network-based computer	90	The model is faster than other models

A model for face recognition based on a deep learning network combined with MAML is presented in Table 2.5. To begin, the dataset was split into two parts: meta-training and meta-test. Second, the fine-tuning model is obtained using the quadratic gradient updated model for model optimization. Third, the MAML method is used to categorize the new category's face pictures. The algorithm is evaluated on the face picture database to ensure its efficacy. The model accuracy was 92.6%, the model was tested on six faces only. The authors mentioned that the models were satisfying with speed, but the model is fast because the dataset only has six faces (Peng et al., 2021). Another model was founded used the meta-learning-based classifier combined with a neural network-based computer to determine whether the incoming users were wearing a mask or not; the model used CelebA as a database that has 5000 images. The model accuracy reaches 90% (Zehtab et al.,2020).

2.7.5 Aging Process Techniques Recent Development in Face Recognition

Meta-learning, often known as learning- to learn, may be defined as improving a learning technique via rounds of learning. Computer methods are regarded as the primary domain for employing meta-learning techniques, owing to their impact on training, which can solve some recognition and vision challenges. As shown in Table 2.6, Classification is one of the problems that meta-learning has helped translate for computer methods. In 2016, a model for few-shot and few-classes learning was based on the LSTM-based model for meta-learning. The model improves accuracy to 60.60% of the time after five shots. The difficulties lie in increasing the number of datasets. The Mini-ImageNet dataset evaluated the model (Ravi & Larochelle, 2016). In 2017, a model for few-shot learning was created using a neural network with meta-learning. The CUB-20 dataset was used to test the model. The model may be used to create a model for few-shot learning since it is simple and effective (Snell et al., 2017).

MT-net, a model based on a neural network and a meta-learning approach, was released in 2018. The accuracy was 96% with twenty methods and one shot. Great accuracy is achieved with a small quantity of data and a mix of meta-learning and a network. Omniglot and MiniImagenet, two databases, assessed the model (Lee & Choi, 2018). In the year 2020, a categorization model was employed—a convolutional neural network-based model. The accuracy improves when both CNN and MAML are used. Omniglot assessed a model based on 20 pictures (Pawar et al., 2020).

Another challenge for computer methods is object recognition, which meta-learning has helped solve. In 2018, a model based on model regression networks (MRN) and model-agnostic meta-learning (MAML) for human motion prediction was developed. The Human 3.6M (H3.6M) model was used to evaluate the model. The accuracy is 68.19% for a 5-shot and 53.26% for a 1-shot. As demonstrated, the accuracy remains low (Gui et al., 2018). For feature extraction and object recognition in 2019, a model based on a Convolutional Neural Network with a low-short learning algorithm was used. The MS-COCO dataset evaluated the model with 5000 images. The model's accuracy is 70.8%, which is low but better than previous studies (Kang et al., 2019).

Landmark prediction is yet another challenge for computer methods, which meta-learning has aided in overcoming; it is a model for motion prediction based on a Deep Neural Network with MAML. The model was evaluated by Berkeley MoCap: new actions. The model accuracy is 73.8%. The accuracy is low, but it is good compared to previous studies (Alet et al., 2018).

Another challenge for computer methods is object segmentation, and different models use meta-learning algorithms to improve the results. In 2017, a model for image classification based on the low shot learning method was combined with the Fully Convolutional Network (FCN). PASCAL-VOC-2012 evaluated the model. The model is faster than other models, supporting dense semantic image segmentation (Shaban et al., 2017). In 2018, a model based on few-shot learning combined with CNN was used for image recognition. The model was evaluated by PASCAL VOC 2012, and the accuracy was 87% (Rakelly et al., 2018).

Image and video generation is one of computer methods challenges, and various models use meta-learning algorithms to improve the results. In 2018, a model was based on an amortization network and a few-shot learning algorithm for image classification. On the two datasets (Omniglot and miniImageNet), the model achieves an average accuracy of 95% (Gordon et al., 2018). In 2019, a Convolution Neural Network CNN model with few-shot learning would extract human face parameters and generate talking face images. The model provides a real solution. The authors used limited time for training to control the training phase. Voxcelec1 evaluated the model with selected 50 videos, and Voxceleb2 with 32 videos (Zakharov et al., 2019).

Table 2.6: The different face challenges

Challenges	**Model**	**Citation**	**Method**	**Dataset**	**Limitation**
Classification	Modified Model-Agnostic Meta-Learning	(Pawar, 2020)	A model based on a convolutional neural network	Omniglot :20 images	When combined with both CNN and MAML, the accuracy improves.

	MT-net is a model based on neural network and meta-learning methods.	(Lee & Choi, 2018)	Neural Network	Omniglot and MiniImagenet	With 20 ways and one shot, the accuracy was 96%. The combination of meta-learning and a network results in high accuracy with a small amount of data.
	A model for few-shot learning based on neural network combined with meta-learning	(Snell et al., 2017)	Neural Network	CUB-200	Because the model is simple and effective, it can be used to build a model for few-shot learning.
	A model for few-shot and few-classes learning based on n LSTM-based model for meta-learning	(Ravi & Larochelle, 2016)	Deep neural networks	Mini-ImageNet	The model improves accuracy 60.60% of the time after five shots. The difficulties associated with increasing the number of datasets
Object recognition	A model based on model regression networks (MRN)and model-agnostic meta-learning(MAML) for human motion prediction	(Gui et al., 2018)	Model Regression Networks	Human 3.6M (H3.6M)	The accuracy is 68.19 for a 5-shot and 53.26 for a 1-shot. The precision is still lacking.
	A model for detecting the objects by features extracting based on CNN	(Kang et al., 2019)	Convolutional Neural Network	MS-COCO With 5000 images	70.8% low accuracy
Landmark Prediction	A model for motion prediction based on Neural Network with MAML	(Alet et al., 2018)	Deep Neural Network	Berkeley MoCap: new actions	73.8% low accuracy

Object Segmentation	A model for image classification based Low-shot learning method with FCN	(Shaban et al., 2017)	Fully Convolutional Network (FCN)	PASCAL VOC 2012	The model is faster than other models, and the model supports dense semantic image segmentation.
	Few-shot learning combine with CNN used for image recognition	(Dong & Xing, 2018)	Convolutional Neural Networks (CNNs)	PASCAL VOC 2012	When compared to previous models, performance has improved.
	Few-shot learning with guided networks used for video object segmentation	(Rakelly et al., 2018)	Guided networks	PASCAL VOC 2012	87% accuracy
Image and Video Generation	A model based on an amortization network and a few-shot learning algorithm for image classification	(Gordon et al., 2018)	Amortization network	Omniglot and miniImageNet	95 accuracy
	A model based on CNN with few-shot learning used to extract human face parameters and generate talking face picture.	(Zakharov et al., 2019)	convolutional neural network	Voxcelec 1 50videos , Voxceleb 232 videos	The model offers an efficient solution. They make the most of their limited time by training.

2.7.6 Evaluation methods for face recognition models

Different methods were used for evaluating the face aging recognition models as holdout, cross-validation, comparing the accuracy, precision and recall, True Positive (TP), False Positive (FP), Ture Negative (TN) and False Negative (FN).

A model based on different techniques used for face recognition used the accuracy as an evaluation method (Shyam et al., 2014) with an accuracy of 95%. Also, generative adversarial networks had an accuracy of 82.9% (Antipov et al., 2017), similar to the recurrent neural network, which used accuracy of 98.40 % (Wang et al., 2016), also artificial neural network with an accuracy of 80% (Mukhaiyar & SafitrI, 2019) and same artificial neural network with 91% accuracy (Alhashmi, 2020).

Another way of evaluation is to calculate the Mean Absolute Error (MAE), the deep neural network model achieves MAE of 5.13 (Wang et al., 2017), and Neural Network combined with particle swarm optimization reaches MAE of 3.015 (Sable, 2021). An additional way of evaluation is to calculate the recall and precision of the models used in other research (Sundaram et al., 2016).

2.7.7 Summary of Current models Reviewed (Strengths and Limitations)

Different models with different techniques are used for face recognition. These models used different databases with different characteristics. They are used for face recognition with different datasets to recognize faces with different illuminations and poses but do not detect face aging. For example, the following models detect the same face with a maximum difference of two years, which is not considered face aging recognition. A model based on eigenvalues with high accuracy 99.40 % (Gaidhane et al., 2014), Quantum Neural networks (QNN) with 97.8% (Y. Xu et al., 2011), Vector Projection Length with 96.88% (C. Hu et al., 2014), Genetic Algorithm with 95.89% (Geng et al.2018), SVM with 98.4% (Biswas & Sil, 2017), and Kernel Collaborative Representation (KCR) with 99.3% (Dong Wang et al., 2015).

ANN models are used for face recognition with different accuracy and limitations. It was used to detect face expression with an accuracy of 65% (Owayjan, 2016). A hybrid network with ten hidden nodes improves the accuracy to 82% (Murphy, 2016). For occlusion and pose variations, CNN accuracy becomes 95% (Ding & Tao, 2017).

The wavelet transforms, and artificial neural networks are used in a model to perform human face recognition. As shown in Figure 2.11, the architecture of ANN has two hidden layers beside the input and output layer.

The model flow shown in Figure 2.12 has five stages: image acquisition, pre-processing, feature extraction, training and testing, and the last stage is classification. At the first stage, different images with different angles are loaded into the model. The second stage is to transfer these images to the grey level as the pre-processing stage. The third stage is feature extraction using wavelet transform. The fourth stage is training the ANN. Finally, the fifth stage is to decide whether the input image belongs to the same class (Burman et al., 2020). The accuracy was 95% for face recognition poses, and the model got challenges dealing with low pass and high pass frequencies. Therefore, the researchers recommended using a wavelet with ANN for face recognition in the future.

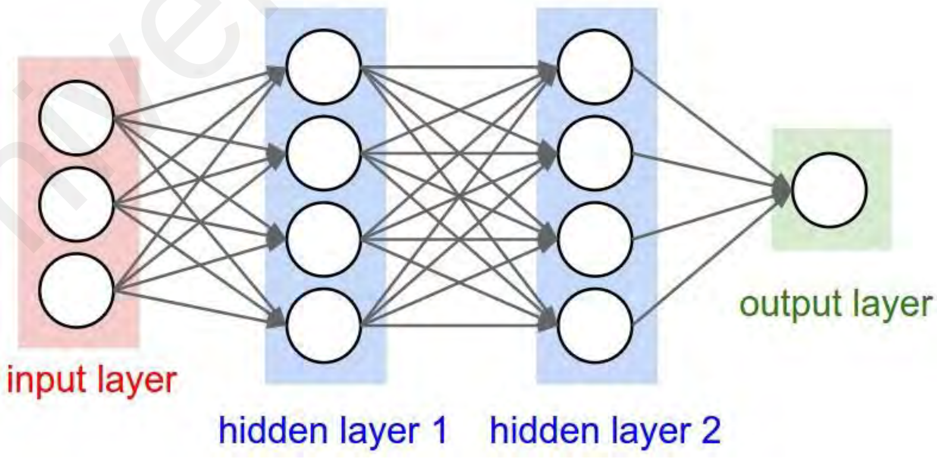

Figure 2.11: ANN Layers

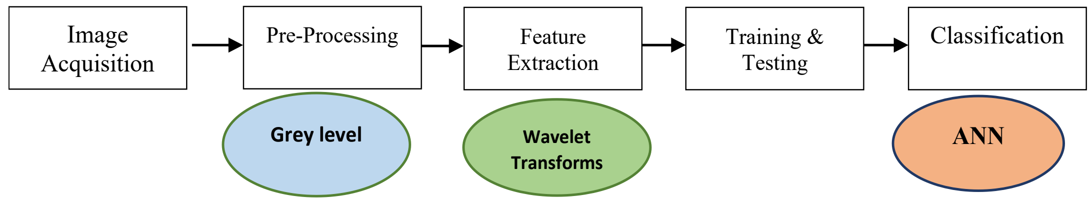

Figure 2.12: The ANN stages

2.8 Limitations of Current Research: (Open Issues in Face Recognition models)

Face models face a variety of challenges and issues, most of which arise in everyday life situations.

2.8.1 Availability of Images

Face recognition accuracy, age estimate accuracy, and facial aging accuracy may all be improved in various ways. Increasing the number of pictures necessary in databases for the same face in different postures is one of these ways. Nonetheless, collecting multiple photos for the same individual or finding images at different ages is difficult. As a result, high-performance techniques and approaches that are more difficult to adapt to such difficulties are necessary.

Collecting labeled data for accurate age estimation is more complicated than collecting image data for age classification (Dehghan et al., 2017) or face recognition issues (Everingham et al., 2015). The high prevalence of human mistakes in real-world age estimate subjects causes this problem. The estimated error made by humans is more important than the estimated mistake made by computerized programs. It is also difficult to rely on human annotators to correctly categorize all of the faces in the databases based on age. Even though several different techniques have proven their accuracy, there are still some open issues, problems, and limitations. Many academics have begun to solve these issues, and they have applied various methods on the same image, such as cropping, scaling, and rotation. As a result, different pictures of the same individual can be acquired in various postures. Furthermore, the picture lighting may be altered to provide diverse images. This approach has the potential to expand the number of pictures, and as a result, it might be considered one solution to the problem. However, human annotators are not feasible because labeled images include a large amount of human error, necessitating a review of the annotated image.

2.8.2 Computation Time

Face recognition is accomplished using the Multiple Convolution Neural Network (CNN). Despite the remarkable accuracy of CNN findings, processing takes a long time and requires many pictures to be processed at a sluggish rate. Because of the numerous CNN layers, each picture requires 88 forward passes on the CNN channel, making real-time or near-real-time operations impractical. Most models suffer from similar limitations and challenges (G. Antipov et al., 2016). To decrease the running time, a Graphics Processing Unit (GPU) can be used, which is effective in image processing, where a considerable number of images is performed in parallel. In addition, the GPU can help to make the training phase of the model faster.

2.8.3 Low Resolution

Any typical image should have a minimum resolution of 16 x 16 pixels. A low-resolution image is one with a resolution of less than 16 x 16 pixels. Small scale standalone cameras, such as CCTV cameras in streets, ATM cameras, and supermarket security cameras, can provide these low-resolution pictures. These cameras can only catch a tiny portion of the human face, and because the camera is not very close to the face, they can only capture a 16 x 16 face region. Because most of the pixels are lost in such a low-resolution image, it does not give much information; hence recognizing peoples' faces may be a difficult task (Zangeneh et al., 2020).

2.8.4 Expressions

The face is one of the essential biometrics since its distinctive characteristics are critical in determining human identity and emotions. Diverse conditions produce various moods, leading to various emotions and changes in facial expressions (Fontaine et al., 2017).

Another essential element to consider is the different expressions of the same individual. In particular, macro-expressions such as pleasure, sorrow, rage, contempt, fear, and surprise are examples of human expressions. Micro-expressions are involuntary face patterns that display fast facial patterns. Due to changes in one's emotional state, macro and micro-expressions appear on one's face, and efficient recognition becomes problematic in the aftermath of such emotions, which are numerous (X. Liu et al., 2021).

2.8.5 Plastic surgery

Face recognition techniques have been shown to fail to recognize individuals' faces following cosmetic surgery. This is a circumstance in which an individual's facial appearance is wholly altered, transforming them into entirely different people. The plastic surgery on the human face causes differences in skin texture across photographs of the same person, making facial identification difficult. Rhytidectomy is a typical example of cosmetic surgery and how it may alter the appearance of the face. This technique changes the look of the face (Rathgeb et al., 2020), and it could speed up the transition from an older to a younger skin texture, resulting in a change in skin texture. In addition, eye lifts, nose contouring, and jaw augmentation are substantial alterations to the face that change facial appearance. Face recognition algorithms have been designed to resist the effects of cosmetic surgery (Sable et al., 2018). The plastic surgery face dataset was developed to test face recognition systems on the subject of plastic surgery. The authors presented an entropy-based volume SIFT technique for identifying faces after plastic surgery (Sable et al.,2019). Their method isolates the scale-space structure's important areas and volume for which the information rate is known. Because entropy is a higher-order statistical characteristic, it has the most negligible impact on unpredictable changes in the face. The support vector machine is used to classify the data using the associated characteristics. On the plastic surgery dataset, experiments were conducted for various plastic surgery situations. They reported an 88% recognition rate for blepharoplasty, an 87% recognition rate for brow lifts, and an 86% recognition rate for lip shaving, respectively. The authors used an edge-based Gabor feature representation approach to identify surgically altered faces (Oloyede et al.,2020). To treat the textural difference created by cosmetic surgery, edge information that is reliant on the shape of critical components of the face is employed.

2.8.6 Occlusion

Occlusion refers to blocking one or more portions of the face, preventing the entire face from being used as an input picture. One of the most challenging issues in face recognition systems is occlusion (Zeng et al., 2020). It is common in real-world scenarios and is caused by beards, mustaches, and accessories (e.g., goggles, caps, masks). Including such components diversifies the subject, making computerized facial recognition challenging to crack (Zeng et al., 2020).

2.8.7 Pose

Pose changes are pretty sensitive to facial recognition systems. When a person's head moves and their viewing angle changes, the posture of their face changes. Changes in face appearance caused by head motions or different camera POVs usually create intraclass variances, dramatically lowering automated face recognition rates. In addition, when the rotation angle is increased, identifying the true face becomes more difficult. If the database includes the frontal image of the face, it may result in poor or no recognition (Yoganand et al., 2020).

2.8.8 Aging Process

There are a variety of reasons why automated age estimate is difficult. It is one of the most significant challenges to automate age estimation between faces on the same age scale and aging features on a person's face (G. Antipov et al., 2016). It is a wild aging nature. However, progress in face aging is slower than in other areas due to the lack of a dataset with thousands of images to collect and label. This dataset is essential for training deep network models (G. Antipov et al., 2016). Most current age estimate research focuses on calculating a person's biological age (defined as the number of years from their birth date) (G. Antipov et al., 2016).

Previous research indicates that there is a scarcity of evaluations of aging systems. Additionally, it is challenging to gather photographs of the same person over a long period. Age changes are frequently combined with all feature changes, and they are frequently mixed with other variation changes, e.g., illumination (Suo et al., 2010). Consequently, aging is one of the primary reasons for the changes in facial appearance (H.-F. Yang et al., 2015). Generative Adversarial Networks (GANs) are used to produce actual pictures. These pictures are indistinguishable from the original images. GANs can also change the appearance of the human face, such as aging, wrinkles, hair color, and eye color. Hence, many models have deployed the GAN approach to extract the aging features of the character.

2.8.9 Age

These days, there is much commercial software for age estimation. However, they work efficiently with images for older children and adults (the minimum age is 18), although the new iPhone has a minimum age of 15 (Fu et al., 2010). Additionally, a substantial number of papers showed that the accuracy rate is deeply connected to age. Many different models have been applied to deal with children.

Table 2.7: The problems in face recognition with the proposed solution

Problem	Citation	Proposed solution
Availability of Images	(Dehghan et al., 2017)	Collect images
Time	(Antipov et al., 2016)	Using GPU
Low Resolution	(Zangeneh et al, 2020)	Try to improve the images by adding pre-processing image processing algorithm
Expressions	(Fontaine et al, 2017) (Liu et al., 2021)	Build models detect the details of facial features
Plastic surgery	(Rathgeb et al., 2020)	Build models detect the details of facial features with the original measurement
Occlusion	(Zeng et al., 2020)	Build models detect the details of facial features
Pose	(Yoganand et al., 2020)	Build models detect the details of facial features
Aging Process	(Antipov et al., 2016)	Build models detect the details of facial features
Age	(Fu et al., 2010)	

2.9 Summary

This chapter studied the different algorithms for face recognition and the various datasets used on face recognition. Also, this chapter has shown the datasets that can be used for face aging recognition. Theoretical concepts have been presented after reviewing the face recognition models, artificial neural networks models, and agnostic meta-learning (MAML) models.

Many models have been reviewed for face recognition and face aging recognition. It was observed from the different reviewed models that most of them concentrated on face recognition in general, not on face aging recognition. Also, it is proven that face age recognition is still a challenge in the face recognition area. Furthermore, it shows the face recognition open issues.

The next chapter presents the research methodology that helped in building a proposed model. This model's aim is to solve the face aging recognition problem.

CHAPTER 3: RESEARCH METHODOLOGY

3.1 Introduction

The open issues and the challenges in face aging recognition were justified in the first chapter. Also, the problem statement was defined, and the objectives were determined to solve this problem statement. A review of the existing methods for the relevant topics is shown in Chapter Two. Additionally, different types of databases were discussed, besides the architecture of the different models used for recognition. Now, the construct of the proposed model and the steps to build this model have become vivid. This chapter presents the main phases of the research methodology, as shown in Figure3.1. Phase 1 examines the literature evaluation and issue identification. Phase 2 describes the study's goals. The suggested model architecture and execution are shown in Phase 3. Finally, phase 4 displays the results as well as the analyses and evaluations for these outcomes. The following subsections describe each step.

3.2 Research roadmap

The research study's plan is shown in Table 3.1. The goals are to summarize the essential parts and map out the connections between them.

Table 3.1: Research Roadmap

Thesis Title	- Face Aging recognition based on ANN and MAML techniques		
Research Problem	- The absence of a high-performance model for automated facial aging recognition. - When the age difference between the query image and the target image is significant, low accuracy in aging occurs. - The inability to extract facial characteristics that are required to improve the accuracy of face aging recognition. - A scarcity of datasets with a gap of more than ten years.		
Aim/Motivation	- To build a model for face aging recognition with acceptance accuracy, can detect the same person's face with more than ten years gap		
Research	1. What are	2. How to develop an	3. How to

Questions-RQ	the limitations of existing face aging recognition techniques?	efficient model for face aging recognition that can detect the same face but in different gap times?		evaluate the proposed model regarding the performance?
Thesis Objectives - TO	1. To identify the limitations of existing frameworks and techniques used in face aging recognition.	2. To develop a unified model for aging recognizes the face using an artificial neural network with (Vedel et al.) and enhancement MAML.		3. To evaluate the performance of the face aging recognition model using the previous studies' approaches.
Research Methodology	- Follow the systematic review papers. - Review the papers. - Extract the limitations.	- Choose the techniques ANN and MAML. - Build the model using Matlab. - Test the integrated compounded. - Collect the datasets. - Train the model.		- Evaluate the models by three datasets. - Compare the results with previous studies.
Thesis Chapters	Chapter 3 Research Methodology	Chapter 4 Framework Design and Implementation	Chapter 5 Evaluation Results and Discussion	Chapter 6 Conclusion and Future Research Work
Chapter Objectives	- Present the methodology research. - Identify the stages	- Present the proposed model design. - Show the implementation of the	- Test the proposed model by the three datasets. - Compare the	- Summarize the proposed study. - Provide recommendations for the proposed

	of research.	proposed model.	proposed model with previous results. - Provide analysis for the result	model. - Present the future work to improve the proposed model.
Thesis Contribution Thesis Objective	C1:The effective face-aging recognition framework that adopts an ANN combined with MAML (TO1 AND TO2): Describe how can build an effective model based on the previous studies	C2: Improve the input images (TO1, TO2) : Describe how can improve the input images based on previous studies. Choose a suitable algorithm for the pre-processing image	C3: Strengths and limitations of existing models (TO1).	C4: Description of Face aging recognition challenges (TO1)

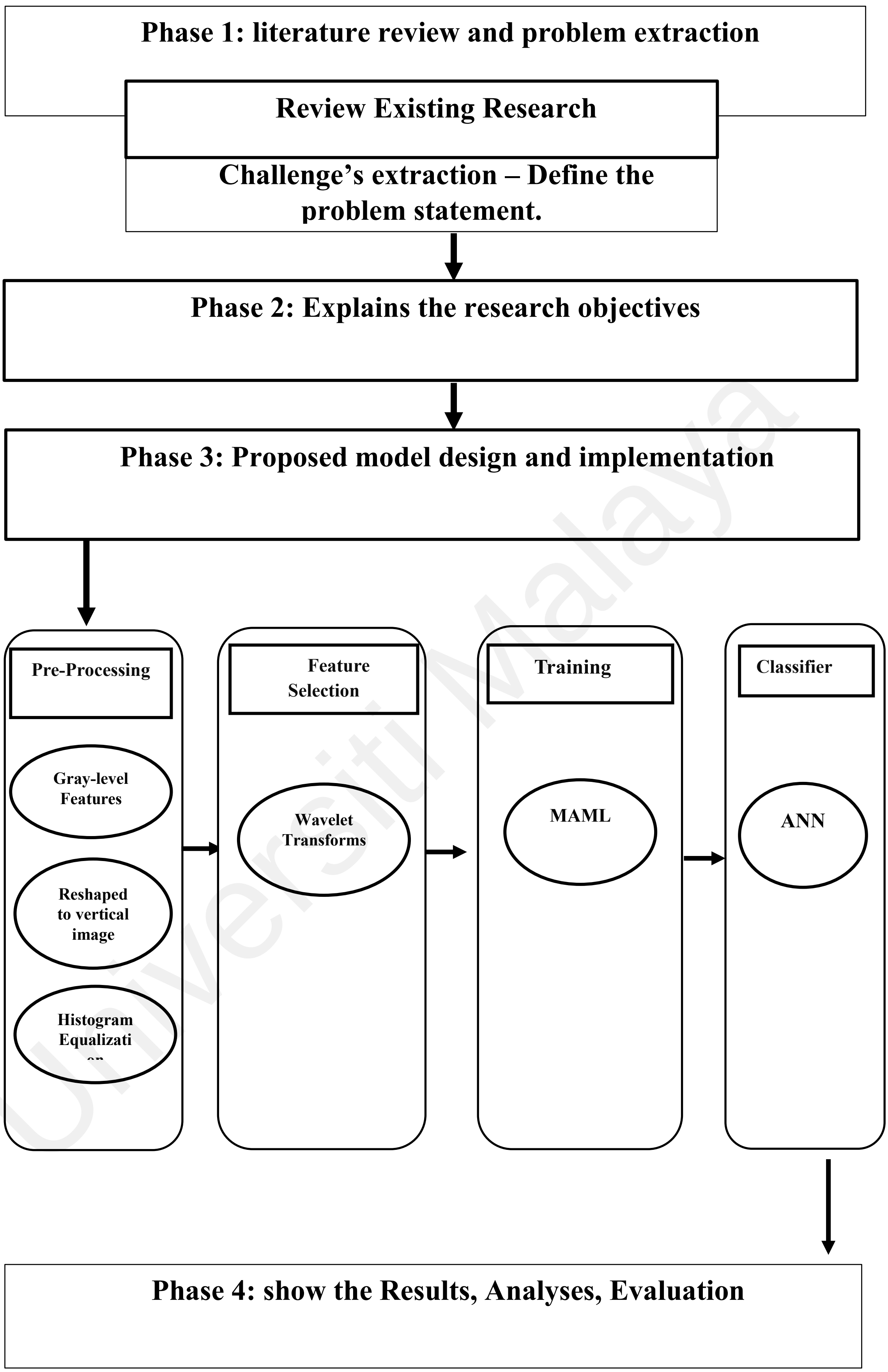

Figure 3.1: Research Methodology stages

3.3 Research Phases

As shown in Figure 3.1, The research methodology has four phases. The first phase is the literature review and problem extraction and discussed in Chapter 2. The first phase has two subsections that review existing research, extract the challenges and open issues, and then define the problem statement. The second phase is writing and explaining the research objectives. During this phase, the objectives have been written. The research becomes more determined after preparing the objectives, and the scope becomes clear. The suggested model design and implementation procedure is the third phase. Pre-process, features selection, training, and classifier are the four subsections of the third phase.

Pre-processing for the input picture is the first step. The model was founded to improve the input images, using the Gary-level features method, a vertical reshape method and a histogram equalization method. The specifics of the methods are explained in Chapter 4. The second section is feature selection. Feature extraction is essential in image recognition in face recognition. This step describes the wavelet transforms algorithm, which is used to extract information from pictures. The network is trained in the third phase using modified MAML, one of the most used meta-learning techniques. MAML calculates the optimum starting weights for the network to learn new tasks quickly, even if a few labeled samples are used in training. The algorithm was used to train the model to learn any new function quickly using only a few datasets. The updated MAML was created to be able to choose the components of the feature substantially. The classifier stage is the fourth part. The ANN was employed as a classifier to determine whether or not it was the same face for the same individual. An artificial neural network (Vedel et al.) is made up of three layers of neurons: input layer, hidden (responsible for pattern extraction and the majority of internal processing), and output (produces and presents final network outputs).

The results, analysis, and assessment are presented in the fourth step. This section's goal is to double-check the suggested model written in MATLAB. Using the Cross-Age Labeled Faces in the Wild (CALFW) database and its data set, this study tested the model's accuracy in face-aging recognition (middle east data set meds). The three datasets are as follows:

1. **Cross-Age LFW (CALFW)** is built to evaluate face verification algorithms under a significant age gap. Positive pairs- (the difference between the age is less than ten years). Negative pairs- (the difference between the age is more than ten years).

2. **The AT&T** face database comprises 400 pictures, all of which are of the same person, same age, and different positions. All of the photos are greyscale and saved in PGM format and are 92 x 112 pixels in size.

3. **Middle East Data set MEDS**: The author collected a data set from people of different ages from the Middle East. The constraint in this picture is that the person should be older than 15 years, and the difference between the two pictures is ten years or more for the same person. It contains 100 person faces, and every person has at least two pictures; hence, 200 pictures were collected.

At the verification process, the model must decide if the input image is for the same person or not. Additional evaluation methods were used, such as specificity (True Negative Rate), sensitivity (recall or true positive rate), false Positive Rate (FPR), precision, f1- score (F- measure) and Fales Negative Rate (FNR).

3.3.1 Phase 3: Proposed model design and implementation

Building the proposed model started with the analysis and collecting information. The model development process is a step-by-step process that comprises phases including planning, analyses, design, deployment, and assessment. Phase 3 started with designing the model and focusing on how to achieve the model's objective. This phase is mainly based on two components:

3.3.1.1 Subcomponent

This section explains the subcomponent of the model and the relations between them to achieve a specific goal. Moreover, at this stage, the input and output should be clear. The model design is based on four processes: pre-processing, feature selection, training, and classifier. They should be explained in detail in Chapter 4.

3.3.1.2 Technology

At this point, the various technologies that would be utilized to construct the model should be determined. It also analyzes the technical viability of each implementation option at this level. It examines the solution and evaluates if it can be supported by the current technology. The researcher should consider if existing technological resources should be updated or new ones introduced to meet the new needs. It also guarantees that the candidate system responds appropriately to the technological upgrade to the degree it can do so.

- **Hardware and Software Setup**

The model was built using MATLAB2016A software/Intel® Core™ i7 CPU @ 2.70 GHz and 8 GB memory. A total of 500 images were tested with PGM format. The circumstances of how the model was evaluated are discussed in another section.

3.3.2 Phase 4: Results, Analyses, Evaluation

Phase 4 of the research methodology shows three stages: results, analyses, and evaluation.

3.3.2.1 Result stage

After the model has been built at the implementation phase, the model should be tested before the evaluation stage.

Testing is a method of evaluating software's functionality and correctness against predefined criteria to improve the model's quality and reliability. Aside from the testing method, quality assurance for the model should occur while being produced.

The model results are presented in detail in Chapter 5; three datasets evaluated the model. Every dataset with its own characteristics will be explained in detail in Chapter 5.

3.3.2.2 Analyses stage

The result of the analysis phase helped in the construction of the most effective algorithms to achieve the objectives acknowledged by this research. In this section, the data collection for the proposed model was designated. However, the analyses of this stage depended on the outcome of the proposed model. To confirm the results of the analyses stage, many experiments were done.

3.3.2.3 Evaluation stage

A model evaluation is a kind of measurement used to determine if the model is the best possible fit for the needs of a given client. Model evaluation tries to approximate a mode's overview accuracy, performance, precision, and recall.

Then the model result by comparing with the result of similar models from the previous studies are shown in Chapter 5.

3.4 Summary

Chapter 3 presents the research methodology that followed in choosing the problem statement and justifying this problem statement until the model is built and evaluated.

The research methodology has four stages; the first stage is a literature review, which determines the open issues, followed by choosing the problem statement. The second stage is explaining the research objectives. Then the third stage is the proposed model design and implementation, which has four stages: pre-processing, feature selection, training and the classifier stage. Finally, the fourth stage is the analyses and evaluation stage.

So the research methodology reflects the thesis objectives and helps achieve them, leading to solving the problem statement.

CHAPTER4: FRAMEWORK DESIGN AND IMPLEMENTATION

4.1 Introduction

In the previous chapter, the research methodology stages were presented. This chapter proposes the new framework for face aging recognition based on artificial neural networks (Vedel et al.) with modified model-agnostic meta-learning. It is clear from the related work that face recognition has many open challenges. It seems evident that some of the challenges got solved by different techniques. ANN is used for face recognition. In this thesis, a modified MAML is added to ANN to improve the aging recognition issue at the training phase. This study motivated the discovery of the same model in this chapter, which is organized as follows: Section 4.1 is the Introduction, Section 4.2 presents the framework ANN_MAML, and Section 4.3, the summary.

4.2 Proposed model: ANN_MAML

This section introduces the proposed framework, depicted in Figure 4.1, with the human image serving as the framework's input image.

The main framework steps:

Step 1: The first step is the image pre-processing stage. The pre-processing stage has three stages:

1. Grey-level features
2. reshaped to a vertical
3. Histogram equalization

Step 2: The second step is to select the features to use Wavelet Transforms.

Step 3: The third step is the training stage used MAML.

Step 4: The fourth step is the recognition phase based on Artificial Neural Network (Vedel et al.).

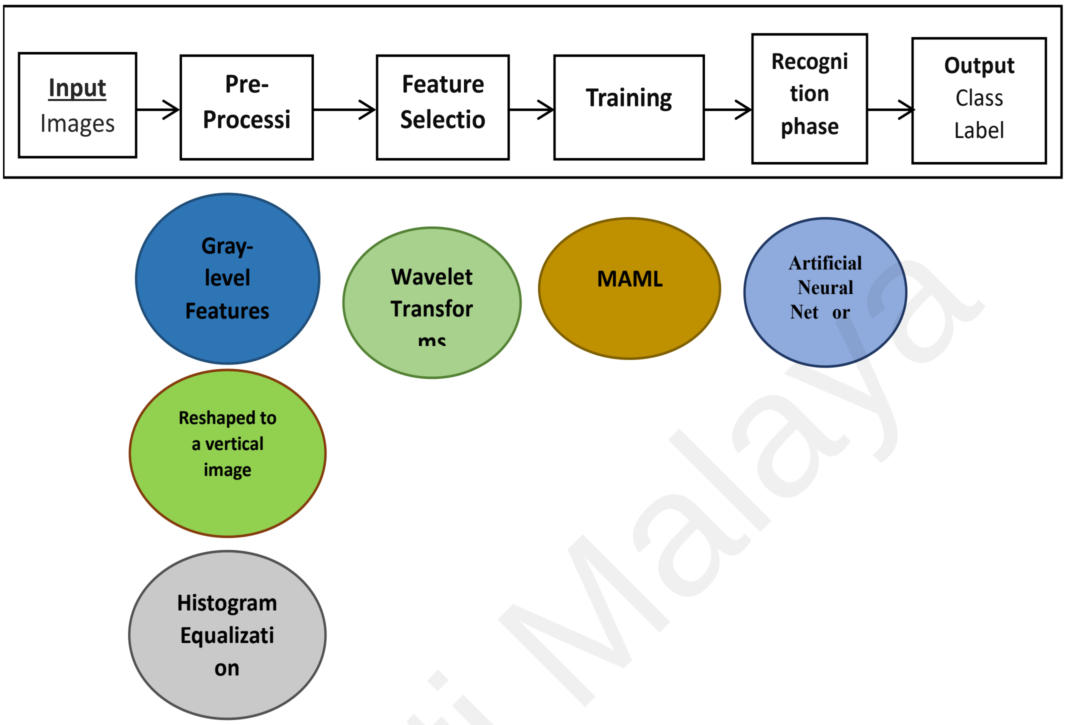

Figure 4.1: Architecture for face aging recognition framework MAML Artificial Neural Network

4.2.1 Step 1: Pre-processing Phase: Image segmentation and detection

The first step is image pre-processing. The pre-processing stage has three stages:

4.2.1.1 Grey-level features

The pixel on the binary image can only be value 0 or value 255. Compared to the greyscale image, the pixel can be valued between 0 and 255.

A histogram of the grey level shows how many pixels of an image share the same level of grey. The x-axis displays the grey levels from 0 to 255, and the y-axis displays the picture frequency. For determining a threshold, this information can be used.

Unfortunately, geometrical and structural elements are susceptible to segmentation errors. Furthermore, that results in an unclear solution depending on the point of view. The following three characteristics become evident given the grey-level picture and the strength transition for line segments:

1. The edges have different intensities.
2. The pixels around the edges replace with the mean.
3. As a binary feature, the direction of intensity changes, or the sign of the gradient.

View-point independence, simplicity, and resilience against segmentation faults are among the characteristics' advantages.

First, the edge image must be computed with an operator with the following properties: a one-pixel-width edge image, proper placement, and noise resistance. A simplified edge operator based on Haralick's facet model is employed in the tests, satisfying these requirements. Line segments will also be retrieved from edge pictures, which may then be segmented using sequential or more global segmentation techniques like the Hough transform.

The grey-level co-occurrence matrix (GLCM) is a popular texture feature descriptor derived from grey-level data.

Figure 4.2: Effect of grey level

4.2.1.2 Reshaped to a Vertical

As a result, specific photos taken by a camera and supplied to AI algorithms vary in size. Therefore an actual size for all images fed to the neural network algorithms should be established. The reshaped vertical method is used to unify the input images size. So all the input images need to be resized to be the same size before inputting them to the ANN.

4.2.1.3 Histogram equalization

Histogram equalization is a contrast adjustment approach in image processing that uses the picture's histograms (data illustrated in Figure 4.3). This approach generally improves the global contrast of many pictures, mainly when a small number of intensity values represents the image. The intensities can be better dispersed on the histogram with this modification, equitably employing the whole range of intensities. This enables locations with poor local contrast to get a boost in contrast. This is accomplished by histogram equalization, which effectively spreads out the densely packed intensity values that decrease picture contrast. The approach works well in photos with both bright and dark backgrounds and foregrounds. The technique, in particular, can improve the visibility of bone structure in x-ray pictures and the detail in photographs that are either over or under-exposed. The approach has the benefit of having a simple methodology that is adaptable to the input picture and has an invertible operator.

Let ni be the number of occurrences of grey level I in a discrete grayscale picture x. The likelihood of a pixel of level I appearing in the picture is

$$px(i) = p(x=1) = ni/n\ ,\ 0<i<L \qquad 4.1$$

px(i) is the picture's histogram for pixel value I normalized to [0,1], n is the total number of pixels in the image, and n is the total number of grey levels in the image (usually 256).

Figure 4.3: Effect of histogram equalization

4.2.2 Step 2: Feature Extraction

Feature extraction is the process of extracting features from a picture to improve accuracy and ease of identification. According to Yu Su, Shiguang Shan, Xilin Chen, and Wen Gao, global and local characteristics are essential for face representation and identification (Su et al., 2009). Therefore, there are two approaches for extracting features: global feature extraction and local feature extraction.

Image recognition, like face recognition, relies heavily on feature extraction. Facial feature extraction removes face component characteristics such as eyes, nose, and mouth from a human face picture. Eye localization and detection are crucial among all face features since they are used to identify the positions of all other facial features.

4.2.2.1 Feature extraction local, feature extraction global

Face feature extraction is required to initiate processing techniques such as face tracking, facial expression recognition, and face recognition. Because the facial pictures have similar geometrical characteristics, it is difficult to tell one from another. Face recognition relies on both global and local characteristics. Both global and local characteristics are retrieved from the input face pictures in the suggested technique. Feature extraction is necessary for face recognition in order to convert high-dimensional picture data into low-dimensional feature vectors. In general, there are two types of feature extraction methods: global and local. Principal Component Analysis (PCA) is the most well-known global technique. PCA is characterized as a global technique for face identification since it pulls facial features from the bases that describe the entire face. The bases, known as eigenfaces, are eigenvectors of the covariance matrix of face pictures and are considered face models.

The overall structural layout of face organs, as well as facial shape, are global characteristics. The low-frequency coefficients of the Fourier transform are used to extract global features from full face pictures. Real and imaginary components are concatenated in the low-frequency range into a single feature vector called Global Fourier Feature Vector (GFFV). Local characteristics occur often and are influenced by the position and orientation of facial pictures. Gabor wavelets are used to extract local characteristics. Local Gabor Feature Vectors are a set of feature vectors that organize Gabor features geographically (LGFV). The linear combination weights for eigenfaces are determined by projecting a face image onto the eigenfaces. These weights are used to represent the features of a person's face. Local features such as the nose, eye, jawline, and cheekbone are detected by the projection.

4.2.2.2 Wavelet transforms algorithm

This section describes the wavelet transforms algorithm, which is used to extract information from pictures. In the last ten years, the Wavelet Transform (WT) has become a widely popular technique for image analysis. WT was chosen for image decomposition due of the following reasons:

1. The resolution of the subpictures is decreased when an image is decomposed using WT. As a result, by working with a lower quality image, the computational complexity is drastically decreased. Harmon proved that a 16x16-pixel picture is sufficient for identifying a human face. The sub-image size is decreased by 64 times compared to the original picture resolution of 128x128, implying a 64-fold reduction in recognition computing burden.
2. WT decomposes pictures into sub-bands that correspond to distinct frequency ranges. The input requirements for the following significant step are easily met by these sub-bands, reducing the computational cost in the suggested system.
3. While Wavelet decomposition supports local information in both the space and frequency domains, Fourier decomposition only supports global information in the frequency domain.

As illustrated in Figure 4.4, a picture is split into four subbands. The LL band is a rougher representation of the original picture. The bands LH and HL record the image's changes in horizontal and vertical directions, respectively, while the HH band displays the image's higher frequency component.

This is the initial breakdown level. The decomposition can be extended to include the LL subband. An image is divided into subbands of various frequencies after performing a three-level Wavelet transform, as illustrated in Figure 4.3. If the resolution of an image is 128x128 pixels, subbands 1,2,3,4 are 16x16 pixels, subbands 5,6,7 are 32x32 pixels, and the subbands 8,9,10 are 64x64 pixels.

Figure 4.4: Face image with one-level, two-level and three-level wavelet decomposition

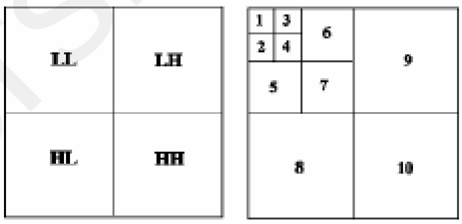

Figure 4.5: Wavelet Decomposition

The wavelet transform compresses the picture signals' energy into a limited number of wavelet coefficients. An image f (x,y) can be represented as follows using two-dimensional wavelet transforms:

$$F(x,y) = SJ + \sum_{j=1}^{J} D_j^V + \sum_{j=1}^{J} D_j^h + \sum_{j=1}^{J} D_j^d$$ 4.2 (Kudratilloev & Akhmedov, 2021)

where the two-dimensional wavelets are the tensor product of the one-dimensional wavelets as below:

$$\varphi(x,y) = \varphi(x) X \varphi(y)$$ 4.3 (Kudratilloev & Akhmedov, 2021)

$$\psi^v(x,y) = \varphi(x) X \psi(y)$$ 4.4

$$\psi^h(x,y) = \psi(x) X \varphi(y)$$ 4.5

$$\psi^d(x,y) = \psi(x) X \psi(y)$$ 4.6

Where J denotes the wavelet phases, the approximation picture is the first step, while the vertical, horizontal, and diagonal images are the next three.

Within the approximation picture, the original image's energy concentrates. Using only three wavelet steps, images exhibit the most essential components (significant component map).

The function $\varphi(r)$ and $\phi(r)$ are used to refer for scaling function and corresponding wavelet function, and both of them satisfy the dilation equations, with $\phi^m{}_n(r)$ and with $\varphi^m{}_n(r)$ being their dilations and translations, respectively.

$$\phi^m{}_n(r) = 2^{-m/2} \phi(2^{-m/2} t - n), n \in Z$$ 4.7 (Ryan, 2019)

$$\varphi^m{}_n(r) = 2^{-m/2} \varphi(2^{-m/2} t - n), n \in Z$$ 4.8 (Ryan, 2019)

It refers to a clockwise-oriented plane curve with parametric coordinates:

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} x_a^M(t) \\ y_a^M(t) \end{bmatrix} + \sum_{m=M-m0}^{M} \begin{bmatrix} x_d^m(t) \\ y_d^m(t) \end{bmatrix} \qquad 4.9 \text{ (Ryan, 2019)}$$

$$x_a^M(t) = \sum_n a_n^M \Phi_n^M(t).\, y_a^M(t) = \sum_n c_n^M \Phi_n^M(t) \qquad 4.10 \text{ (Ryan, 2019)}$$

are called the approximation signals at scale m and

$$x_d^m(t) = \sum_n r_n^m \Phi_n^m(t).\, y_d^m(t) = \sum_n d_n^m \Phi_n^m(t) \qquad 4.11 \text{ (Ryan, 2019)}$$

The detailed signals of scale m are referred to as detailed signals. Because detailed signals, which have a high frequency, are usually accompanied by noise. We used the approximation coefficients an $^M{}_n$, c $^m{}_n$ defined above as a planar curve descriptor for matching.

The steps required in extracting image features are as follows:

Normalize the pictures W (x1, x2) first, then subtract the average value of the image from the normalization. As a result, the image is correctly focused on the effective pixels. The images are separated using wavelets of varying sizes. This was done by extracting the appropriate signal properties on a frequency range spanning from low to high.

To extract signals of each frequency, rebuild wavelet decomposition coefficients.

D_0, D_1, ….., $D_{M\ is}$ are used to express the decomposed reconstructed signals from low-frequency and high-frequency coefficients. D can be referring to signal as

$$D = D_0 + D_1 + \ldots\ldots + D_M \qquad 4.12 \text{ (Ryan, 2019)}$$

The approximation component of the wavelet coefficients was used in the feature extraction approach in principal component analysis. If the approximation is ai, i=1,….., M, and that there are M images in the training set, then we have an image feature such that

$$y=w^T (a - A) \quad 4.13 \text{ (Ryan, 2019)}$$

where the mean image is:

$$A = \frac{1}{M} \sum_{i=1}^{M} a_i \quad 4.14 \text{ (Ryan, 2019)}$$

Furthermore, *w* represents the eigenvectors corresponding to large eigenvalues of the covariance matrix $(a-A)(a-A)^T$.

The Gabor wavelets family and the results of applying them to a facial image are shown in Figure 4.6.

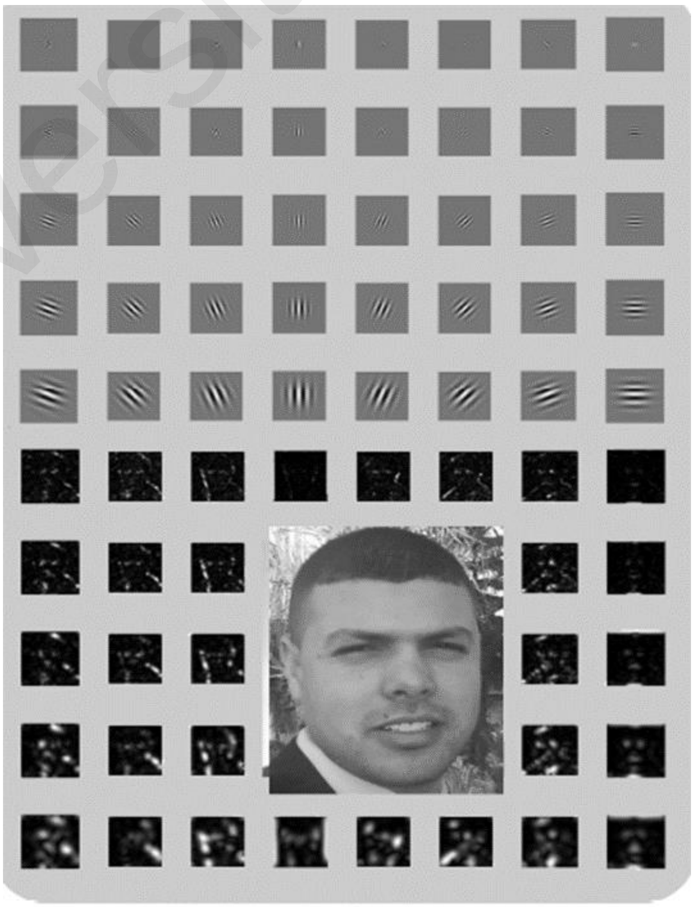

Figure 4.6: Gabor wavelets family

The dataset pictures extracted characteristics such as the mouth, nose, chin, and eyes. These feature points' Euclidean distances are computed for 2-level haar and coif and are listed in Tables 1 and 2 correspondingly.

The different distance face features that are examined are FFD1, FFD2, FFD3, and FFD4. FFD1 measures the distance between the extreme left and right corners of the eyes, while FFD2 measures the distance between the nose and the center of the eyes. The distance between the center of the eyes and the mouth is measured in FFD3. The distance between the midpoints of the eyes and the chin is computed as FFD4 in Table 4.1. The assessed feature points were put into an Artificial Neural Network to compare the two images. The distance between these feature points was calculated using Euclidean distance. The distance between pixels (a,b) and the center of the image (x,y):

$$D_E[(a,b)\ ,\ (x,y)\ ] = [(a-x)^2 + (b-y)^2\ ]^{1/2}$$

Table 4.1: Calculated the distance between the features

Face Images	FFD1	FFD2	FFD3	FFD4
IMAGE12	2.1356	0.8292	1.3503	2.1342
IMAGE13	2.5347	0.9218	1.6995	2.4290
IMAGE14	2.4498	0.8488	1.6894	2.3289
IMAGE15	2.5665	0.9321	1.6734	2.5447

Wavelet transform is used for feature extraction with a different algorithm for face recognition and detection. Wavelet transform was used with neural network for the 22D and 3d face images occlusion recognition with accepted accuracy of 95% and 80% (Yuan & Niemann, 2000). Also, Gabor Filters are used for feature extraction. The model is used for facial expression. This led to an increase in the detection rate from 86.23% to 90.3%) (Da'San et al., 2015).

The face classification happens by geometric face similarity. Thirty facial feature points have been identified. Also, they have used full convolutions of Gabor wavelets at each pixel in the image. This minimum overlap distance between facial points was selected to be 9 pixels.

4.2.3 Step 3: The third step is the training stage used MAML

The final level is model agnostic meta learning training (MAML). MAML is a meta-learning method that is frequently utilized. MAMIL determines the network's optimal beginning weights, helping it learn new jobs fast, even if just a few labeled samples are utilized in training. MAML is the algorithm used to train any model, allowing it to learn any new function fast with only a few datasets. In meta-learning, existence functions are utilized as training examples.

A model f, the input x maps to outputs a. Assumed a meta-model f parameterized by meta-parameters θ. The model may be trained with a variety of dataset sizes thanks to meta-learning. Our model-agnostic meta-learning method (MAML) optimizes for a representation that can swiftly adapt to new tasks, as shown in Figure 4.3.

The main aim is to adapt Model Agnostic Meta-Learning (MAML) $f(T, \theta)$ with weights θ to perform better on ANN using data α, Where $R_\alpha = (X_\alpha, Y_\alpha)$. So, the adaption data R_α produces adapted weights $\theta`$:

$$adapt(f, \theta, R_\alpha) \rightarrow \theta` \qquad 4.15$$

The input X_α corresponds to labels Y_α. The performance of the Model $R_u = (X_u, Y_u)$ is measured as the value of a loss function:

$$L(Y_u, f(X_u; \theta)) \qquad 4.16$$

The performance of adapted Model $R_u = (X_u, Y_u)$ is measured as the value of a loss function

$$L (Y_u , f (X_{u\,;} \text{adapt} (f , \theta , R_\alpha)) \qquad 4.17$$

To train the ANN function using the meta-learning approach, we add the parameter ϕ

$$\text{adapt} (f, \theta, R_{\alpha;} \phi) \rightarrow \theta ` \qquad 4.18$$

The meta-loss learners may be calculated as the total of all parameter losses:

$$k = \Sigma_R \; L (Y_u , f (X_{u\,;} \text{adapt} (f , \theta , R_{\alpha\,;} \phi))) \qquad 4.19$$

where R consists of series of data for adaption of the model:

$$R = \{(R_{a1}, R_{x1}) , \ldots.., (R_{aN} , R_{xN})\} \qquad 4.20$$

We used several photos as adaptation data and images as unseen data to approximate the data. The data is divided into two sets; meta-training set S_{train} and meta-validation set O_{val}.

Finally, H as loss is used to optimize the parameter ϕ of the function using gradient descent :

$$\phi = \arg\min_{\phi} H \qquad 4.21$$

To formulate the training as a meta-learning task, the weights of the model θ and the parameter of the meta-learners ϕ are calculated, so minimizing the loss of H:

$$\theta^{\wedge}\phi^{\wedge} = \arg\min_{\theta\,\phi} H \qquad 4.22$$

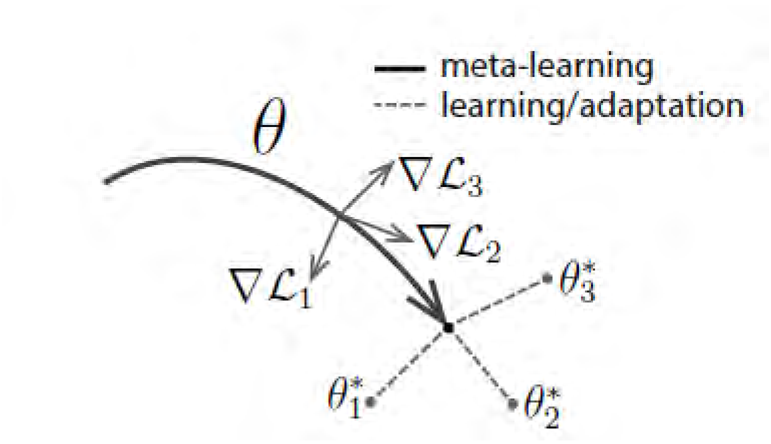

Figure 4.7: Diagram of MAML (Finn et al., 2018)

The model-agnostic meta-learning method (MAML) is depicted in Figure 4.7, and it seeks to create a representation that can quickly adapt to new tasks. The authors describe a method that uses meta-learning to learn the parameters of any standard model to prepare it for quick adaptation. To put it another way, we want to make model parameters that respond to changes in the job. Even small changes in the parameters would improve when the gradient of the loss function of any task chosen from p(T) is modified in the direction of the gradient of that loss. The objective is calculated using the new model parameters, and the meta-optimization is performed on those parameters. The objective of the method we propose is to improve the model's parameters. On a new job, a small number of gradient steps would yield the most efficient behavior. The objective is calculated using the new model parameters, and the meta-optimization is performed on those parameters.

Algorithm 1: Model-Agnostic Meta-Learning :

Require: p(T): distribution over tasks

Require: α, ß: step size hyper parameters randomly initialize θ

while not done, do

Sample batch of tasks $T_i \sim p(T)$

for all Ti do

Evaluate $\nabla_\theta L_{Ti} (f_\theta)$ concerning K examples

Compute adapted parameters with gradient descent : $\theta_i' = \theta - \alpha \nabla_\theta L_{Ti} (f_\theta)$

End for

Update $\theta \leftarrow \theta - \beta \nabla_\theta \Sigma_{Ti \sim p(T)} L_{Ti} (f_{\theta'})$

end while

Algorithm 2 Adaptation Function :

function ADAPT (f , Θ , D_a ; ϕ)

$\Theta_0 \leftarrow \Theta$

for j ϵ { 1 ….. adaptation steps } do

$L_j \leftarrow L (Y_a , f (X_a ; \Theta_{j-1}))$

$\Theta_j \leftarrow \Theta_{j-1} - \phi \nabla_{\Theta_{j-1}} L_j$

return $\Theta_{\text{adaptation steps}}$

4.2.4 Step 4: Recognition phase

An ANN is a short form of artificial neural network that shares variables from a series of conditional distributions. There can be one or many layers in an ANN(Vedel et al.).

ANNs model is made up of three layers of 100 neurons: input images, hidden (which is responsible for pattern extraction and the majority of internal processing), and output (produces and presents final network outputs).

The aggregate weight inputs activate a neuron, and the activation signal travels through a transfer function to generate a single output in each layer. The transfer functions, learning rules, and architecture all influence the neural network's overall behavior.

In the late nineteenth and early twentieth century, some basis of artificial neural networks (ANNs) were reported. This mainly consisted of interdisciplinary study, including physics, psychology, and neurophysiology. Rather than specific mathematical models of neuron activity, this early work concentrated on generic learning, vision, conditioning, and other concepts. As a result of these new developments, the field of neural networks has been reinvigorated. Many researches have been published in the last two decades, and many different ANN types of ANNs have been investigated. Neural networks have been utilized in various industries, including aerospace, automotive, banking, military, electronics, entertainment, finance, insurance, manufacturing, medical, oil and gas, speech, securities, telecommunications, transportation, and the environment. The application of ANN models in the ecological field began in the early 1990s.

1. Over the last 10–15 years, several varieties of ANNs have been developed, but two major categories may be identified depending on how the learning process is carried out:
2. i. During supervised learning, a 'teacher' informs the ANN how well it performs or the appropriate behavior during the learning phase.
3. ii. In 'unsupervised' learning, the ANN autonomously evaluates the properties of the data set and learns to reflect these aspects in its output.

A feedforward neural network is an artificial neural network in which the connections between the nodes do not create a loop. The feedforward neural network was the first and simplest artificial neural network to be developed. In this network, information only flows in one direction: forward, from the input nodes to the output nodes through the hidden nodes. There are no loops or cycles in the network.

This is seen in Figure 4.8. In a feedforward network, information always flows in one direction and never backward.The architecture of ANN of is shown in figure 4.8 which have three layers.

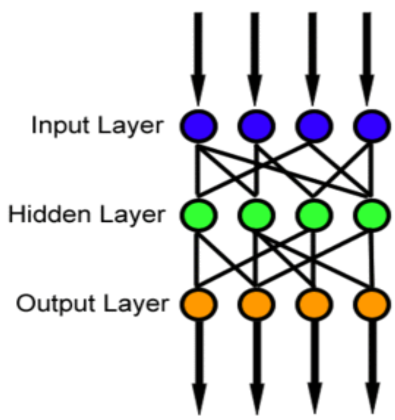

Figure 4.8: Feedforward network

4.2.4.1 Multi-layer perceptron

In this type of network, many layers of computer units are connected in a feed-forward way. Directed connections connect each neuron in one layer to the neurons in the next layer. In many applications, the units of these networks employ a sigmoid function as an activation function.

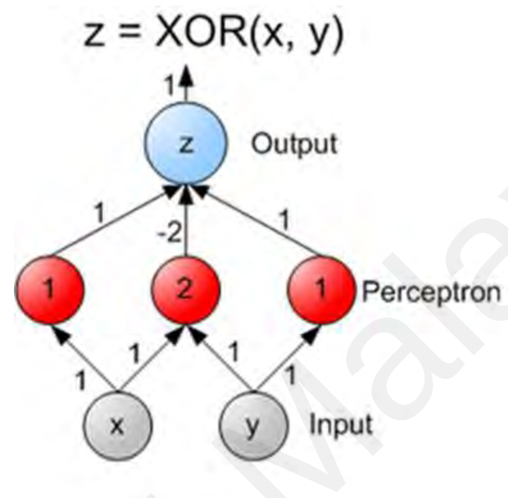

Figure 4.9: A two-layer neural network capable of calculating XOR

In general, correctly training a network is a sensitive task. This is particularly important when there are just a few training instances available. The problem is that the network overfits the training data and misses the underlying statistical process in the data.

Computational learning theory aims to train classifiers using a small amount of data. However, when it comes to neural networks, a simple heuristic known as early stopping almost always ensures that the network generalizes successfully in real-time rather than only on the training set.

Two additional typical issues with neural network techniques are their convergence speed and the risk of ending up in a local minimum of the error function. There are now practical techniques to make multi-layer perceptron neural networks the preferred tool for many machine learning applications.

A series of independent neural networks controlled by an intermediary can also be employed, analogous to how the brain works. These neurons may operate separately and complete an enormous task, with the outputs eventually being combined.

4.2.4.2 Initialize Network

Each neuron has a unique set of weights that must be balanced. There is a bias weight for each input connection and a weight for each input connection. A network is structured using layers. The input layer is a row from the training dataset. The first real layer is the hidden layer. The output layer follows, with one neuron for each class value.

It is a good idea to set the network weights to small random numbers. In this example, we will use random integers in the range of 0 to 1.

4.2.4.3 Forward Propagate

- The output of a neural network may be determined by propagating an input signal through each layer until the output layer emits its results.

- This is the technique that will be used to make predictions on new data once the network has been trained and the methodology that is used to make predictions during training that needs to be corrected.

- Forward propagation is divided into three stages:

1. Neuron Activation

2. Neuron Transfer

3. Forward Propagation

(a) Neuron Activation

The first step is to determine a single neuron's activity in response to a specific input. The input, like the hidden layer, might be a row from the training dataset. It may also be the outputs from each neuron in the hidden layer in the case of the output layer. Neuron activity is calculated using the weighted sum of the inputs.

(b) Neuron Transfer

The activation of a neuron must be transferred once it has been triggered to identify the cell's output. There are a variety of transfer functions available. The sigmoid activation function is usually used to transfer outputs, although the tan h (hyperbolic tangent) function can also be utilized. Large deep learning networks have recently been interested in the rectifier transfer function.

(c) Forward Propagation

The outputs of each neuron are listed below. The network's network has been calculated at each tier. The neurons in the following layer get all of the preceding layer's outputs as inputs. The outputs of each neuron are listed below. The network's network has been calculated at each tier. The neurons in the following layer get all of the preceding layer's outputs as inputs.

4.2.4.4 The training phase inside the neural network

In addition to the input pattern, feedforward neural networks (FFNet) use a supervised learning approach that needs the neural network to know which category the pattern belongs to. Learning can occur when a pattern is provided at the inputs. The pattern will vary as it progresses through the levels of the network until it reaches the output layer.

The output layer's units are all categorized differently. The current outputs of the network are compared to what they would have been if this pattern had been correctly classified; in this case, the unit with the correct category would have had the most excellent output value, while the output values of the other output units would have been minimal.

As a result of this comparison, all of the connection weights are slightly adjusted so that the next time this same pattern is presented at the inputs, the value of the output unit that corresponds to the correct category is slightly higher than it is now, while the output values of all the other incorrect outputs are slightly lower (differences between actual and idealized outputs are sent down from the top to the lower layers, used to alter connection weights). This is why a backpropagation network is commonly used to describe this type of neural network.

The size of the neural network, the number of patterns to be learned, the number of epochs, the minimizer's tolerance, and the computer's speed all influence the time of the learning phase.

4.2.4.5 Neural Network algorithm

The "Back Propagation Algorithm" is one of the most popular and extensively used methods for training feed-forward neural networks. In essence, it is a technique for updating synaptic weights in networks by back-propagating a gradient vector, each element of which is given as the derivative of an error measure concerning a parameter. An error signal is a signal that describes the difference between the actual network outputs and the desired outputs.

Neuron j is a neuron in the output layer. Equation 4.23 defines the error signal at the output of neuron j for the nth iteration:

$$e_j(n) = d_j - y_j(n) \quad 4.23$$

Where d_j is the desired output for neuron j, and y_j (n) is the actual output for neuron j calculated by using the current weights of the network at iteration n. There is a desired output for each input, which the network is intended to generate. Equation 4.24 gives the instant value of the error energy for neuron j:

$$\varepsilon_j(n) = ½ * e^2_j(n) \quad 4.24$$

Because the output layer's neurons are the only ones accessible, error signals for those neurons may be computed directly. Hence, the instantaneous value, $\boldsymbol{\varepsilon}$(n), of the total error energy is the sum of all ε_j (n) 's for all neurons in the output layer, as given in Equation 4.25:

$$\boldsymbol{\varepsilon}(n) = ½ \sum_{j\in Q} e_j^2(n) \quad 4.25$$

where Q is the set of all neurons in the output layer.

Assume the training set has N patterns. Equation 4.26 calculates the network's average squared energy:

$$\boldsymbol{\varepsilon}_{av} = 1/N \ \sum_{j\in Q} e_j^2(n) \quad 4.26$$

It is worth noting that the instantaneous error energy is not the same as the total error energy $\boldsymbol{\varepsilon}$(n), and therefore, the average error energy $\boldsymbol{\varepsilon}_{av}$ is a function of all the free parameters of the network. The back-propagation technique allows the user to change the network's free parameters to reduce the average error energy $\boldsymbol{\varepsilon}_{av.}$ Backpropagation algorithms come in two flavors: "sequential mode" and "batch mode."

After each training example is given in sequential mode, weight updates are done. A full presentation of the training set is referred to as an "epoch." Once all training has been provided, weight updates are done in batch mode.

Sequential mode is sometimes referred to as an online, pattern, or stochastic model. This is the most often used mode of operation, which is explained further below.

Let us start with the output equation for neuron j in Equation 4.27:

$$Y\ j\ (n) = \int\left[\sum_{i=0}^{m} w_{ji}\ (n)\ y_i\ (n)\right] \qquad 4.27$$

where m is the total number of inputs (excluding the bias) from the preceding layer to the neuron j, and f is the nonlinear activation function used in the neuron j. Here, w_{j0} equals the bias b_j applied to the neuron j, corresponding to the fixed input $y_0 = +1$.

The weight updates to be applied to the weights of the neuron j is proportional to the partial derivative of the instantaneous error energy ε(n) with respect to the corresponding weight, i.e., $\partial\varepsilon(n) / \partial w_{ji}\ (n)$, Equation 4.28 may be represented using the chain rule of calculus:

$$\frac{\partial \varepsilon(\mathrm{n})}{\partial \mathrm{wji}\ (\mathrm{n})} = \frac{\partial \varepsilon(\mathrm{n})}{\partial \mathrm{ej}\ (\mathrm{n})} \frac{\partial \mathrm{ej}(\mathrm{n})}{\partial \mathrm{yj}\ (\mathrm{n})} \frac{\partial \mathrm{yj}(\mathrm{n})}{\partial \mathrm{wji}\ (\mathrm{n})} \qquad 4.28$$

The correction $\Delta w\ ji\ (n)$ applied to $wji\ (n)$ is defined by the delta rule, given in

Equation 4.29.

$$\Delta w\ ji\ (n) = -\ \eta\ \frac{\partial \varepsilon(\mathrm{n})}{\partial \mathrm{wji}\ (\mathrm{n})} \qquad 4.29$$

In Equation 4.29, η corresponds to the learning-rate parameter of the backpropagation.

The algorithm is generally set to a pre-determined value and maintained constant during the algorithm's operation.

Feed-forward neural networks are the most common form of artificial neural networks, and they are used in a wide range of applications. Following a lecture on feed-forward neural networks, a quick introduction to artificial neural networks was given in this thesis.

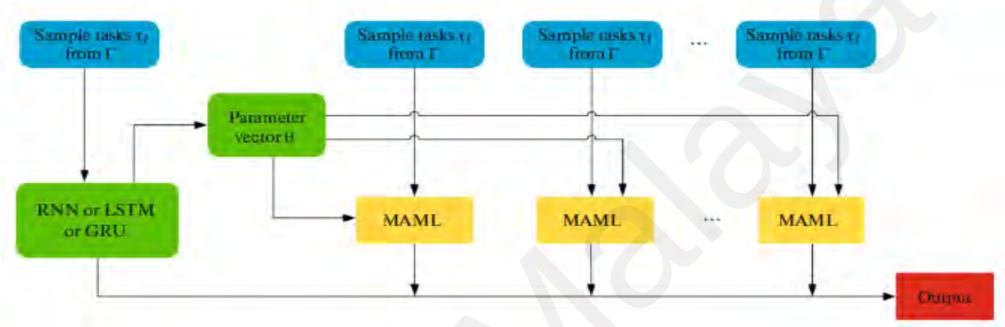

Figure 4.10: ANN with MAML

The proposed framework subcompounds start from the pre-processing phase for unify the input image ,also to have the same size and characteristics . The second phase was feature extraction to extract the feature , the third phase was MAML to training the ANN , which select the important features .The last phase is ANN phase for classify the face.

4.3 Summary

Chapter 4 presents the main framework for the proposed model. The architecture of the framework has four steps. The first step is the image pre-processing stage. The pre-processing stage has three stages: first, Grey-level features reshaped to a vertical, and histogram equalization. The second step is to select the features using wavelet transforms with three topics: feature selection definition, feature extraction local and feature extraction global explanation, and then show the wavelet transform algorithm.

The third step is the training using MAML. Finally, the fourth step is the recognition phase based on Artificial Neural Network (Vedel et al.). Also, the requirements, analyses, preliminary design, and implementation were explained.

CHAPTER5: RESULTS AND ANALYSES

5.1 Introduction

As stated in the previous chapters, the innovation of this research is to build a framework capable of detecting the human face and comparing two faces of the same person with different ages, also at the same time improve system's performance and additionally reduce the detection time. The evaluation phase presents the performance of the proposed framework based on several experiments performed. Three evaluation stages show the effectiveness of the proposed framework concerning the performance of various models incorporated in the proposed framework.

The assessment looks at how well the ANN model performs on three datasets: (1) CALFW (Cross-Age LFW), (2) Middle East Dataset (EDS), and (3) AT&T. The performance of the ANN model with adding modified MAML method is investigated in the second level of the assessment using the same three datasets.

This chapter is divided into several stages of evaluation, each with its own set of objectives. Each level is demonstrated based on the results and analysis of the suggested model's performance. The evaluation process has specific standard criteria for experimental methodologies; this section discusses the commonalities in each stage's components. All available tools, datasets, and evaluation metrics are equivalent.

All of the tests were carried out on a PC running Windows 10 Enterprise N. The machine is equipped with an 8.00 GB random access memory (RAM), a 2.90 GHz Intel(R) Core (G. Hu et al.), and an i7-7500U processor. The assessment process has specific standard requirements; this section details the commonalities in each stage's sections. General tools, datasets, and assessment measures are all comparable.

Three types of datasets are utilized to assess the framework's performance at each stage of the experiment.

CALFW (Cross-Age LFW), My Own Dataset (Middle East Dataset MEDS), and AT&T are the datasets. Each of the three datasets has its distinct feature. The datasets' descriptions are utilized in this evaluation step.

The following is how this chapter is structured: The introduction is in Section 5.1. Section 5.2 introduces the datasets CALFW, AT&T, and MEDs, Section 5.3 demonstrates the generic tool used to construct the framework, Section 5.4 explains the various evaluation matrices, and Section 5.5 demonstrates the training procedure. Finally, Section 5.6 summarizes the findings of the experiment.

5.2 Data Set

The datasets are CALFW, AT&T, and the Middle East, and each has its distinct characteristic. The datasets utilized in this evaluation step are described as follows:

5.2.1 Dataset 1: Cross-Age LFW (CALFW) Database

The Cross-Age LFW (CALFW) dataset was created by selecting 3,000 favorable face pairings with age gaps to enhance intra-class diversity in the aging process. Negative pairs of the same gender and race are also chosen to lessen the impact of attribute differences between positive and negative pairs and accomplish face verification rather than attribute categorization.

Figure 5.1: The aging process is more evident in CALFW dataset

CALFW was founded to create a real-world face recognition situation and applied to reach face verification. As a result, different pictures of the same individual have been added to this collection.

5.2.1.1 Construction Details:

The method for creating the CALFW dataset may be broken down into the following stages:

1. Collecting raw images from the Internet.

Google, Getty Images, and Bing are used to search for different face images of different ages. The images were of different celebrities of different ages.

2. When many people are in a photo, a face detector is used, and the findings are adjusted manually.

3. Trim and resize the faces that have been discovered. Using the Matlab function 'imresize', the pictures were shrunk to 250 by 250 pixels. The pictures were then saved in JPEG format.

4. The duplicated pictures have been removed. The repeated picture is when two images are identical at every pixel. So, the pictures that were repeated were removed.

5. To make sure whether the labels of the picture are correct, which means to manually check that the pictures for the same person are correctly matched with the same person and have different ages.

6. Images of famous landmarks

7. Try to figure out how old each photograph is. Positive couples will have an enormous age gap, while negative pairs will include individuals of the same gender and race.

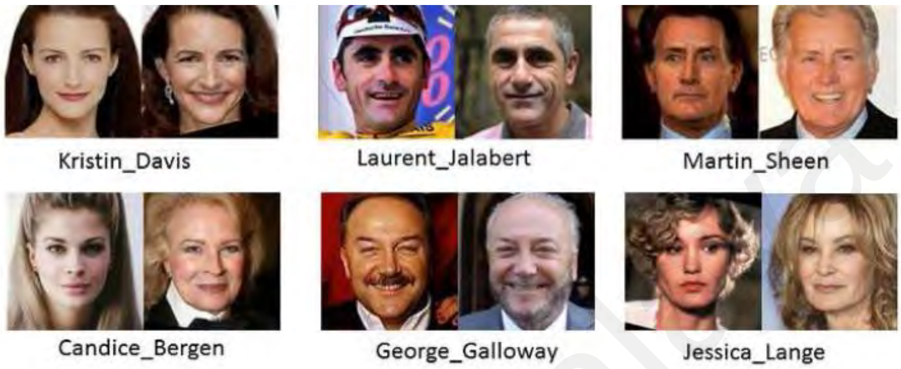

Figure 5.2: Positive Pairs in CALFW

Figure 5.3: Negative Pairs in CALFW

5.2.2 DATASET 2: AT&T Database of faces

The AT&T database has ten different photos for each of 40 different themes for 400 images. Face details (glasses / no glasses), facial emotions (open/closed eyes, smiling / not smiling), various times, and altered face lighting may all be used to create diverse themes. The photos had a distinctive grey-scale backdrop, and the individuals were in a vertical, forward-facing posture for the face. There are ten pictures for each individual in this database. The collection contains 400 example pictures with a size of 92×112 pixels and a grey level of 256.

AT&T comprises a series of facial pictures collected at the Cambridge University Engineering Department lab between April 1992 and April 1994. The database was utilized in a face recognition study conducted with the same university's Speech, Vision, and Robotics Group.

Figure 5.4: AT&T images

5.2.3 Middle East DataSet (MEDs)

- Collecting raw pictures from various individuals in the Middle East by requesting them to send me their images of various ages.

- When there are many people in a photo, a face detector is used to manually adjust the findings.

- Trim and resize the faces that have been discovered. Using the Matlab function imresize, the pictures were shrunk to 250 by 250 pixels. The pictures were then saved in PGM format.

- Images of famous landmarks

- Determine how old each image is. Positive couples will have the largest age gap, while negative pairs will include individuals of the same gender and race.

Figure 5.5: Middle East Dataset

5.3 General Tools

MATLAB software was used as a primary programming package to carry out the experiments in this chapter. The following sections provide some basic information about these programs:

5.3.1 Matlab

MATLAB (short for "matrix laboratory") is a commercial multi-paradigm programming language created by Math Works. MATLAB allows you to create user interfaces, perform algorithms, and interact with programs written in other languages. It first appeared in 1970, and the stable release is R2021a on 17 march of 2021. MATLAB users are with various specializations; engineering, science, computer science and economics. In 2020 Matlab users reached 4 million users from the whole world.

5.3.2 Photoshop

Photoshop was first released in 1988, and its term has now become a generic trademark, allowing it to be used as a verb (for example, "to photoshop an image," "photoshopping"). Photoshop supports masks, alpha compositing, and many color models, including RGB, CMYK, CIELAB, spot color, and duotone, and can edit and composite raster pictures on several layers.

5.4 Evaluation Metrics

This section outlines the assessment measures utilized in this study to assess the suggested model's performance. Performance metrics are helpful tools for determining a model's efficacy. A confusion matrix can be used to assess a type's categorization performance.

Table 5.1 shows the general form of the confusion matrix for the binary class classification issue. True Positive (TP) and True Negative (TN) denoted the number of successfully identified spam and actual samples in this table. The number of valid instances categorized as spam is known as False Positive (FP), whereas the number of spam instances classed as legitimate is known as False Negative (FN).

Table 5.1: Confusion matrix

		Class= correct	Class= Wrong
Actual Class	Class= correct	TP	FN
	Class= Wrong	FP	TN

The parameters TP, TN, FP, and FN in this table can be used to calculate standard metrics like True Positive Rate (TPR), True Negative Rate (TNR), False Positive Rate (FPR), and False Negative Rate (FNR), as demonstrated in Eqns. 5.1, 5.2, 5.3, and 5.4. TPR, also known as detection rate (Faraki et al.), sensitivity, or recall, measures a classification model's accuracy on labeled samples. The F-measure or F1-score is a composite statistic that has been frequently used to assess the effectiveness of classification systems. As demonstrated in Eq. 5.6, this measure is computed as the harmonic mean of accuracy and recall. AUC-ROC has also been used, a metric that displays TPR and FPR on a single graph to produce another robust assessment measure.

In practice, each classifier's categorized hand signs may or may not correspond to the actual sign status. As a result, four scenarios are defined:

- True Positive (Sukhija et al.): correctly classified signs.
- False Positive (FP): incorrectly classified signs.
- Ture Negative(TN): correctly misclassified signs.
- False Negative (FN): incorrectly misclassified signs.

Specificity (True Negative Rate): is the number of valid negative predictions that are divided by the total number of negatives. It is also known as a real negative rate (TNR). The highest level of specificity is 1.0, while the lowest level is 0.0. Equation 5.1 assesses the accuracy with which erroneous instances are classified:

$$\text{Specificity} = TN / (TN + FP) \qquad 5.1$$

The sensitivity (recall or actual positive rate) is the number of correct positive predictions divided by the total number of positives. It is also known as the recall rate (REC) or the actual positive rate (TPR). The highest level of sensitivity is 1.0, while the lowest level is 0.0. In Equation 5.2, the classification accuracy of actual instances is measured:

Sensitivity = TP / (TP + FN) 5.2

False Positive Rate (FPR): This_is computed by dividing the total number of negatives by the number of erroneous positive predictions. The best rate of false positives is 0.0, while the worst rate is 1.0. It is also possible to compute it as 1-specificity:

FPR = FP / (FP + TN) = 1- Specificity 5.3

Precision: is computed by dividing the number of positive predictions by the number of correct positive forecasts. It is also known as positive predictive value (PPV). The best precision is 1.0, while the poorest precision is 0.0. It describes random errors and measures the statistical variability.

Precision = TP / (TP + FP) 5.4

F1- Score (F- measure): It measures the balance among the sensitivity and precision, in which its best value is 1.

F1- Score = 2* ((precision * sensitivity)/(precision + sensitivity)) 5.5

False Negative Rate (FNR): It is the percentage of positives that offered negative results

FNR = FN/ (TP + FN)= 1 – Sensitivity 5.6

5.5 Training and testing phases (Proposed Model Evaluation ANN_MAML)

Training phase

ANN have 4 + 16 + 5 = 25 input nodes. Its goal is to find out important face features: horizontal blocks to find out mouths and eyes, square blocks to find out each of the eyes, noses, and mouths. The system uses one hidden layer with 25 nodes to represent local features that characterize faces well . Its activation function is the Tanh function with the learning rate Ɛ=0.3.

To define face component feature vectors, we detect two image regions: (a) region 1 contains eyes, which was detected by coordinates of feature points: left temple and right temple features, left outer top eyebrow and left inner top eyebrow features, right outer top eyebrow and right inner top eyebrow features, left eye bottom and right eye bottom features; (b) region 2 contains labeled face without mouth, which was detected by coordinates of feature points: left jaw and right jaw features, left outer top eyebrow and left inner top eyebrow features, right outer top eyebrow and right inner top eyebrow features, and left nose bottom, nose base, and right nose bottom. After locating image regions of the face image, region 1 and 2 are normalized in a standard size of 30 × 30 pixels. We have vectors that represent eyes image region:

$X_{eyes} = (x_1, x_2, \ldots, x_{900})^T$

and face without mouth image region:

$X_{face\_no\_mouth} = (x_1, x_2, \ldots, x_{900})^T$

These vectors are called “Geometric-component feature vectors.”

Evaluating MAML- ANN Model Using K-Fold Cross-Validation

K-Fold Cross-validation is a statistical method used to estimate the skill of machine learning models, and it works by grouping a given data sample by splitting them based on k. In this research, k=10, so 1000 rows, it was separated into 100 rows * 10, and each fold will be the test fold like the image below. The proposed model trains the 10 folds sets and is trained twice using k-fold cross-validation.

To increase the training speed and use all CPU power, the parameter n_jobs=-1.

The three tests were carried out at this stage to determine the performance of the suggested model. The following are the goals of this evaluation phase:

a. To rank the categories of features proposed in this study

b. To investigate a suitable model for face recognition

c. To identify the particular features using the proposed model.

5.5.1 Experiment and Procedure Description

The conclusion of the suggested model in this study was used to evaluate the outcomes of the experiments on the three datasets completed in this stage. Therefore, the experiments that follow could help achieve the primary goal:

- Experiment 1: The goal of this experiment is to see how fast you can recognize images with a varying number of them—the framework's dependability

- Experiment 2: The goal of this experiment is to determine the correctness of the suggested framework.

- Experiment 3: Calculate the Confusion Matrices, the error analysis and comparison with other previous work.

5.5.1.1 Experiment 1: The aim of this experiment is to evaluate the speed of the recognition framework

Different sizes of datasets have been used to evaluate our framework ANN -MAML. Figure 5.6 shows the relation between the size of the dataset and the time needed for training. It is proven that when the size of the images increases, the time of training increases. For example, when the data size was 180 images, the training time was 83.518 seconds. Also, when the data size was 220 images, the training time was

180.532 seconds—similarly, the time increased to 346.129 seconds when the data size was 440 images. The time becomes 2146.492 seconds when the size of images becomes 500.

Table 5.2: Face recognition rates at CALFW dataset

No. Images	Time
180	83.518
220	180.532
440	346.129
500	2146.492

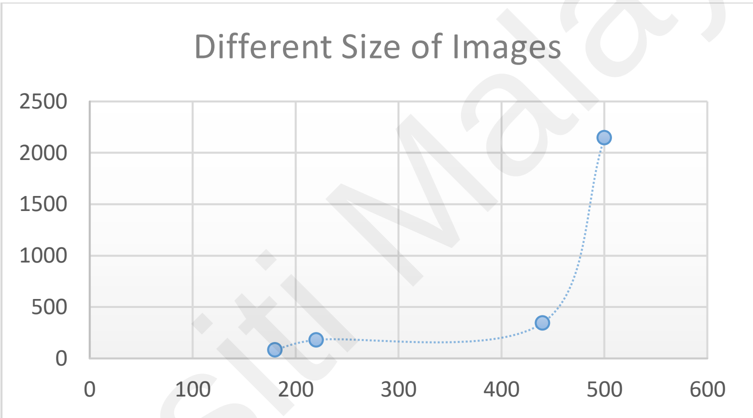

Figure 5.6: Relation between size of dataset and time need for training

5.5.1.2 Experimental 2: Recorded Accuracy and Evaluation Metric for the framework for the three datasets

The second experiment calculates the proposed framework's accuracy using three datasets CALFW, AT&T and MEDS.

Table 5.3: Clarify the Framework Accuracy for three Datasets

Algorithm	CALFW	AT&T	MEDS
Deep Metric Learning (Ni et al., 2019)	87.57%	91%	85%
AA-CNN(Y. Huang & Hu, 2020)	90%	92%	87%
KISSME(Koestinger, Hirzer, Wohlhart,	84.46%	89%	80%

Roth, & Bischof, 2012)			
Noisy Softmax(Chen, Deng, & Du, 2017)	82.52%	88%	81%
Linear Discriminant Analysis (Zangeneh et al.)	82%	95%	79%
Kernel Fischer Discriminant Analysis (KFDA)	83.40%	96%	80%
Kernel Principal Component Analysis (KPCA	30.50%	50.83%	30%
Principal Component Analysis (PCA)	71.60%	79.11%	70%
Artificial Neural Network	85%	92%	75%
Proposed framework (ANN-MAML)	90.40%	96.20%	86.95%

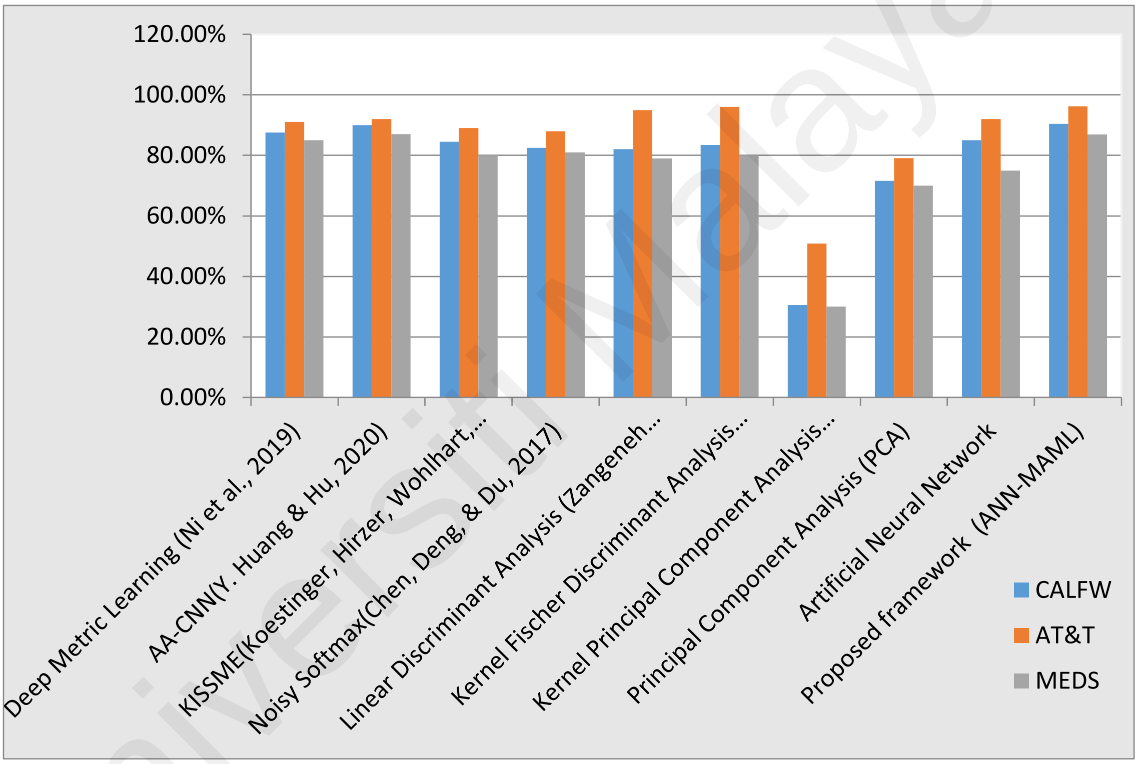

Figure 5.7: Explain the Framework Accuracy for the three datasets compared to the previous work

As indicated in the chart, there are a variety of ways to deal with facial recognition. Although the Deep Metric Learning system has an accuracy rate of 87.57% (Ni et al., 2019), the performance needs to be improved by adding more features and images. Age Adversarial Convolutional Neural Network with parallel network architecture (AA-CNN) increases the performance to 90 % (Huang & Hu, 2020), but they labeled the images with the age that gives ease to the framework.

KISSME framework used for face recognition is a combination between Mahalanobis distance functions and SVM; the system gave 84.46% (Koestinger et al., 2012). Also, a deep convolutional neural network was supported by softmax, used for face recognition, with a system accuracy of 82.52% (Chen et al., 2017). However, all these frameworks use LFW, LFW datasets that do not include images for the same person of different ages, so they did not detect aging. Another application used Linear Discriminant Analysis (Zangeneh et al.) gives 82%. Also, Kernel Fischer Discriminant Analysis (KFDA) gave 83.40%, Kernel Principal Component Analysis (KPCA) gave 30.50%, Principal Component Analysis (PCA) gave 71.60%, Artificial Neural Network gave 85%. The proposed framework ANN –MAML accuracy is 90.4 % using the dataset CALFW; this dataset includes pictures of the same person at different ages (more than ten years), and the framework can learn wrinkles.

AT&T is the second dataset used for evaluating the proposed framework. A&T is a standard dataset from the internet that has images for the same person in a different position but the same age. The results were obtained from related studies. Deep Metric Learning (Ni et al., 2019) give accuracy of 91%, AA-CNN (Huang & Hu, 2020) gives 92%, KISSME (Koestinger et al., 2012) gives 89%, Noisy Softmax (Chen, Deng, & Du, 2017) gives 88%, Linear Discriminant Analysis (Zangeneh et al.) gives 95%, Kernel Fischer Discriminant Analysis (KFDA) gives 96%, Kernel Principal Component Analysis (KPCA) gives 50.83%, Principal Component Analysis (PCA) gives 79.11%, Artificial Neural Network gives 92, and finally, the Proposed framework (ANN-MAML) gives 96.20%.

The results' accuracy with the MEDS dataset is as follows: Deep Metric Learning (Ni et al., 2019) gives 85%, AA-CNN (Huang & Hu, 2020) gives 87%, KISSME (Koestinger, al at, 2012) gives 80%, Noisy Softmax (Chen, Deng, & Du, 2017) gives

81%, Linear Discriminant Analysis (Zangeneh et al.) is 79%, Kernel Fischer Discriminant Analysis (KFDA) is 80%, Kernel Principal Component Analysis (KPCA) is 30%, Principal Component Analysis (PCA) is 70%, Artificial Neural Network is 75%, and the Proposed framework (ANN-MAML) gives 86.95%.

The expreince shown that the model based AA-CNN(Y. Huang & Hu, 2020) give higher aaccuracy from the proposed framework only with MEDS dataset , the AA—CNN give 87% and for the proposed framework 86.5%.

The proposed model show superiority over others techniques as shown at table5.3 and figure 5.7 which explain the framework accuracy for the three datasets compared to the previous work .The accuracy for the propoosed model was 90.40% with the dataset CALFW , the accuracy was 96.20% with the dataset AT&T , and the accuracy 86.5% with the dataset MEDS.

5.5.1.3 Experimental 3: Calculate the Confusion Matrices, the error analysis and comparison with other previous work

The third experiment calculated the confusion matrices for the three datasets and the error analysis for the experiments.

(a) ***Results and Discussion for Framework First Dataset (CROSS-AGE LFW (CALFW) DATABASE):***

- Confusion Matrices for CALFW: CALFW dataset has 12000 images, with 600 pairs of images. Table 5.4 presents the confusion matrix for CALFW.

Table 5.4: Confusion Matrices for CALFW

Proposed model / Actual	Same person	Not the same person
Same person	500	70
Not the same person	35	490

- Specificity (True Negative Rate) = TN / (TN + FP) =490 / (490+35)=0.9333.

- Sensitivity (recall or true positive rate) = TP / (TP + FN) = 500/(500+70)=0.87719.

- False Positive Rate (FPR) = FP / (FP + TN) == 35/(35+490)= 0.6666.

- Precision = TP / (TP + FP) = 500/(500+35) = 0.9345.

- F1- Score (F- measure) = 2* ((precision * sensitivity)/(precision + sensitivity)) = 2*((.81973)/(1.81169)) = .45246.

- Fales Negative Rate (FNR) = FN/ (TP + FN) = 70/(500+70)= 0.12280.

- Accuracy = (TP+TN) / (TP+TN+FP+FN) = (500+490)/ (500+490+70+35) = 0.9041.

(b) ***Experimental Results and Discussion for Framework Second Dataset (AT&T DATASET)***

• Confusion Matrices for AT&T: AT&T database contains 40 faces; this means 40 individuals, ten images of each face. So the total number of images for this dataset is 400 images. Table 5.5 presents the confusion matrix for AT&T.

Table 5.5: Confusion Matrices for AT&T

Actual / Proposed model	Same person	Not the same person
Same person	TP = 76	FN = 3
Not the same person	FP = 3	TN = 76

• Specificity (True Negative Rate) = TN / (TN + FP) = 76 / (76+3) = 0.9620

• Sensitivity(recall or true positive rate) = TP / (TP + FN) = 76 / (76+3) =0.9620

- False Positive Rate (FPR) = FP / (FP + TN) = 3 / (3+76) = 0.03797

- Precision = TP / (TP + FP) = 76 / (76+3) = 0.9620

- F1- Score (F- measure) = 2* ((precision * sensitivity)/(precision + sensitivity))= 2

- Fales Negative Rate (FNR) = FN/ (TP + FN)= 1 – Sensitivity= 0.03797

- Accuracy = (TP+TN) / (TP+TN+FP+FN) = 152/158 = 0.9620

(c) Experimental Results and Discussion for Framework Third Dataset (Middle East Data (MEDS))

MEDS database contains 23 faces, this means 23 people, and every face has 2 images. So the total number of images for this dataset is 46 images. Table 5.6 presents the confusion matrix for MEDS.

Table 5.6: Face recognition rates at MEDS dataset

Actual \ Proposed model	Same person	Not the same person
Same person	TP = 30	FN= 6
Not the same person	FP= 3	TN = 30

- Specificity (True Negative Rate) = TN / (TN + FP) = 30 / (30+3) = 30 / 33 = 0.9090

- Sensitivity(recall or true positive rate) = TP / (TP + FN) = 30 /(30+6) = 30/36=0.8333

- False Positive Rate (FPR) = FP / (FP + TN) =3/(33)= 1- Specificity= 1- 0.9090

- Precision = TP / (TP + FP) = 30 / (30+3)=0.83

- F1- Score (F- measure) = 2* ((precision * sensitivity)/(precision + sensitivity))= 2*(0.83* 0.83)/(0.83+ 0.83)= 2

- Fales Negative Rate (FNR) = FN/ (TP + FN)= 1 – Sensitivity= 1-0.86= 0.14.
- Accuracy = (TP+TN) / (TP+TN+FP+FN) = (30+30) / 69 = 0.8695 = 86.95%.

5.6 Discussion

This section discusses the result and why this result comes out as observed. From the result section, there are three datasets, CALFW, AT&T and MEDS, with different characteristics.

5.6.1 Dataset 1: Cross-Age LFW (CALFW) Database

The dataset Cross-Age LFW has the prefix ”Cross-Age” which suggests that the age gap of the same individual has been considered a crucial intra-class variation that better simulates real-world face verification situation. To increase intra-class variability, CALFW emphasizes the age gap. Additionally, negative pairings are chosen to eliminate gender or racial differences. The substantial intra-class variation and the modest inter-class variance are both taken into account by CALFW at the same time. Each of the 4,025 people in the CALFW dataset has two, three, or four pictures.

When it comes to creating positive pairings, there are two principles to follow. To begin, the pairs comprise as many individuals as possible to replicate the variation in real-world face verification; thus, if the fold has more than 300 people, one positive pair is chosen for everyone. Second, to account for the age disparity in our sample. For each individual, the positive pairings with the most significant age gap were chosen. To prevent the attribute difference between positive (same gender and race) and negative (random race and gender) pairs, each person's race and gender was labeled manually in CALFW; then negative pairs were picked with individuals of the same gender and race at random.

For this dataset images description, the proposed model accuracy is 90.40 %. Therefore, it is accepted accuracy compared with other models; it has high accuracy.

The neutral of the images even the positive or negative gap effect on the result. Increasing the gap between the images lead to a decrease the accuracy.

Also, the number of images increases the processing time; as the number of images increases, the processing time increases. As shown in Table 5.8, the CALFW dataset has an average gap for positive pairs of 16.61±10.78, and 16.14±11.88 for the negative pair.

Table 5.7: Age gap average value of positive pairs and negative pairs in CALFW

Pairs	CALFW
Positive Pairs	16.61±10.78
Negative Pairs	16.14±11.88

5.6.2 DATASET 2: AT&T Database of faces

The images were taken at various times on the AT&T dataset by adjusting the lighting, the detail of the face (no glasses/glasses), and facial expressions (smiling or not smiling closed or open eyes). The photographs were shot against a black background and in a frontal, upright stance (with tolerance for some side movement).

The proposed method has 96.20% accuracy, but due to the nature of the images in the AT&T dataset, there is no gap between the images of the same individual. That means the individual has different images but almost the same age (almost two years difference as the dataset images were collected between April 1992 and April 1994). This is the reason of having higher accuracy than other techniques.

5.6.3 MIDDLE EAST DATASET (MEDS)

MEDs dataset is collected between 2008 to 2021 with a gap of more than 16 years for some faces. Also, the dataset has images of faces younger than 15 years, as the faces under 15 years are not stable, and the facial feature changes. For example, Figure 5.10 presents a face for a person when he was 12 years old and then when he became 28 years old.

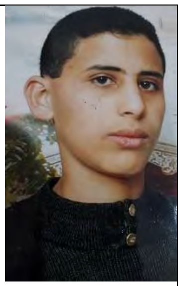	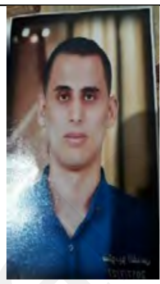
12 years old	28 years old

Figure 5.8:Example for the dataset for face 12 years old and 28 years old

According to the nature of the dataset, images that contain faces less than 15 years and gaps of more than 16 years lead to reduced accuracy of the proposed model.

The proposed framework used three reprocessing techniques, which led to improving the image and determining the boundaries of the face; then, the framework used wavelet transform as a feature selection, resulting in a selection of the most effective facial features. A modified MAML with ANN was used for training the framework.

5.7 Summary

Chapter 5 presents the results, evaluation, and discussion of why these results were obtained. Also, this chapter shows the results for the three datasets: CALFW, AT&T

and the Middle east dataset. The third section also demonstrates the general tools used to build the proposed model such as Matlab and Photoshop.

Moreover, the experimental results and discussion for the proposed framework with the three datasets are presented in this chapter. Different methods used or the evaluation calculated the accuracy, Confusion Matrices and Error Analysis. Finally, the findings of the proposed framework were compared to those of prior research.

CHAPTER 6: CONCLUSION AND FUTURE RESEARCH WORK

6.1 Introduction

Chapter 6 describes the summary of this research by displaying the achievements and presents the limitations. This chapter also discusses the recommendations for future work to enhance the proposed model presented in this thesis.

6.2 The Achievements of the proposed work

The proposed model achieved its goals by conforming to the research objectives and the research questions mentioned in Chapter 1. A new model based on neural network technique with modified MAML technique was proposed for face aging recognition.

The first objective of this thesis was to analyze and identify the limitations of existing frameworks and techniques used in recognition. This objective was achieved in Chapter 2. In Chapter 2, three taxonomies based on different databases were discussed, such as neural network techniques, Meta-learning techniques and face aging recognition. The taxonomies give a review of relevant research that has already been done. This classification would help researchers acquire more information in detail in the domain of face recognition. Chapter 2 presented the previous related works, the different image recognition techniques, and the current approach for face recognition. In this chapter, the problems in Face Recognition were presented. Also, it shows the different artificial neural networks models that are used for face recognition. Then it presents the usage of Model Agnostic Meta-Learning (MAML) methods in solving computer sciences problems. Finally, it reviewed the current systems for aging process techniques.

To achieve the second objective, the design and development model for recognizing the face using enhancement Artificial Neural network. The proposed research introduced a novel enhancemed ANN with modified MAML, presented in Chapter 3. Chapter 3 shows the methodology used for achieving this objective and explains the phases in detail. The first phase is the requirements phase, then the development phase, presenting the hardware and software needed to set up this model. After that, the analyses phase and then the last phase is the evaluation phase.

After that, Chapter 4 shows the technical framework design and implementation. The framework ANN-MAML contains 3 phases: pre-processing, feature extraction, and recognition.

The third objective is to evaluate the face recognition model using the previous studies' approaches. Chapter 5 presents the evaluation methods for the model. The model was evaluated by three datasets: CALFW, AT&T, and the Middle east. The model is implemented by MATLAB. For each dataset, the accuracy has been calculated, the Confusion Matrices have been analyzed, then the Error Analysis has been extracted. In the end, a comparison has been made between the model and previous works.

6.3 Contribution of the proposed model

This thesis contributed to face aging recognition. The framework ANN-MAML is a model based on an Artificial neural network combined with modified MAML. The first step in the framework is image pre-processing. The pre-processing phase has three stages: grey-level features, reshaped to a vertical, and histogram equalization. The critical pre-processing phase is used for contrast enhancement in the images. The second step is modified Model Agnostic Meta-Learning (MAML). MAML is considered one of the most popular algorithms at meta-learning. A modified MAML helps choose the ideal initial weights for the parameter ANN network. Therefore, the ANN network technique can learn better with the new tasks, even if the training is with just a few images.

One of the neural networks approaches is the artificial neural network (Vedel et al.). To categorize the faces, ANN was employed as a classifier. The suggested framework has been trained using conventional ANN parameters and meta-learning to help the model adapt quicker. The suggested framework has been trained using conventional ANN parameters and meta-learning to help the model adapt quicker.

The tests were carried out with the three datasets, as indicated in Chapter 5: Evaluation Results and Discussion. The model's performance was assessed using the CALFW datasets, and it was found to have a 98% accuracy rate, indicating that the framework outperformed previous state-of-the-art models.

On the other side, the model's performance was assessed using AT&T datasets, and it was shown to have a 98.8% accuracy rate. The model's performance was also examined using the MEDS datasets, and it was shown to have a 92% accuracy rate.

Based on this, the suggested model outperformed prior models in terms of results. The accomplishment of acceptable outcomes was the most significant contribution of the suggested model.

6.4 Limitation of the proposed model

The accomplishments of this research, which is focused on refining a model for face aging recognition, were shown in the preceding part.

Throughout the reasons behind this research, several limitations were encountered even though the proposed model in this thesis worked well for the different situations. These limitations are discussed as follows for future development:

6.4.1 Input Images

The input images should undergo pre-processing before entering the model. This process is as the following:

- Trim is one of the pre-processing functions. Trim is used to crop the images to concentrate on the face.

- Resize is the second pre-process function. Using the Matlab function imresize, the pictures were shrunk to 250 by 250 pixels.

- The grey-level feature is a greyscale image; in the grey-level, the pixel can be valued between 0 and 255.

- The images were saved in PGM format.

6.4.2 Performance evaluation

Even though the evaluation of the proposed model has been presented over the three datasets, the model's performance can still be improved to achieve better accuracy and decrease the false recognition rate.

6.4.3 Feature extraction time

Gabor wavelet is a function used for feature extraction. This feature is used to improve ANN recognition. This method will take some time to extract features of all of the picture datasets. In the future, the model can improve the complexity of the feature extraction stage.

6.4.4 The Running time

Graphics Processing Unit (GPU) has been used in this study to decrease the running time. GPUs are effective in image processing because they can process a large number of pictures in parallel. The GPU can assist in speeding up the model's training process.

6.5 Recommendations for future work

In this section, a list of suggestions for the future is given as there is scope for improving the proposed model. These are discussed as follows:

1. Real-Time application

Subsequently, the focus is to build a model that can be used for real-time applications. So, to reach this task, cooperation between the software and hardware is required.

2. Real-Time images (Videos)

Afterward, the datasets used to evaluate the model can change to use it in real-time and use input images by a camera, either still images or videos. In this thesis, the model was trained and evaluated with three datasets with different characteristics as the gap between the images, but all three datasets are still images.

3. Feature features

The performance accuracy of feature extraction on the model needs to be improved to provide a more efficient model. This, in turn, will improve the performance of the training and the classification model proposed in this research.

4. Real-time recognition system

The main future work of this research is to develop a framework for the recognition system in real-time with the webcam so that it can be applied at airports, police stations, and banks in real-time.

5. Increase the accuracy

Also, one of the important future works is to increase the framework accuracy. The higher accuracy of the framework, the more the frameworks are guaranteed.

6. Mobile aging system

Mobile applications have become one of the most powerful technologies in the world that make the framework more user-friendly and accessible. So, one of the future works is to build a mobile application for the face aging framework system.

7. Huge database up to 1M

The more images that can train the framework, the better the training process becomes. So in the future, collecting 1M images to train and test the framework is recommended.

8. Increase the number of the collected images

Another proposed work in the future can be to collect numerous images for the learning and training of the neural network.

9. Image type

As stated in the previous sections, the input image should be (PGM), while future studies can improve the model to accept different image formats.

REFERENCES

Abate, A. F., Nappi, M., Riccio, D., & Sabatino, G. (2007). 2D and 3D face recognition: A survey. *Pattern recognition letters*, *28*(14), 1885-1906.

ALDayel, A., & Magdy, W. (2021). Stance detection on social media: State of the art and trends. *Information Processing & Management*, *58*(4), 102597.

Alet, F., Lozano-Pérez, T., & Kaelbling, L. P. (2018). Modular meta-learning. Conference on Robot Learning, pp. 856-868.

Alqudah, A., & Al-Zoubi, H. (2015). Efficient k-class approach for face recognition. *Computers & Electrical Engineering*, *45*, 260-273.

Antipov, G., Baccouche, M., Berrani, S. A., & Dugelay, J. L. (2016, June 26 2016-July 1 2016). Apparent Age Estimation from Face Images Combining General and Children-Specialized Deep Learning Models. 2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 96-104.

Antipov, G., Baccouche, M., & Dugelay, J.-L. (2017a). Face aging with conditional generative adversarial networks. Image Processing (ICIP), 2017 IEEE International Conference on,pp96.

Azmeen, J., & Borah, D. J. (2021). Face Recognition Techniques, Challenges: A Review. *Soft Computing for Intelligent Systems*, 369-374.

Ben Fredj, H., Bouguezzi, S., & Souani, C. (2021). Face recognition in unconstrained environment with CNN. *The Visual Computer*, *37*(2), 217-226.

Bendjillali, R. I., Beladgham, M., Merit, K., & Taleb-Ahmed, A. (2020). Illumination-robust face recognition based on deep convolutional neural networks architectures. *Indonesian Journal of Electrical Engineering and Computer Science*, *18*(2), 1015-1027.

Bi, Y., Lv, M., Wei, Y., Guan, N., & Yi, W. (2016). Multi-feature fusion for thermal face recognition. *Infrared Physics & Technology*, *77*, 366-374.

Biswas, S., & Sil, J. (2017). An efficient face recognition method using contourlet and curvelet transform. *Journal of King Saud University-Computer and Information Sciences* , pp*718-729*.

Bouguila, J., & Khochtali, H. (2020). Facial plastic surgery and face recognition algorithms: interaction and challenges. A scoping review and future directions. *Journal of stomatology, oral and maxillofacial surgery*, *121*(6), 696-703.

Chu, W.-S., Chen, J.-C., & Lien, J.-J. J. (2011). Kernel discriminant transformation for image set-based face recognition. *Pattern Recognition*, *44*(8), 1567-1580.

Da'San, M., Alqudah, A., & Debeir, O. (2015). Face detection using Viola and Jones method and neural networks. 2015 International Conference on Information and Communication Technology Research (ICTRC),pp. 40-43.

Dehghan, A., Masood, S. Z., Shu, G., & Ortiz, E. (2017). View Independent Vehicle Make, Model and Color Recognition Using Convolutional Neural Network. *arXiv preprint arXiv:1702.01721*.

Ding, C., & Tao, D. (2017). Trunk-branch ensemble convolutional neural networks for video-based face recognition. *IEEE transactions on pattern analysis and machine intelligence*, *40*(4), 1002-1014.

Diyasa, G., Fauzi, A., Idhom, M., & Setiawan, A. (2021). Multi-face Recognition for the Detection of Prisoners in Jail using a Modified Cascade Classifier and CNN. Journal of Physics: Conference Series, Vol. 1844, No. 1, p. 012005

Dong, N., & Xing, E. P. (2018). Few-Shot Semantic Segmentation with Prototype Learning. BMVC, InBMVC 2018 Sep (Vol. 3, No. 4).

Elmahmudi, A., & Ugail, H. (2019). *The Effect of FaceAgeing in Face Recognition* (2516-2314).

Faraki, M., Yu, X., Tsai, Y.-H., Suh, Y., & Chandraker, M. (2021). Cross-Domain Similarity Learning for Face Recognition in Unseen Domains. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, p. 15292-15301. 2021.

Finn, C., Xu, K., & Levine, S. (2018). Probabilistic model-agnostic meta-learning. *Advances in neural information processing systems*, *31*.

Fu, Y., Guo, G., & Huang, T. S. (2010). Age synthesis and estimation via faces: A survey. *IEEE transactions on pattern analysis and machine intelligence*, *32*(11), 1955-1976.

Gaidhane, V. H., Hote, Y. V., & Singh, V. (2014). An efficient approach for face recognition based on common eigenvalues. *Pattern Recognition*, *47*(5), 1869-1879.

Gao, Q., Hao, X., Zhao, Q., Shen, W., & Ma, J. (2013). Feature extraction using two-dimensional neighborhood margin and variation embedding. *Computer Vision and Image Understanding*, *117*(5), 525-531.

Gelb, A., & Metz, A. D. (2018). *Identification Revolution: Can Digital ID be Harnessed for Development?* Brookings Institution Press.

Geng, X., Zhou, Z.-H., Zhang, Y., Li, G., & Dai, H. (2006). Learning from facial aging patterns for automatic age estimation. Proceedings of the 14th ACM international conference on Multimedia,

Gokturk, S. B., Anguelov, D., Vanhoucke, V. O., Lee, K.-c., Vu, D. T., Yang, D., Shah, M., & Khan, A. (2016). System and method for providing objectified image renderings using recognition information from images. In: Google Patents.

Gooya, E. S., Al Falou, A., & Kaddah, W. (2020). Robust and discriminating face recognition system based on a neural network and correlation techniques. 2020

Tenth International Conference on Image Processing Theory, Tools and Applications (IPTA),

Gordon, J., Bronskill, J., Bauer, M., Nowozin, S., & Turner, R. E. (2018). Meta-learning probabilistic inference for prediction. *arXiv preprint arXiv:1805.09921*.

Gui, L.-Y., Wang, Y.-X., Ramanan, D., & Moura, J. M. (2018). Few-shot human motion prediction via meta-learning. Proceedings of the European Conference on Computer Vision (ECCV),

Hu, C., Ye, M., Du, Y., & Lu, X. (2014). Vector projection for face recognition. *Computers & Electrical Engineering*, *40*(8), 51-65.

Hu, G., Yang, Y., Yi, D., Kittler, J., Christmas, W., Li, S. Z., & Hospedales, T. (2015). When face recognition meets with deep learning: an evaluation of convolutional neural networks for face recognition. Proceedings of the IEEE International Conference on Computer Vision Workshops,

Huang, B., Chen, W., Wu, X., Lin, C.-L., & Suganthan, P. N. (2018). High-quality face image generated with conditional boundary equilibrium generative adversarial networks. *Pattern Recognition Letters*, *111*, 72-79.

Huang, Z., Zhang, J., & Shan, H. (2021). When age-invariant face recognition meets face age synthesis: A multi-task learning framework. *arXiv preprint arXiv:2103.01520*.

Iqbal, M. T. B., Shoyaib, M., Ryu, B., Abdullah-Al-Wadud, M., & Chae, O. (2017). Directional Age-Primitive Pattern (DAPP) for Human Age Group Recognition and Age Estimation. *IEEE Transactions on Information Forensics and Security*.

Jana, R., & Basu, A. (2017). Automatic age estimation from face image. Innovative Mechanisms for Industry Applications (ICIMIA), 2017 International Conference on,

Kang, B., Liu, Z., Wang, X., Yu, F., Feng, J., & Darrell, T. (2019). Few-shot object detection via feature reweighting. Proceedings of the IEEE/CVF International Conference on Computer Vision,

Koch, G., Zemel, R., & Salakhutdinov, R. (2015). Siamese neural networks for one-shot image recognition. ICML deep learning workshop,

Kudratilloev, N., & Akhmedov, B. (2021). Application of communication-cluster technologies in pedagogical institutions: interactive methods of processing graphic data. *Scientific Progress*, *1*(5).

Lang, P., Steines, D., Liew, S.-W., & Vargas-Voracek, R. (2015). Methods for the compensation of imaging technique in the processing of radiographic images. In: Google Patents.

Lee, Y., & Choi, S. (2018). Gradient-based meta-learning with learned layerwise metric and subspace. International Conference on Machine Learning, Li, Z., Gong, D., Li, X., & Tao, D. (2016). Aging face recognition: A hierarchical learning model based on local patterns selection. *IEEE Transactions on Image Processing*, *25*(5), 2146-2154.

Liu, C., Cao, Y., Luo, Y., Chen, G., Vokkarane, V., & Ma, Y. (2016). Deepfood: Deep learning-based food image recognition for computer-aided dietary assessment. International Conference on Smart Homes and Health Telematics,

Liu, H.-D., Yang, M., Gao, Y., Yin, Y., & Chen, L. (2014). Bilinear discriminative dictionary learning for face recognition. *Pattern Recognition*, *47*(5), 1835-1845.

Liu, T., Mi, J.-X., Liu, Y., & Li, C. (2016). Robust face recognition via sparse boosting representation. *Neurocomputing*, *214*, 944-957.

Liu, X., Jin, L., Han, X., Lu, J., You, J., & Kong, L. (2021). Identity-aware facial expression recognition in compressed video. 2020 25th International Conference on Pattern Recognition (ICPR),

Merler, M., Ratha, N., Feris, R. S., & Smith, J. R. (2019). Diversity in faces. *arXiv preprint arXiv:1901.10436*.

Mi, J.-X., Huang, D.-S., Wang, B., & Zhu, X. (2013). The nearest-farthest subspace classification for face recognition. *Neurocomputing*, *113*, 241-250.

Mukhaiyar, R., & Safitri, R. (2019). Implementation of artificial neural network: Back propagation method on face recognition system. 2019 16th International Conference on Quality in Research (QIR): International Symposium on Electrical and Computer Engineering,

Ni, T., Gu, X., Zhang, C., Wang, W., & Fan, Y. (2019). Multi-Task Deep Metric Learning with Boundary Discriminative Information for Cross-Age Face Verification. *Journal of Grid Computing*, 1-14.

Oh, B.-S., Toh, K.-A., Teoh, A. B. J., & Lin, Z. (2018). An analytic Gabor feedforward network for single-sample and pose-invariant face recognition. *IEEE Transactions on Image Processing*, *27*(6), 2791-2805.

Oloyede, M. O., Hancke, G. P., & Myburgh, H. C. (2020). A review on face recognition systems: recent approaches and challenges. *Multimedia Tools and Applications*, *79*(37), 27891-27922.

Pandya, J. M., Rathod, D., & Jadav, J. J. (2013). A survey of face recognition approach. *International Journal of Engineering Research and Applications (IJERA)*, *3*(1), 632-635.

Panjaitan, C., Silaban, A., Napitupulu, M., & Simatupang, J. W. (2018). Comparison K-Nearest Neighbors (K-NN) and Artificial Neural Network (ANN) in Real Time Entrants Recognition. 2018 International Seminar on Research of Information Technology and Intelligent Systems (ISRITI), Pawar, A. (2020). Modified Model-Agnostic Meta-Learning. 2020 IEEE International Conference on Machine Learning and Applied Network Technologies (ICMLANT),

Peña, A., Serna, I., Morales, A., Fierrez, J., & Lapedriza, A. (2020). Facial Expressions as a Vulnerability in Face Recognition. *arXiv preprint arXiv:2011.08809*.

Peng, S., & Zheng, K. (2021). Application of Meta Learning in Face Recognition. Journal of Physics: Conference Series,

Pujol, F. A., & García, J. C. (2012). Computing the Principal Local Binary Patterns for face recognition using data mining tools. *Expert Systems with Applications*, *39*(8), 7165-7172.

Rakelly, K., Shelhamer, E., Darrell, T., Efros, A. A., & Levine, S. (2018). Few-shot segmentation propagation with guided networks. *arXiv preprint arXiv:1806.07373*.

Ravi, S., & Larochelle, H. (2016). Optimization as a model for few-shot learning.

Ryan, Ø. (2019). Using Tensor Products to Apply Wavelets to Images. In *Linear Algebra, Signal Processing, and Wavelets-A Unified Approach* (pp. 317-342). Springer.

Sabharwal, T., & Gupta, R. (2019). Human identification after plastic surgery using region based score level fusion of local facial features. *Journal of Information Security and Applications*, *48*, 102373.

Sable, A. H. (2021). Customized Adaptive Gradient and Orientation histogram for faces altered by Face surgery. 2021 7th International Conference on Electrical Energy Systems (ICEES),

Saeed, U., Masood, K., & Dawood, H. (2021). Illumination normalization techniques for makeup-invariant face recognition. *Computers & Electrical Engineering*, *89*, 106921.

Salunke, V. V., & Patil, C. (2017). A new approach for automatic face emotion recognition and classification based on deep networks. 2017 International Conference on Computing, Communication, Control and Automation (ICCUBEA),

Schwartz, W. R., Guo, H., & Davis, L. S. (2010). A robust and scalable approach to face identification. European Conference on Computer Vision, pp. 476-489.

Shaban, A., Bansal, S., Liu, Z., Essa, I., & Boots, B. (2017). One-shot learning for semantic segmentation. *arXiv preprint arXiv:1709.03410*.

Singh, R., & Om, H. (2017). Newborn face recognition using deep convolutional neural network. *Multimedia Tools and Applications*, 1-11.

Snell, J., Swersky, K., & Zemel, R. S. (2017). Prototypical networks for few-shot learning. *arXiv preprint arXiv:1703.05175*.

Sukhija, P., Behal, S., & Singh, P. (2016). Face recognition system using genetic algorithm. *Procedia Computer Science*, *85*, 410-417.

Tabassum, F., Islam, M. I., Khan, R. T., & Amin, M. (2020). Human face recognition with combination of DWT and machine learning. *Journal of King Saud University-Computer and Information Sciences*.

Taigman, Y., Yang, M., Ranzato, M. A., & Wolf, L. (2014). Deepface: Closing the gap to human-level performance in face verification. Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1701-1708.

Vedel, A., Wellnitz, K. B., Ludeke, S., Soto, C. J., John, O. P., & Andersen, S. C. (2020). Development and Validation of the Danish Big Five Inventory-2. *European Journal of Psychological Assessment*.

Wang, D., Lu, H., & Yang, M.-H. (2015). Kernel collaborative face recognition. *Pattern Recognition*, *48*(10), 3025-3037.

Wang, D., Otto, C., & Jain, A. K. (2017). Face search at scale. *IEEE transactions on pattern analysis and machine intelligence*, *39*(6), 1122-1136.

Wang, J., Liu, Y., Hu, Y., Shi, H., & Mei, T. (2021). Facex-zoo: A pytorch toolbox for face recognition. *arXiv preprint arXiv:2101.04407*.

Wang, Q., & Guo, G. (2019). Benchmarking deep learning techniques for face recognition. *Journal of Visual Communication and Image Representation*, *65*, 102663.

Wang, W., Cui, Z., Yan, Y., Feng, J., Yan, S., Shu, X., & Sebe, N. (2016). Recurrent face aging. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, p. 2378-2386.

Wang, X., Zhou, Y., Kong, D., Currey, J., Li, D., & Zhou, J. (2017). Unleash the black magic in age: a multi-task deep neural network approach for cross-age face verification. Automatic Face & Gesture Recognition (FG 2017), 2017 12th IEEE International Conference on, p. 596-603

Xie, W., & Zisserman, A. (2018). Multicolumn networks for face recognition. *arXiv preprint arXiv:1807.09192*.

Xu, Y., Zhang, X., & Gai, H. (2011). Quantum neural networks for face recognition classifier. *Procedia Engineering*, *15*, 1319-1323.

Xu, Z., Li, S., & Deng, W. (2015). Learning temporal features using LSTM-CNN architecture for face anti-spoofing. 2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR),

Yang, M., Zhu, P., Liu, F., & Shen, L. (2015). Joint representation and pattern learning for robust face recognition. *Neurocomputing*, *168*, 70-80.

Yoganand, A. V., Kavida, A. C., & Devi, D. R. (2020). Pose and occlusion invariant face recognition system for video surveillance using extensive feature set.

International Journal of Biomedical Engineering and Technology, *33*(3), 222-239.

Yuan, C., & Niemann, H. (2000). An appearance based neural image processing algorithm for 3-D object recognition. Proceedings 2000 International Conference on Image Processing (Cat. No. 00CH37101),

Zakharov, E., Shysheya, A., Burkov, E., & Lempitsky, V. (2019). Few-shot adversarial learning of realistic neural talking head models. Proceedings of the IEEE/CVF International Conference on Computer Vision, p. 9459-9468.

Zang, F., Zhang, J., & Pan, J. (2012). Face recognition using Elasticfaces. *Pattern Recognition*, *45*(11), 3866-3876.

Zangeneh, E., Rahmati, M., & Mohsenzadeh, Y. (2020). Low resolution face recognition using a two-branch deep convolutional neural network architecture. *Expert Systems with Applications*, *139*, 112854.

Zeng, D., Veldhuis, R., & Spreeuwers, L. (2020). A survey of face recognition techniques under occlusion. *arXiv preprint arXiv:2006.11366*.

Zhang, Z., Song, Y., & Qi, H. (2017). Age progression/regression by conditional adversarial autoencoder. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 5810-5818.

Zhao, X., He, Z., Zhang, S., Kaneko, S., & Satoh, Y. (2013). Robust face recognition using the GAP feature. *Pattern Recognition*, *46*(10), 2647-2657.

Zhou, B., Lapedriza, A., Khosla, A., Oliva, A., & Torralba, A. (2017). Places: A 10 million image database for scene recognition. *IEEE transactions on pattern analysis and machine intelligence 2017 Jul 4;40(6):1452-64*.