



University of Zawia
Faculty of Science
Department of Computer Science

Face Recognition using Multi-Resolution Multiple Color Fusion model based on Genetic Algorithm (MMCF-GA)

Student Name: Bushra Ali Omar Abosloge

Name of Supervisor: Dr. Mustafa Alrjebi
(Associate Professor)

**Thesis Was Submitted in Partial Fulfillment of The Requirements for
The Degree of Master in Computer Science**

Date (1/1/2026)



قرار لجنة البحث لإجازة (/) العالبة () الدقيقة ()

عملاً بقرار السيد: رئيس الجامعة رقم (1281) م لسنة (2025) الصادر بتاريخ 10 - 12 - 2025 م قامت اللجنة المشكلّة بمناقشة الرسالة المقدّمة من الطالبة : بشرى علي عمر ابوسلوقي لنيل درجة الإجازة العالية "الماجستير" في (الحاسوب).
عنوانها: (التعرف على الوجه باستخدام نموذج الصور متعدد الألوان والأحجام اعتماداً على الخوارزمية الجينية)

وتتكون اللجنة من الأساتذة :

- | | | |
|-------------------------|----------------|----------------|
| د. مصطفى المختار الرجبي | مشرقاً ومقرراً | جامعة الزاوية. |
| د. جبريل رمضان امبارك | عضواً | جامعة الجفارة. |
| د. نهى عمران ابوخير | عضواً | جامعة الزاوية. |

وبعد مناقشة الرسالة على تمام الساعة (العاشرة صباحاً) من يوم الخميس الموافق 2026/1/1 م بمدرج كلية العلوم - جامعة الزاوية ، وتقويم مستواها العلمي والمنهج الذي اتبعته الباحثة والمصادر والمراجع التي استخدمتها في دراستها قررت اللجنة ما يلي :-

القرار

بعد إتمام الطالبة : بشرى علي عمر ابوسلوقي ، لمتطلبات الدراسات العليا وإمتحانها ومناقشة الرسالة وتقويمها تقرر:

ويمنح الطالب فرصة للأخذ بالملاحظات خلال 1 أشهر من تاريخ المناقشة)

ويمنح الطالب فرصة أخرى للمناقشة خلال أشهر (

التوقيع بتاريخ المناقشة التوقيع بعد الأخذ بالملاحظات - التاريخ

..... التوقيع /

..... التوقيع /

..... التوقيع /

1. إجازتها بدون تعديلات ☐

2. إجازتها بتعديلات ☒

3. عدم إجازتها ☐

أسماء أعضاء اللجنة :

1. د. مصطفى المختار الرجبي

2. د. جبريل رمضان امبارك

3. د. نهى عمران ابوخير

مدير إدارة الدراسات العليا والتدريب

د. حمزة الهادي كشلاف

..... التوقيع /

مدير مكتب الدراسات العليا بالكلية

د. نوري الأمير كشلاف

..... التوقيع /

منسق التخصص

د. نهى عمران ابوخير

..... التوقيع /

صورة ..

إدارة الدراسات العليا والتدريب

منسق التخصص / الملف الشخصي للطالب

zainab -1/2026

Declaration

I Bushra Ali Omar Abosloge confirm that the work contained in this thesis / dissertation, unless otherwise referenced is the researcher's own work, and has not been previously submitted to meet requirements of an award at this University or any other higher education or research institution, I furthermore, cede copyright of this thesis / dissertation in favour of University of Zawia

Student name: Bushra Ali Omar Abosloge

Signature:

A handwritten signature in purple ink, consisting of stylized, overlapping loops, enclosed within a thin purple oval border.

Date: 1 / 1 / 2026

Abstract

Facial recognition is a critical area in computer vision with wide-ranging applications in security, identity verification, and human-computer interaction. Recent research has shown that leveraging multiple color spaces and multi-resolution representations can significantly enhance recognition accuracy compared to traditional RGB or grayscale approaches. However, existing multi-channel models such as the Multi-resolution Multiple Color Fusion (MMCF) framework depend on a greedy search algorithm to select optimal channel combinations, which is prone to local optima and limited exploration of the search space.

This thesis proposes a novel optimization framework that replaces the greedy search with a Genetic Algorithm (GA) to identify the most effective combination of MMCF for Face Recognition. Two GA-based models, MMCF-GA-1 and MMCF-GA-2, were developed using different classifiers Sparse Representation-based Classification (SRC) and Linear Discriminant Analysis (LDA) and evaluated on multiple benchmark datasets under diverse and challenging conditions, including pose variations, facial expressions, and unconstrained environments.

Experimental results demonstrate that the proposed GA-based model consistently outperforms the baseline RGB as well as the MMCF representations, achieving higher recognition accuracy and improved robustness across all tested scenarios. Furthermore, integrating the genetic models with advanced pose correction techniques, such as the Stacked Progressive Autoencoder (SPAEC) and the Random Mask Attention Generative Adversarial Network (RMAGAN), further enhanced their adaptability to large pose variations. This robustness was particularly evident in the highly demanding Single Sample Per Person (SSPP) scenario, where the models maintained high accuracy despite minimal training data.

The findings confirm that combining multi-resolution color information with robust optimization strategies can significantly strengthen Face Recognition systems.

Keywords: Face Recognition, Genetic Algorithm, Multi-Resolution Multiple Color Fusion, Feature Optimization, Color Models.

Dedication

To my beloved mother, your prayers have been a light illuminating my path, and the beat of your heart a source of my strength. You instilled in me a love for knowledge and taught me patience and perseverance throughout my academic journey. You have always been my greatest support and refuge.

To my brothers and sisters, my constant source of strength and companions along the way, who never ceased to offer their help and encouragement. I extend to you my deepest gratitude and appreciation for your love and support.

To the soul of my dear father, though your body is gone, your wisdom and kindness continue to light my way.

May Allah bless your soul with eternal mercy.

I dedicate to you the fruit of this effort, in recognition and gratitude for every moment of love, care, and encouragement you have given me.

Acknowledgement

All praise is due to Allah, by whose grace good deeds are completed and through whose blessings this scientific work has been accomplished. He has granted me the strength, patience, and determination to continue the journey despite all challenges.

I would like to express my sincere gratitude and deep appreciation to my esteemed supervisor, Dr. Mustafa Alrjebi, for his continuous academic guidance, invaluable advice, and unwavering support, which had a profound impact on the completion of this research.

My heartfelt thanks also go to my beloved mother, whose prayers have been my shield and whose patience has been my driving force throughout every stage of my studies. She embraced my worries and fueled my ambition.

I am also deeply grateful to my dear family members my brothers and sisters for their constant moral support and encouragement, which have been a cornerstone in my academic journey.

TABLE OF CONTENTS

chapter	TITLE	page
	DECLARATION	i
	ABSTRACT	ii
	DEDICATION	iii
	ACKNOWLEDGEMENT	iv
	TABLE OF CONTENTS	v
	LIST OF TABLES	ix
	LIST OF FIGURES	x
	LIST OF ABBREVIATIONS	xii
1	Introduction	1
	1.1. Introduction	2
	1.2. Problem Statement	3
	1.3. Research Objectives	3
	1.4. Research Questions	4
	1.5. Main Contribution	4
	1.6. Thesis Organization	4
2	Literature Review	5
	2.1. Introduction	6
	2.2. Face Recognition Based on Subspace Approaches	6
	2.2.1. Principal Component Analysis (PCA)	6
	2.2.2. Linear Discriminant Analysis (LDA)	8
	2.2.3. Sparse Representation-based Classification (SRC)	9
	2.3. The Role of Color Information in Face Recognition	11

2.3.1. Color Spaces	11
2.3.2. The RGB Color Model	12
2.3.3. The YIQ Color Space	12
2.3.4. The YcbCr Color Space	12
2.3.5. The I1I2I3 Color Space	13
2.3.6. The CIE XYZ Color Space	13
2.3.7. The YUV Color Space	14
2.3.8. The Normalized Color Spaces: nRGB and nXYZ	14
2.3.9. Discriminant Color Space (DCS)	15
2.3.10. Local Discriminant Color Space (LDCS)	15
2.3.11. Independent Color Space (ICS)	16
2.3.12. Holistic Orthogonal Analysis (HOA)	16
2.4. Multi-Channel Representation	16
2.4.1. Multiple Color Fusion (MCF)	16
2.4.2. Two-Directional Multiple Color Fusion (2D-MCF)	17
2.4.3. Multi-Resolution Multiple Color Fusion (MMCF)	17
2.5. The Greedy Search Algorithm	18
2.6. The Genetic Algorithm (GA)	19
2.7. Summary	20
3 Research Methodology	21
3.1. Introduction	22
3.2. Overall Framework	22
3.3. Defining the Baseline: The Original MMCF Model	22
3.4. Proposed Model: Multi Resolution Multiple Color Fusion using Genetic Algorithm (MMCF-GA)	24
3.5. Dataset and Preprocessing Consistency	24
3.5.1. GA Representation Framework	24
3.5.2. Initial Population and Redundancy Repair Mechanism	26
3.5.3. Fitness Function	28
3.5.4. Evolutionary Operators	28

3.6. Construction of the MMCF-GA-1 Model	34
3.7. Construction of the MMCF-GA-2 Model	34
3.8. Handling Pose Variations: Integration with Stacked Progressive Auto Encoder (SPAЕ)	35
3.8.1. The Principle of Stacked Progressive Auto Encoder (SPAЕ)	35
3.8.2. Training SPAЕ Models for Color Channels	36
3.8.3. Pose Correction and Face Recognition	36
3.9. Utilizing Advanced Pose Correction Methods for Dataset Creation	37
3.10. Software and Hardware	38
3.11. Summary	38
4 Experimental Setup and Results	39
4.1. Introduction	40
4.2. Databases Setup	40
4.2.1. AR database	40
4.2.2. The Curtin Database	41
4.2.3. Multi-Pie Database	41
4.2.4. Labeled Faces in the Wild Face Database	41
4.2.5. Georgia Tech Face Database	42
4.2.6. The Facial Recognition Technology database	42
4.2.7. Psychological Image Collection face databaset	42
4.2.8. The Face Recognition Grand Challenge Database	43
4.3. Experimental Evaluation and Results	43
4.4. Experimental Results Using the LDA Classifier	44
4.4.1. Experiment One	44
4.4.2. Experiment Two	45
4.4.3. Experiment Three	46
4.4.4. Experiment Four	47
4.4.5. Experiment Five	48
4.4.6. Experiment Six	49
4.5. Experimental Results Using the SRC Classifier	51
4.5.1. Experiment Seven	51

4.5.2. Experiment Eight	52
4.5.3. Experiment Nine	53
4.5.4. Experiment Ten	54
4.5.5. Experiment Eleven	55
4.5.6. Experiment Twelve	56
4.6. Face Recognition Experiment under Pose Variations and Single Sample Per Person (SSPP)	57
4.6.1. Experimental Evaluation of SPAE with Genetic Models	57
4.6.2. Evaluation on the Multi-Pie Database	57
4.7. Evaluation of Genetic Models (MMCF-GA-1 and MMCF-GA-2) Using LFW Dataset Corrected by the Random Mask Attention Generative Adversarial Network (RMAGAN)	60
4.8. Discussion	62
5 Conclusion and recommendations for Future Works	66
5.1. Conclusion	67
5.2. Future Works	67
6 References	68
7 Arabic Abstract	73

LIST OF TABLES

Table	TITLE	PAGE
1	The Official MMCF Representation Adopted as the Baseline.	24
2	Example of Chromosome Encoding with 6 Genes.	25
3	The optimal feature representation for the MMCF-GA-1 Model using SRC classifier.	34
4	The optimal feature representation for the MMCF-GA-2 Model using LDA classifier.	35
5	Databases Used for Face Recognition with MMCF-GA-1 and MMCF-GA 2 models.	43
6	Recognition accuracy applied on AR+CURTIN database and LDA classifier.	44
7	Recognition Accuracy Applied the GT Database and LDA Classifier.	45
8	Recognition Accuracy on the LFW Database and LDA Classifier.	47
9	Recognition Accuracy on the FERET and LDA Classifier.	48
10	Recognition Accuracy on the PICS Database and LDA Classifier.	49
11	Recognition Accuracy on the FRGC Database and LDA Classifier.	50
12	Recognition Accuracy on the AR + Curtin Dataset Using SRC Classifier.	51
13	Recognition Accuracy Applied the GT Database and SRC Classifier.	52
14	Recognition Accuracy Applied the LFW Dataset Using SR Classifier.	53
15	Recognition Accuracy Applied the FERET Dataset Using SRC Classifier.	54
16	Recognition Accuracy Applied the PICS Dataset Using SRC Classifier.	55
17	Recognition accuracy applied the FRGC Dataset Using SRC Classifier.	56
18	Face Recognition on Multi-Pie Using SPAE-MMCF-GA-1 and SPAE MMCF-GA-2 Compared with SPAE-MMCF, SPAE-RGB, and SPAE-(G).	59
19	Face Recognition Accuracy on LFW Dataset Corrected by the RMAGAN.	60

LIST OF FIGURES

Figure	TITLE	PAGE
1	Illustrative Example of Single-Point Crossover Using 6 Genes (Crossover After Gene 4).	30
2	MMCF-GA model training using GA methodology.	33
3	Face pose correction using SPAE and the genetic model.	37
4	Displays sample images from the AR database.	40
5	Displays sample images from the Curtin database.	41
6	Displays sample images from the Multi-Pie face database.	41
7	Displays sample images from the LFW face database.	42
8	Displays sample images from the GT Face Database.	42
9	Sample images from the FERET face database.	42
10	Sample images from the PICs face database.	42
11	Displays sample images from the FRGC face database.	43
12	Comparison of Error Rates Among Face Recognition Models on the AR+CURTIN Dataset and LDA classifier.	45
13	Comparison of Error Rates Among Face Recognition Models on the GT Dataset and LDA classifier.	46
14	Comparison of Error Rates Among Face Recognition Models on the LFW Dataset and LDA classifier.	47
15	Illustrates the significant error rate reduction on the FERET database and LDA classifier.	48
16	Comparison of Error Rates Among Face Recognition Models on the PICS Dataset and LDA classifier.	49
17	Comparison of Error Rates Among Face Recognition Models on the FRGC Dataset and LDA classifier.	50
18	Comparative Analysis of Error Rates Among Models on the AR + Curtin Dataset Using SRC Classifier.	51
19	Comparative Analysis of Error Rates Among Models on the GT Dataset Using SRC Classifier.	52
20	Comparative Analysis of Error Rates Among Models on the LFW Dataset Using SRC Classifier.	53
21	Comparative Analysis of Error Rates Among Models on the FERET Dataset Using SRC Classifier.	54
22	Comparative Analysis of Error Rates Among Models on the PICS Dataset Using SRC Classifier.	55

23	Comparative Analysis of Error Rates Among Models on the FRGC Dataset Using SRC Classifier.	56
24	Comparative Analysis of Error Rates Among Models on the Pose Corrected LFW Dataset (Gallery = 1) .	61
25	Comparative Analysis of Error Rates Among Models on the Pose Corrected LFW Dataset (Gallery = 3).	61
26	Comparison of Error Rates Among MMCF-GA-1, MMCF-GA-2 models, and MMCF Model Under Various Database Conditions Using SRC Classifiers.	64
27	Comparison of Error Rates Among MMCF-GA-1, MMCF-GA-2 models, and MMCF Model Under Various Database Conditions Using LDA Classifiers.	65

LIST OF ABBREVIATIONS

Abbreviation	Meaning
2D-MCF	Two-Directional MCF
DCS	Discriminant Color Space
GAN	Generative Adversarial Network
GA	Genetic Algorithm
HOA	Holistic Orthogonal Analysis
ICS	Independent Color Space
LDA	Linear Discriminant Analysis
LDCS	Local Discriminant Color Space
MCF	Multiple Color Fusion
MMCF	Multi-resolution Multiple Color Fusion
PCA	Principal Component Analysis
RGB	Red, Green, Blue
RMAGAN	Random Mask Attention Generative Adversarial Network
SSPP	Single Sample Per Person
SRC	Sparse Representation-based Classification
SPAE	Stacked Progressive Autoencoder
SSFR	Single-Sample Face Recognition

Chapter 1

Introduction

1.1 Introduction

This chapter provides the fundamental framework for the research presented in this thesis. It begins with an overview of Face Recognition technology and its various applications, leading to the specific challenges that motivated this study. The core research problem is then identified, followed by a clarification of the research objectives, the central research question, and the main scientific contributions. The chapter concludes with an outline of the thesis structure, paving the way for the subsequent chapters.

The twenty-first century has witnessed an unprecedented technological explosion, with facial recognition standing out as one of its most remarkable achievements. This technology has attracted widespread attention due to the significant advancements it has made over the past few decades, becoming an effective tool for verifying and identifying individuals [1]. The system combines individual Face Recognition with the extraction of high-level features from images such as texture, color, and shape which are then compared against faces stored in a database to determine whether the face is known or unknown [1]. Face Recognition can be widely applied in public security and law enforcement as well as in credit card verification, criminal identification, access control, commerce, education, healthcare, human computer interaction, digital libraries, and information security [2].

Notwithstanding its potential, the performance of facial recognition systems is often degraded by several intrinsic issues associated with the subject's appearance. They range from extreme facial expressions to partial occlusions e.g. glasses or scarves and large pose variations, as well as the challenging SSPP setting where there's just one reference image per subject [3][4][5].

Early research in the field of facial recognition relied on gray-scale images due to their simplicity [6]. However, research has found that color information can further improve recognition accuracy. Some color spaces are built specifically for Face Recognition [7]. Although these color spaces contain valuable discriminant information compared to typical RGB, Face Recognition, all these-color composed of only three-color components to represent a face image [7]. Realizing that three color channels might not be informative enough for all recognition problems, researchers started to look into techniques that use a richer color models [7]. The Multiple Color Fusion (MCF) [8] model was an early attempt in this direction that showed major gains in accuracy over conventional three-channel methods. Building on this success, more advanced architectures were developed,

most notably the Two-Directional MCF (2D-MCF) [9] and the MMCF [10] models. The MMCF model provides a much richer facial representation by featuring both multiple color channels as well as multiple resolutions [10].

Despite these advancements, the optimal way to fuse information from multiple color channels and resolutions is still an open research question, highlighting the need for a more systematic approach to design and evaluate these complex models.

1.2 Problem Statement

While previous color models such as MCF, 2D-MCF, and MMCF have achieved notable results, their development is constrained by a significant limitation. The construction of these models has predominantly relied on the greedy search algorithm to identify the optimal combination of color channels [11]. The primary drawback of this approach is that the greedy algorithm is inherently myopic. It makes locally optimal choices at each step, selecting what appears best at the current moment without regard to future consequences [12]. Although this algorithm is simple to implement, its tendency to converge on sub-optimal solutions constitutes a fundamental problem, limiting the maximum potential performance of these advanced Face Recognition models.

1.3 Research Objectives

The primary objective of this research is to develop and enhance image representation to improve the accuracy and robustness of Face Recognition systems.

This is achieved by designing an optimal framework that leverages a GA [12], to optimize the selection of multi-resolution color channels. The specific, measurable objectives are:

1. To improve the feature representation of facial images by designing a search strategy based on a GA. This strategy will be developed to identify the optimal combination of MMCF, serving as a robust replacement for the sub-optimal greedy algorithm.
2. To conduct a rigorous experimental evaluation to validate the effectiveness of the new representation, by comparing the recognition accuracy of the proposed models against the original MMCF model on benchmark datasets.
3. To analyze and demonstrate that the enhanced representation, discovered via the GA-based approach, leads to superior Face Recognition performance.

1.4 Research Question

This research seeks to answer the following primary question:

1- To what extent can a GA based search strategy improve upon the greedy search algorithm in identifying an optimal combination of multi-resolution color channels, as measured by Face Recognition accuracy?

1.5 Main Contributions

The main contribution of this thesis lies in the development and enhancement of the search method used to construct multi-resolution, multi-color Face Recognition models. Specifically, this research makes the following contributions: First, a GA is proposed and implemented to replace the sub-optimal greedy search algorithm used in previous models (e.g., MMCF). This new methodology is designed to more effectively overcome the problem of local optima, and explore a larger solution space. Second, as a direct result of this enhanced search method, a novel color model is developed: Multi Resolution Multiple Color Fusion using Genetic Algorithm MMCF-GA. Through rigorous experimentation, this model is shown to exhibit superior performance, achieving a significant increase in Face Recognition accuracy compared to its greedy-based counterparts.

1.6 Thesis Organization

The remainder of this thesis is organized as follows:

- Chapter 2 reviews the relevant literature on color spaces, multi-channel Face Recognition models, and the optimization algorithms central to this research.
- Chapter 3 details the proposed research methodology, including the design of the GA and the step-by-step construction of the new MMCF-GA model.
- Chapter 4 presents the experimental setup, including the datasets and evaluation protocols. It then reports the results and provides a comprehensive discussion and analysis of the findings.
- Finally, chapter 5 concludes the thesis, and offers recommendations for future work.

Chapter 2

Literature Review

2.1 Introduction

This chapter provides a concise review of the literature on Face Recognition, beginning with classical feature extraction methods such as Principal Component Analysis (PCA) [13], LDA [14], and SRC [15]. Followed by the role of color information and color spaces [16], as well as multi-channel fusion strategies including MCF [8], the Two-Directional Multiple Color Fusion (2D-MCF) [9], and the MMCF [10].

The discussion extends to modern architectures such as the SPAE framework [17], which has been integrated with MMCF to address challenges related to face pose variations and the SSPP problem, alongside alternative approaches based on Generative Adversarial Network (GAN) algorithms [18]. The chapter concludes by identifying the limitations of greedy search algorithms [11], as a key research gap.

2.2 Face Recognition Based on Subspace Approaches

Subspace methods have been a cornerstone in Face Recognition for several decades. In image processing, Face Recognition data is often represented as high-dimensional feature vectors [19]. This is because images can be of a large size, meaning they contain a high number of redundant pixels and irrelevant information. High-dimensional data is known to be susceptible to noise and feature redundancy, which can lead to inaccurate results [20]. To overcome these problems, researchers have developed several dimensionality reduction techniques aimed at reducing the dimensions of face images while retaining the most important information. This section will review some of the most commonly dimensionality reduction techniques used in Face Recognition.

2.2.1 Principal Component Analysis (PCA)

Principle Component Analysis (PCA) [13] is a fundamental linear technique used for dimensionality reduction. It employs an orthogonal transformation to convert a set of potentially correlated variables into a new set of linearly uncorrelated variables known as principal components. The primary objective of PCA is to reduce the dimensionality of a dataset while retaining the maximum amount of variance, which is considered to be the most significant information [20].

In Face Recognition, PCA is a powerful tool for feature extraction. It achieves this by identifying the directions the principal components in which the data varies the most. By projecting the original high-dimensional data onto a smaller number of these principal

components, a lower-dimensional representation can be obtained with minimal information loss [13][20].

The procedural steps of applying PCA to a set of face Images are as follows:

1. Data Representation: Given a training set of M facial images, each image is first vectorized into a column vector Γ_i of size N^2 (where $N \times N$ is the image resolution).
2. Compute the Mean Face: The average face vector, Ψ , is calculated by averaging all training image vectors:

$$\Psi = (1/M) * \sum(\Gamma_i) \quad \text{for } i=1 \text{ to } M \quad (2.1)$$

3. Centralize the Data: Each image vector Γ_i is centered by subtracting the mean face Ψ . This results in a set of zero-mean vectors, $\Phi_i = \Gamma_i - \Psi$. These vectors can be aggregated into a matrix $A = [\Phi_1, \Phi_2, \dots, \Phi_m]$.
4. Compute the Covariance Matrix: The covariance matrix C is computed to capture the variance within the dataset: $C = A * A^T$ (2.2)

However, the size of C is $N^2 \times N^2$, which is computationally prohibitive for typical image sizes. A more efficient method, known as the snapshot or trick method, is to compute a smaller $M \times M$ matrix, $L = A^T * A$.

The eigenvectors v_i of L can then be used to find the eigenvectors u_i of the original covariance matrix C by the relation $u_i = A * v_i$.

5. Identify Principal Components (Eigenfaces): The eigenvectors u_i of the covariance matrix C are the principal components, often referred to as eigenfaces in the context of Face Recognition. These eigenvectors form an orthogonal basis for the new, lower-dimensional feature space.
6. Dimensionality Reduction: The eigenvectors are sorted according to their corresponding eigenvalues in descending order. To reduce dimensionality, only the top M' eigenvectors (where $M' < M$) are selected to form the projection matrix $W = [u_1, u_2, \dots, u_{\{M'\}}]$. The choice of M' is often determined heuristically to retain a desired percentage of the total variance.

7. Feature Extraction and Classification: To obtain the feature vector Ω for a given test image Γ_{test} , it is first centered and then projected onto the eigenface space using the projection matrix W : $\Omega = W^T * (\Gamma_{\text{test}} - \Psi)$ (2.3)

The resulting feature vector Ω is a compact, lower-dimensional representation of the face. Recognition is then performed by comparing this vector to the feature vectors of the training images.

Despite its effectiveness in data representation, PCA has a significant drawback from a classification perspective. It operates in an unsupervised manner, meaning it does not consider class labels during dimensionality reduction. By focusing solely on maximizing variance, it may discard components that, while having low variance, which is crucial for discriminating between different individuals [20][21].

This limitation is addressed by LDA, which is presented in the next section.

2.2.2 Linear Discriminant Analysis (LDA)

Unlike the unsupervised PCA, LDA [14] is a supervised dimensionality reduction technique. Its primary objective is to find a lower-dimensional subspace that maximizes the separability between different classes/ individuals while minimizing the variance within each class. This is achieved by finding a projection that maximizes the ratio of between-class scatter to within-class scatter [14][22].

The main steps for applying LDA in Face Recognition are as follows:

- **Compute Scatter Matrices:** The first step is to quantify the within-class and between-class scatter.
- The within-class scatter matrix (S_w) measures the scatter of samples around their respective class means. It is computed as:

$$S_w = \sum_{i=1}^C \sum_{k=1}^{N_i} (x_k - \mu_i)(x_k - \mu_i)^T \quad \text{for } i=1 \text{ to } C, \text{ for each } x_k \text{ in class } I \quad (2.4)$$

- The between-class scatter matrix (S_b) measures the scatter of the class means around the overall mean of all data. It is computed as:

$$S_b = \sum_{i=1}^C N_i (\mu_i - \mu)(\mu_i - \mu)^T \quad \text{for } i=1 \text{ to } C \quad (2.5)$$

Here, C is the number of classes, μ_i is the mean vector of class I , N_i is the number of samples in class I , and μ is the total mean vector of all samples.

Formulate the Optimization Criterion: LDA seeks a projection matrix, W_{lda} , that maximizes the ratio of the determinant of the between-class scatter to the determinant of the within-class scatter in the projected space. This is known as Fisher's criterion:

$$J(W) = |W^T * S_b * W| / |W^T * S_w * W| \quad (2.6)$$

Solve the Generalized Eigenvalue Problem: The projection matrix W_{lda} that maximizes this ratio can be found by solving the generalized eigenvalue problem:

$$S_b * v_i = \lambda_i * S_w * v_i \quad (2.7)$$

The columns of the optimal projection matrix W_{lda} are the eigenvectors (v_i) corresponding to the largest eigenvalues (λ_i).

Dimensionality Reduction and Feature Extraction: The dimensionality is reduced by selecting the top $C-1$ eigenvectors to form the projection matrix W_{lda} . A high-dimensional input face x is then projected to a lower-dimensional feature vector y as follows: $y = W_{lda}^T * x$ (2.8)

This feature vector y can then be used with a classifier, for recognition.

The Singularity Problem in LDA: A major challenge in applying LDA to Face Recognition is the singularity problem. This occurs when the number of training samples is less than the dimensionality of the feature space, causing the within-class scatter matrix S_w to become singular (i.e., its determinant is zero) and thus non-invertible. This is problematic because solving the eigenvalue problem in Equation (2.7) requires the inverse of S_w . A common solution is to first apply PCA to reduce the dimensionality of the data to a point where S_w becomes non-singular, and then apply LDA on the resulting PCA subspace. However, this two-stage approach has a significant drawback: the PCA step, being unsupervised, might discard some discriminant information that is valuable for LDA. To address this, alternative methods like Regularized LDA (R-LDA) have been proposed [7][23]. R-LDA modifies the within-class scatter matrix by adding a scaled identity matrix, ensuring it is always invertible. The R-LDA objective function is:

$$J(W) = |W^T * S_b * W| / |W^T * (S_w + \alpha I) * W| \quad (2.9)$$

where α is a regularization parameter and I is the identity matrix. This allows LDA to be applied directly to the original data without the risk of singularity, potentially preserving more discriminant information [7][23].

2.2.3 Sparse Representation-based Classification (SRC)

Moving beyond the subspace projection methods of PCA and LDA, Sparse Representation-based Classification SRC [15] offers a fundamentally different paradigm for Face Recognition. Instead of learning a compact, low-dimensional subspace, SRC operates directly in the high-dimensional space and leverages the principle of sparsity. The core assumption is that a test image of a particular individual can be efficiently represented as a sparse linear combination of all training samples, where the only significant (non-zero) coefficients in this combination will correspond to the training samples of that same individual [7][15]. This approach has demonstrated remarkable robustness against challenges like occlusions, illumination changes, and facial expressions, making it a highly effective classification framework [24].

The mathematical formulation of SRC can be outlined as follows:

1. **Problem Formulation:** Let A be a matrix whose columns are the training samples from K different classes (individuals). A test sample y from one of these classes can be expressed as a linear combination of all training samples:

$$y = A * x \quad (2.10)$$

Here, x is a vector of coefficients. The "sparsity" assumption implies that x should be a sparse vector, with most of its entries being zero, except for those corresponding to the class of y .

2. **Finding the Sparsest Solution:** The goal is to find the sparsest possible coefficient vector x that satisfies the equation. This is formulated as an ℓ_0 -minimization problem, which seeks to minimize the number of non-zero elements in x :

$$\hat{x}_0 = \arg \min \|x\| \text{ subject to } y = A * x \quad (2.11)$$

However, solving the ℓ_0 -minimization problem is NP-hard. It has been shown that if the solution x is sufficiently sparse, this problem is equivalent to the more tractable ℓ_1 -minimization problem, which can be solved efficiently using convex optimization techniques:

$$\hat{x}_1 = \arg \min \|x\|_1 \text{ subject to } y = A * x \quad (2.12)$$

3. **Classification via Reconstruction Error:** Once the sparse coefficient vector $\hat{x}_1$ is found, the classification is not done by directly inspecting the coefficients, especially in the presence of noise. Instead, SRC uses a more robust method based on reconstruction error. For each class I , we reconstruct the test image y using only the coefficients associated with that class.

Let $\delta_i(\hat{x}_1)$ be a function that selects only the coefficients belonging to class I from the vector $\hat{x}_1$. The reconstructed image for class I is: $\hat{y}_i = A_i * \delta_i(\hat{x}_1)$

The classification is then performed by assigning the test sample y to the class that yields the smallest reconstruction error, calculated as the residual $r_i(y)$:

$$\text{Identity}(y) = \arg \min_i \|y - \hat{y}_i\|_2 = \arg \min_i \|y - A_i * \delta_i(\hat{x}_1)\|_2$$

In essence, SRC classifies a test image by determining which class's training samples can "reconstruct" it most accurately, guided by the principle of sparsity. This reliance on reconstruction error provides its robustness, as minor occlusions or noise will have a limited impact on the overall error calculation for the correct class [15][25].

2.3 The Role of Color Information in Face Recognition

While the subspace methods discussed previously, such as PCA, LDA, and SRC, laid the foundational groundwork for Face Recognition, they were predominantly developed and evaluated on grayscale images. For many years, grayscale was the standard representation, favored for its computational simplicity and lower data dimensionality [6]. However, this approach inherently discards a rich source of discriminative information available in color images. Research has consistently demonstrated that color is not merely an aesthetic property but a powerful feature for enhancing recognition accuracy [26][27][28]. Different color channels can encode distinct information related to skin tone, hair color, and subtle texture variations, which are often lost in a grayscale conversion. This additional information proves particularly valuable in challenging, real-world scenarios involving variations in lighting, facial expression, and even age estimation. Consequently, a significant body of research has shifted focus towards developing models that can effectively leverage color [26][27][28].

The following sections will review the fundamental building blocks of color-based recognition, starting with the color spaces used to represent and manipulate this information.

2.3.1 Color Spaces

A color space is a mathematical model that defines how colors can be represented as tuples of numbers, typically as a combination of three-color channels. Its primary purpose is to provide a standardized framework for specifying, creating, and visualizing color. While many color spaces were not originally designed for pattern recognition tasks, their unique properties make them valuable for Face Recognition. The choice of color space can significantly impact how features are extracted, as each space encodes color and luminance information differently. The foundational model for most digital color representation is the RGB (Red, Green, and Blue) space [16], from which nearly all other color spaces can be derived through linear or non-linear transformations. The following subsections will provide an overview of the RGB model and other key color spaces relevant to this research [16].

2.3.2 The RGB Color Model

The RGB model is structured around a Cartesian coordinate system, where each color is defined by its three primary components: Red (R), Green (G), and Blue (B). An RGB image with dimensions of M rows and N columns is represented as an $M \times N \times 3$ array of colored pixels. In this context, the 3 indicates the three layers of intensity for Red, Green, and Blue [15]. The RGB representation can be denoted as shown in Eq. (2.13).

$$I_{\text{RGB}} = \begin{bmatrix} R \\ G \\ B \end{bmatrix} \in \mathbb{R}^{3 \times d}, \quad d = M \times N \quad (2.13)$$

While the RGB model is well-suited for various technical applications, it has limited effectiveness for image segmentation and analysis due to the high correlation among the R, G, and B components. Additionally, it is not a perceptual model; that is, the R, G, and B components include both color information and illumination data. The primary advantage of the RGB color space is its simplicity [16]. For these reasons, a variety of mathematical transformations can be applied to the RGB color space to derive other color spaces, some common examples of which will be discussed next.

2.3.3 The YIQ Color Space

The YIQ color space, formerly used in the National Television System Committee (NTSC) television standard [29], is derived from the RGB color space through a linear transformation matrix, as defined in Eq. (2.14).

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.275 & -0.321 \\ 0.212 & -0.523 & 0.311 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (2.14)$$

This transformation separates the image data into a luminance component (Y) and two chrominance components (I and Q).

The key advantage of this separation, particularly for Face Recognition, is that it allows for the independent processing of illumination (the Y channel), which can help in building models that are more robust to lighting variations [29][30].

2.3.4 The YcbCr Color Space

Similar to YIQ, the YcbCr color space is widely used in digital video and image compression standards [31]. It also separates image data into a luminance component (Y)

and two color-difference chrominance components: Cb (blue-difference) and Cr (red-difference). The conversion from RGB to YcbCr is defined by the transformation matrix in Eq. (2.15).

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 65.481 & 128.533 & 24.966 \\ -37.797 & -74.203 & 112.000 \\ 112.000 & -93.786 & -18.214 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (2.15)$$

The key motivation for its application in Face Recognition is identical to that of YIQ: the decoupling of luminance (Y) from chrominance (Cb, Cr). This separation makes the representation more robust to illumination variations, a critical advantage over the standard RGB space for recognition tasks [31].

2.3.5 The I1I2I3 Color Space

The I1I2I3 color space is designed to address the high correlation inherent in the RGB model by achieving channel decorrelation. This is accomplished by applying the Karhunen-Loève Transform (K-L Transform) a method statistically equivalent to PCA to the RGB data.

The resulting transformation, defined in Eq. (2.16), produces three statistically uncorrelated components (I1, I2, I3). By minimizing data redundancy, this creates a more efficient and compact representation, which is highly advantageous for feature extraction algorithms in Face Recognition tasks [29][32].

$$\begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix} = \begin{bmatrix} 0.333 & 0.500 & -0.500 \\ 0.333 & 0.000 & 1.000 \\ 0.333 & -0.500 & -0.500 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (2.16)$$

2.3.6 The CIE XYZ Color Space

The International Commission on Illumination (CIE) developed the 1931 XYZ color space as the first mathematically defined standard based on human visual perception [29]. It is derived from RGB through a linear transformation, as shown in Eq. (2.17), to produce three tristimulus values (X, Y, Z).

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.607 & 0.174 & 0.201 \\ 0.299 & 0.587 & 0.114 \\ 0.000 & 0.066 & 1.117 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (2.17)$$

Its relevance to Face Recognition lies in its design: the Y component directly corresponds

to the luminance (perceived brightness) of a color. This intentional separation allows for illumination normalization to be performed on a perceptually meaningful channel, offering a distinct advantage over purely technical color models for recognition tasks [29][32].

2.3.7 The YUV Color Space

The YUV color space belongs to the same family of luminance-chrominance models as YIQ and YcbCr, and is primarily associated with the PAL analog video standard [29]. It operates on the identical principle of decoupling image information into one luminance channel (Y) and two chrominance channels (U and V). The transformation from RGB to YUV is defined in Eq. (2.18).

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.2990 & 0.5870 & 0.1140 \\ -0.1471 & -0.2888 & 0.4359 \\ 0.6148 & -0.5148 & -0.1000 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (2.18)$$

Given its conceptual similarity to YIQ and YcbCr, the motivation for its application in Face Recognition is the same: to isolate the luminance component (Y) for independent processing. This allows for illumination normalization techniques to be applied without affecting the color data, thereby enhancing the model's robustness against lighting variations [29].

2.3.8 The Normalized Color Spaces: nRGB and nXYZ

Normalization is a technique used to make color representations less sensitive to variations in lighting intensity [32]. In this context, the normalized color spaces nRGB and nXYZ were proposed to enhance Face Recognition accuracy by achieving the Double Zero Sum (DZS) property, a characteristic also found in spaces like I1I2I3 that has been linked to improved performance [32][33].

These spaces are derived from their parent models through the linear transformations defined in Eq. (2.19) and Eq. (2.20), respectively.

$$\begin{bmatrix} nR \\ nG \\ nB \end{bmatrix} = \begin{bmatrix} 1.000 & -0.333 & -0.333 \\ 0.000 & 0.666 & -0.333 \\ 0.000 & -0.333 & 0.666 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (2.19)$$

$$\begin{bmatrix} nX \\ nY \\ nZ \end{bmatrix} = \begin{bmatrix} 0.6070 & -0.0343 & -0.3940 \\ 0.1740 & 0.2537 & -0.3280 \\ 0.2000 & -0.2193 & 0.7220 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (2.20)$$

While normalization primarily addresses illumination changes, other specialized color spaces have been developed to tackle different challenges, such as maximizing the discriminative power of the features themselves. The following section will discuss examples of these discriminant color spaces.

2.3.9 Discriminant Color Space (DCS)

Unlike other methods, the DCS directly optimizes for class separability [33]. It creates new, highly discriminative color channels by forming linear combinations of the original RGB components. This approach is theoretically optimal for classification. For instance, the RGB-r space uses difference channels like R-G to

$$\bullet \quad R_r = R \quad G_r = R - G \quad B_r = R - B$$

capture subtle facial variations, resulting in a representation that improves generalization and classification accuracy [33][34].

2.3.10 Local Discriminant Color Space (LDCS)

The LDCS extends the DCS framework by adapting the discriminant analysis to local data structures. Instead of computing one global transformation, the core principle of LDCS is to first partition the color feature space into several local regions and then derive a distinct, optimized transformation for each region [35]. This approach is implemented through two primary variations:

- **Pixel-Level (PLDCS):** This method operates at the most granular scale, applying a unique transformation to each pixel. It excels at capturing fine-grained, intricate color details but incurs significant computational cost.
- **Block-Wise (BWDCS):** This method offers a compromise by partitioning the image into blocks and computing a shared set of features for each. It balances the need for local information with computational feasibility, reducing the complexity inherent in pixel-level processing [35].

2.3.11 Independent Color Space (ICS)

Shifting the focus from maximizing class separability, another strategy for engineering color spaces is to minimize the statistical dependency between color channels. This is the objective of the ICS, which is derived by applying ICA to the original RGB data [36]. ICA is a computational method that seeks to find a linear transformation that projects the data onto a basis where the resulting components are as statistically independent as possible. Unlike PCA, which only ensures that the components are uncorrelated, ICA imposes a much stronger criterion of statistical independence. When applied to face images, this transformation yields a representation where channels like luminance, skin tones, and other features are separated into distinct, non-redundant components. This reduction in redundancy can lead to more efficient and robust feature extraction for recognition tasks [36].

2.3.12 Holistic Orthogonal Analysis (HOA)

HOA aims to reduce the correlation of color information at the feature level. It achieves this by using the Fisher criterion to obtain discriminant transforms for the red, green, and blue components, while also ensuring these transforms are mutually orthogonal [36].

2.4 Multi-Channel Representation

Since the three-color channels might lack sufficient information to effectively tackle Face Recognition issues, researchers have developed approaches that utilize multiple color channels. The following sections describe these models.

2.4.1 Multiple Color Fusion (MCF)

The study [8] introduced the MCF model as an early systematic attempt to overcome the limitations of traditional three-channel color representations by leveraging an extended color space consisting of 39 components. Although this approach demonstrated effectiveness by achieving an 80% verification rate, the methodology employed for model construction raises several methodological concerns. Specifically, the model relies exclusively on a greedy search algorithm to select the final twelve channels. From an optimization perspective, this strategy is inherently limited, as greedy search focuses on local optima at each step and lacks the ability to capture non-linear interactions among channels. As a result, the selected combination may not correspond to the global optimum.

Furthermore, the model adopts a holistic processing strategy, which subsequent studies have shown to be insufficient for handling local facial variations, such as partial illumination, occlusion, or pose variations. This combination of a constrained search mechanism and a rigid holistic approach highlight a clear research gap. These limitations have led to subsequent advancements such as the 2D-MCF model [9].

2.4.2 Two-Directional Multiple Color Fusion (2D-MCF)

Based on the foundational MCF model, the 2D-MCF model [9] was proposed to mitigate the limitations of holistic processing by adopting a block-wise strategy. Although this approach succeeded in capturing local discriminative information, it reproduced some of the methodological shortcomings of its predecessor. Specifically, applying the greedy search algorithm independently to each non-overlapping block increases the risk of falling into suboptimal local solutions. This fragmented optimization neglects potential global correlations among different facial regions, which may result in representations that are locally optimized but inconsistent at the global level.

Furthermore, although integrating the 2D-MCF model with the SRC classifier [24] improved the ability to handle partial occlusions, the framework remains constrained by a strict dependency on precise facial alignment. Any slight spatial misalignment can disrupt the correspondence between blocks, effectively limiting the model's effectiveness to frontal and carefully controlled environments. This structural rigidity, coupled with the inefficiency of the greedy selection mechanism at the block level, represents a significant research gap. These limitations subsequently motivated the development of the MMCF model [10].

2.4.3 Multi-Resolution Multiple Color Fusion (MMCF)

The MMCF model [10] extends earlier fusion-based approaches, such as MCF and 2D-MCF, by introducing a pivotal new dimension: image resolution. While previous models operated at a single, fixed scale, MMCF is based on the premise that different resolutions capture varying levels of facial detail, and that an optimal representation can be achieved by fusing color channels across multiple scales [7][10].

The core methodology of MMCF begins with converting all input images into nine distinct color spaces, followed by resizing each color channel to four different resolutions (e.g., 32×32 , 24×24 , 16×16 , 12×12). Subsequently, the standard MCF steps are applied,

including feature extraction using SRC and channel selection through a greedy search algorithm. Although this approach achieves notable improvements in recognition accuracy compared to the original MCF [7][10], its reliance on greedy search inherently limits the model’s ability to consistently reach globally optimal representations.

MMCF has been evaluated under pose variation conditions as well as in the challenging SSPP scenario, demonstrating superior performance relative to conventional representations such as RGB or grayscale, particularly at large pose angles. Integration with the SPAE framework [17] further enhanced performance, as SPAE normalizes non-frontal faces to a frontal view through a step-wise transformation [10][17]. Nevertheless, the model’s performance remains heavily constrained by the requirement for precise facial alignment, as even minor spatial misalignments can disrupt block correspondence, limiting practical applicability in unconstrained environments.

Moreover, recent studies suggest that combining MMCF with Generative Adversarial Networks (GANs) [18] can help address some of these limitations, as GANs enable the direct synthesis of high-quality frontal face images, providing a robust alternative for pose normalization. However, the fundamental challenge remains: all fusion-based models (MCF, 2D-MCF, MMCF) continue to rely on the greedy search algorithm, which cannot guarantee globally optimal solutions consistently.

These observations highlight the need for alternative optimization algorithms, such as the genetic algorithm, capable of overcoming the inherent limitations of greedy search, while achieving optimal representations with reasonable computational complexity.

2.5 The Greedy Search Algorithm

The Greedy Search Algorithm [11] is a straightforward and intuitive heuristic strategy designed to solve optimization problems by making a series of locally optimal choices. The fundamental principle governing it is remarkably simple: at each decision-making step, the algorithm selects the option that offers the greatest immediate benefit, without considering the future consequences of that choice. It is greedy in the sense that it always pursues the best local gain, with the hope that this sequence of local gains will lead to a globally optimal solution [11].

Inherent Limitations of the Greedy Search:

Despite its simplicity and relative speed, this mechanism suffers from inherent limitations that render it ill-suited for complex combinatorial problems such as channel selection.

The most significant of these limitations include [11][12]:

- **Myopia:** The most critical flaw of the greedy algorithm is its myopia. A decision that appears optimal at a specific step may lead the search process down a path that prevents it from reaching the global optimum. For instance, there might be a color channel (Channel "B") that performs poorly on its own but, when combined with another channel (Channel "C"), produces a synergistic effect that far surpasses the performance of any other pair of channels. The greedy algorithm, by selecting the best-performing initial channel (Channel "A"), may never explore this hidden synergy between channels "B" and "C".
- **Sensitivity to the Starting Point:** The path taken by the algorithm is entirely dependent on its initial choice.
- **Inability to Backtrack:** Once the greedy algorithm makes a decision and adds an element to the solution, it cannot reverse that decision. If it later becomes apparent that this choice was detrimental in the long run, the algorithm is already locked into a sub-optimal path.

2.6 The Genetic Algorithm (GA)

To overcome the limitations of myopia and convergence to local optima inherent in the greedy search algorithm, this thesis aims to advance the prior color models for Face Recognition by replacing the greedy algorithm with a more capable search space exploration technique: the GA [12].

GA belongs to the class of evolutionary algorithms inspired by Charles Darwin's theory of natural evolution. They rely on mechanisms that simulate biological processes such as selection, crossover, and mutation to find optimal or near-optimal solutions [12][37].

A GA maintains a population of potential solutions and iteratively improves them, which facilitates the exploration of a broader solution space and results in a more efficient search process compared to the greedy search.

The algorithm begins by creating an initial population of potential solutions, which are then evaluated according to a fitness function that reflects the solution's quality for the given problem. Subsequently, the best solutions are selected to produce a new generation through crossover and mutation operations, ensuring the introduction of solution diversity and the exploration of new regions within the search space. This process continues over several generations, with the quality of the solutions gradually improving until an optimal

performance level is reached [37][38]. The GA is distinguished by its ability to achieve an effective balance between the exploration of new areas in the solution space and the exploitation of already discovered good solutions. This significantly reduces the probability of getting trapped in local optima, a common problem in greedy search and certain other metaheuristic algorithms, such as Particle Swarm Optimization (PSO).

From the perspective of comparison with the exhaustive search category, methods such as brute-force search may guarantee finding the optimal solution [39], however, despite their accuracy, these methods are often impractical for large-scale problems due to their high time complexity, which could require years to train the models. In contrast, a GA offers an effective balance between the pursuit of optimal or near-optimal solutions and the reduction of training time. This makes them an ideal alternative to both exhaustive and greedy search algorithms, with the potential to achieve improved outcomes within a reasonable timeframe.

Consequently, the ability of the GA to balance solution quality with temporal efficiency makes it a promising and ideal replacement for the greedy algorithm in selecting the optimal combination of color channels for Face Recognition models.

The following chapter presents a detailed description of the proposed methodology designed to address this challenge.

2.7 Summary

This chapter provided a comprehensive review of Face Recognition literature, from classical subspace methods PCA, LDA, SRC to advanced multi-channel, multi-resolution fusion models MMCF, MCF, 2D-MCF. It highlighted the critical role of color information and revealed a shared limitation: reliance on the greedy search algorithm, prone to myopia and local optima. In addition, this chapter also discussed both the Greedy Search Algorithm and the GA, highlighting their differences and the potential of GA as a more effective alternative.

Chapter 3

Research Methodology

3.1 Introduction

Following the literature review in the previous chapter, it has been established that advanced multi-channel, multi-resolution Face Recognition models, such as MMCF, remain fundamentally constrained due to their reliance on the greedy search algorithm for selecting the optimal color channel combination. This algorithm is locally optimal, focusing on immediate gains without considering the broader context or potential future alternatives. Accordingly, this chapter presents a detailed methodology aimed to address this research gap.

Transitioning from theoretical critique to practical implementation, this study operationalizes the GA as a robust and effective optimization strategy. The main objective is to construct a distinctive feature representation for selecting the optimal color channel combination and to validate its effectiveness by replacing the suboptimal greedy search with the more capable and adaptive GA.

This chapter systematically details the construction of the baseline model (original MMCF) and the proposed GA-enhanced models [10][12], while also outlining the experimental framework, including the design of genetic operators, feature extraction techniques, and evaluation protocols employed to rigorously and objectively test the central hypothesis.

3.2 Overall Framework

A systematic experimental framework was designed to achieve the study objectives and compare the performance of all models. The original MMCF model was adopted as the baseline, using the channel combination identified through the Greedy Search algorithm with the SRC classifier. Two GA-based models were then built: MMCF-GA-1, which applies GA with the SRC classifier for direct comparison with the baseline, and MMCF-GA-2, which applies GA with the LDA classifier to demonstrate the flexibility and adaptability of the proposed methodology across different classifiers.

3.3 Defining the Baseline: The Original MMCF Model

To ensure a fair and scientifically valid comparison, the first step in the experimental framework is to establish a robust baseline. For this purpose, the original MMCF model, as originally defined in the foundational work [10], is adopted. The construction of this model, as documented in the original research, was carried out using a specific training

and evaluation protocol. The original model was constructed using a large training dataset derived from the AR [40] and Curtin [41] databases. For more details on these databases, please refer to Section [4.2]. This dataset comprised 912 frontal images from 152 subjects. For the internal evaluation during the feature selection process, the data was partitioned such that for each subject, two images were used as gallery and four as test images, with the SRC classifier [15] serving as the evaluation tool.

The construction of the MMCF baseline model consists of the following stages:

1. **Color Space Transformation:** Each input RGB image is transformed into nine distinct color spaces: RGB, XYZ, YIQ, YUV, I1I2I3, nRGB, nXYZ, HSV, and YcbCr.
2. **Channel Reduction:** To minimize redundancy, identical or highly correlated color channels are removed, reducing the feature pool to 22 unique color channels.
3. **Multi-Resolution Generation:** Each of the 22 channels is resized into four distinct resolution levels (32×32 , 24×24 , 16×16 , and 12×12), producing a candidate set of 88 unique channel–resolution combinations.
4. **Initial Channel Selection and Redundancy Removal:** Each candidate channel–resolution combination is evaluated individually for its Face Recognition performance. The combination yielding the highest recognition accuracy is retained as the initial component of the MMCF model. All other resolutions of the selected channel are subsequently removed from the candidate pool to avoid redundancy.
5. **Progressive Channel Addition:** The remaining candidate channels are iteratively combined with the already selected set. In each round, the combination that achieves the highest recognition rate is incorporated into the model.
6. **Stopping Criterion:** The iterative process continues until adding further channels no longer improves recognition accuracy or leads to performance degradation. At this point, the construction process is terminated, and the final MMCF representation is established.
7. **Adopted Baseline Representation:** The outcome of this structured search is a stable representation consisting of nine selected channel–resolution combinations. This exact nine-channel representation, detailed in Table 1, is adopted as the official baseline model for this study. It provides the reference framework against which the enhancements proposed in this research are systematically evaluated.

Table 1: The Official MMCF Representation Adopted as the Baseline

Color Channel	V	U	X	I3	B	I	Q	nG	nB
Color Space	YUV	YUV	XYZ	I1I2I3	RGB	YIQ	YIQ	nRGB	nRGB
Resolution	24x24	32x32	12x12	32x32	12x12	12x12	32x32	24x24	32x32

3.4 Proposed Model: Multi Resolution Multiple Color Fusion using Genetic Algorithm (MMCF-GA).

Building upon the limitations identified in the MMCF baseline most notably its reliance on Greedy Search for channel/ resolution selection. This research introduces a GA as a more robust optimization strategy. The design and implementation of the proposed GA framework including solution representation, fitness evaluation, and parameter configurations are elaborated in the following subsections.

3.5 Dataset and Preprocessing Consistency

To ensure a fair and direct comparison with the MMCF baseline, this research adopts the same dataset, partitioning strategy, and evaluation criteria as reported in the original implementation. The training dataset comprises frontal face images sourced from the AR [40] and Curtin [41] databases, totaling 912 images from 152 subjects. For each subject, two images are assigned to the gallery set and four images to the probe set, with the SRC [15] employed as the evaluation tool.

The preprocessing procedure applied in the baseline is also replicated here to preserve methodological consistency. Specifically, all images were transformed from the original RGB color space into nine alternative color spaces: RGB, XYZ, YIQ, YUV, I1I2I3, nRGB, nXYZ, HSV, and YcbCr. Following this transformation, identical and highly correlated channels were eliminated, yielding a refined set of 22 distinct channels.

Each channel was subsequently resized into four resolutions (32×32, 24×24, 16×16, and 12×12), producing 88 unique channel-resolution combinations. These combinations form the standardized input space for feature selection and optimization through the proposed GA framework.

3.5.1 GA Representation Framework.

In the proposed GA framework, each potential solution is encoded as a fixed-length binary chromosome. The problem addressed in this research is the selection of an optimal

subset from the 88 available channel-resolution combinations generated during the pre-processing stage.

Formally, an individual chromosome C is defined as a binary vector of length $L = 88$:

$$C = [g_1, g_2, \dots, g_L] \quad (3.1)$$

Let S be the set of indices of the selected combinations. The value of each gene g_i is then

$$\text{defined as: } g_i = \begin{cases} 1 & \text{if } i \in S \\ 0 & \text{if } i \notin S \end{cases} \quad (3.2)$$

The population P is composed of $N = 100$ individual chromosomes, $P = \{C_1, C_2, \dots, C_n\}$.

This binary encoding allows the GA to efficiently explore the combinatorial search space while maintaining a direct mapping to the feature subsets.

- **Gene:** Each gene corresponds to a specific channel–resolution pair and is represented in binary form, taking the value '1' if the pair is selected and '0' if it is not.
- **Chromosome:** The chromosome is a binary string consisting of 88 bits, where each bit (gene) represents one of the available (channel–resolution) combinations. A value of '1' indicates that the gene is active (selected), whereas a value of '0' indicates that it is inactive (not selected). Thus, each chromosome represents a complete “recipe” for constructing a candidate model. In the adopted configuration, the number of active genes (those with a value of '1') within each chromosome varies according to the random selection process, but it does not exceed 22 genes, which represents the maximum number of (channel–resolution) combinations allowed in each chromosome to form the proposed model in each evolutionary generation. Table 2 illustrates an example of chromosome encoding consisting of 6 genes, where each gene represents a specific (channel–resolution) combination, and the binary values (1/0) indicate whether the corresponding feature is selected or not

Table 2: Example of Chromosome Encoding with 6 Genes

Gene Index	Represents	Genes
1	R 32×32	1
2	B 24×24	1
3	G 32×32	0
4	X 16×16	1
5	Q 24×24	1
6	H 16×16	0

- **Population:** The population is a diverse set of chromosomes evaluated and evolved over successive generations. This diversity enables the GA to systematically explore the search space, balancing the exploitation of promising solutions with the exploration of new possibilities.

This representation ensures that the GA can efficiently navigate the large combinatorial search space, while maintaining a clear and consistent mapping between the binary encoding and the actual feature subsets used in the Face Recognition model.

3.5.2 Initial Population and Redundancy Repair Mechanism

The GA begins by creating an initial population of 100 candidate solutions (chromosomes). The choice of a population size of 100 is deliberate and serves as a balance between two competing requirements:

1. **Genetic Diversity:** A sufficiently large population ensures a wide coverage of the solution space at the start of the evolutionary process. With 100 chromosomes, the GA can explore a broad range of channel–resolution combinations, reducing the risk of premature convergence caused by insufficient diversity.
2. **Computational Efficiency:** While larger populations increase diversity, they also raise the computational burden since each chromosome must be evaluated in every generation. A size of 100 provides adequate exploration without incurring prohibitive computational costs, especially given that each chromosome must be evaluated using the SRC classifier over the dataset.

This balance makes 100 a practical and effective population size for the scale of the current problem, which involves optimizing over 88 possible channel-resolution combinations.

However, a purely random generation of chromosomes can produce invalid solutions, such as selecting multiple resolutions of the same channel simultaneously (e.g., the R channel). To address this, a Redundancy Repair Mechanism is applied immediately after chromosome initialization. The rule is straightforward: for any given channel (e.g., the R channel), only the gene corresponding to the largest selected resolution is allowed to remain ‘1’; all other genes associated with the same channel are forced to ‘0’.

The decision to retain the highest resolution is based on the hypothesis that higher-resolution channels contain more detailed and discriminative information, which is generally more valuable for Face Recognition tasks, whereas lower resolutions provide

less detailed information. Therefore, prioritizing the highest available resolution ensures that no critical information is lost, aligning with the objective of this study to improve image representation.

This ensures that each chromosome represents a logically consistent and non-redundant feature subset.

This process can be formally described. Let C' be a raw, potentially invalid chromosome. The repair mechanism, $Repair(C')$, produces a valid chromosome C where for each gene g_i :

$$g_i = \begin{cases} 1, & \text{if } i = \underset{j \in S_k, g'_j=1}{argmax} Resolution(j) \\ 0, & \text{otherwise} \end{cases} \quad (3.3)$$

S_k : This represents the set of genes that all belong to the same color channel. For example, if the chromosome has genes for the R-channel at four different resolutions (32x32, 24x24, 16x16, and 12x12), then S_k for the R-channel would be the set of indices corresponding to these four genes. Its purpose is to group all variations of a single channel together for evaluation.

Resolution(j): This is a helper function that returns a simple number representing the size of the resolution at a given gene index j . For instance, for a gene representing a 32x32 resolution, $Resolution(j)$ would return 1024 (i.e., 32 x 32). This allows the algorithm to easily compare different resolutions and identify which one is the largest.

This ensures that each chromosome represents a logically consistent and non-redundant feature subset.

Redundancy Repair Example: The repair process can be illustrated with a simple example using the genes corresponding to the R-channel at four resolutions (32x32, 24x24, 16x16, and 12x12):

Step1: Raw Chromosome Generation

A randomly generated chromosome may assign multiple '1's to the same channel, as shown above:

Gene Value	1	1	0	1
Represents	R32x32	R24x24	R16x16	R12x12

This chromosome is considered invalid because it selects more than one resolution for the same channel.

Step 2: Redundancy Detection

The repair mechanism detects multiple active genes '1' for the R-channel. According to the rule, only the largest selected resolution (here, R 32×32) is allowed to remain active.

Step 3: Chromosome Repair

After detecting the redundancy, the chromosome is corrected as follows:

Gene Value	1	0	0	0
Represents	R32x32	R24x24	R16x16	R12x12

Thus, only the largest resolution for the R-channel remains selected, and redundancy is eliminated. This two-step procedure (random generation followed by redundancy repair) is repeated until 100 valid chromosomes are created, forming the initial population (Generation 0).

3.5.3 Fitness Function

The fitness function is a core component of the GA, designed to quantify the performance, or quality, of each candidate solution (chromosome). In this research, the fitness of a chromosome is directly defined as the Face Recognition accuracy achieved by the specific channel–resolution combination it represents.

The fitness score for each chromosome is determined in a two-step process:

1. Feature Vector Construction: Based on the active genes '1' in the chromosome, the corresponding channel-resolution combinations are identified. For each image in the dataset, the selected channels are extracted and concatenated to form a single feature vector.
2. Fitness Evaluation: The constructed feature vectors are used to train and test for Face Recognition using the SRC classifier. The resulting recognition accuracy is then assigned as the chromosome's fitness score.

3.5.4 Evolutionary Operators

Following the fitness evaluation of the current population, a set of evolutionary operators is applied to generate a new population for the subsequent generation. These operator elitism, selection, crossover, and mutation are the core mechanisms that drive the algorithm search toward optimal solutions.

1. Elitism

The principle of elitism is implemented to preserve the best-performing solutions discovered so far. Specifically, the top 2% of the population, corresponding to the two fittest chromosomes, are identified and copied directly and unaltered into the next generation. This strategy guarantees that the highest fitness achieved by the GA is monotonically non-decreasing, preventing the loss of valuable genes or channels and accelerating convergence toward a high-quality solution.

2. Parent Selection

To generate the 98 individuals(chromosomes) required for the new population, parents are selected from the current generation using Roulette Wheel Selection, a fitness-proportionate selection method. This technique ensures that chromosomes with higher fitness defined in this research as higher recognition accuracy have a greater probability of being chosen to create the next generation (via crossover). While this approach favors stronger solutions(chromosomes), it does not entirely discard weaker ones, thereby maintaining crucial genetic diversity within the population.

The probability, $P(C_i)$, of selecting a specific chromosome C_i from the population P (of size N) is calculated as the ratio of its fitness, $f(C_i)$, to the total fitness of the entire population:

$$P(C_i) = \frac{f(C_i)}{\sum_{j=1}^N f(C_j)} \quad (3.4)$$

To illustrate this process, consider a simplified population of six chromosomes with the following recognition accuracies (fitness values):

- Chromosome A: $f(C_A) = 90\%$
- Chromosome B: $f(C_B) = 75\%$
- Chromosome C: $f(C_C) = 60\%$
- Chromosome D: $f(C_D) = 40\%$
- Chromosome E: $f(C_E) = 30\%$
- Chromosome F: $f(C_F) = 5\%$

The total fitness of this population is $(90 + 75 + 60 + 40 + 30 + 5) = 300$.

The selection probability for each chromosome is then calculated as follows:

- $P(C_A) = 90 / 300 = 30\%$
- $P(C_B) = 75 / 300 = 25\%$

- $P(C_C) = 60 / 300 = 20\%$
- $P(C_D) = 40 / 300 \approx 13.3\%$
- $P(C_E) = 30 / 300 = 10\%$
- $P(C_F) = 5 / 300 \approx 1.7\%$

This selection process is repeated 98 times to choose the parents (chromosomes) required to generate the offspring (chromosomes) for the next generation. The process operates with replacement, meaning that a single, highly-fit chromosome (such as Chromosome A) can be selected multiple times, while less fit chromosomes (such as Chromosome F) still retain a small, non-zero chance of being selected. This mechanism effectively balances the exploitation of high-performing solutions (i.e., chromosomes with high recognition accuracy) with the exploration of the broader search space.

3. Crossover

The selected parents are randomly paired to produce offspring via a Single-Point Crossover operator. For each pair of parents, a single random crossover point is selected along the chromosome's length. The genes beyond this point are then swapped between the two parents to create two new offspring, as illustrated in Figure 1. Crucially, since the crossover operation can result in infeasible offspring (e.g., by combining genes to select multiple resolutions for the same channel), the redundancy repair mechanism (previously described in Section 3.7.2) is applied to each new offspring. This step ensures that all generated individuals are valid and logically consistent before they enter the new population.

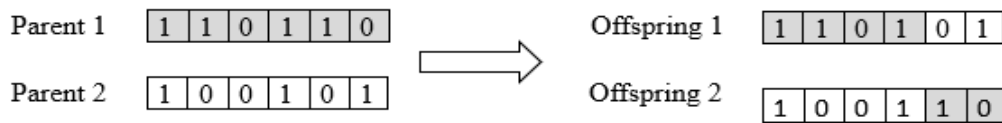


Figure 1: Illustrative Example of Single-Point Crossover Using 6 Genes (Crossover After Gene 4).

The figure 1 provides an illustrative example of the crossover process using two simplified chromosomes, each composed of 6 genes. A single crossover point is randomly selected after the fourth gene, where Parent 1 and Parent 2 exchange the segments that follow this point. As a result, Offspring 1 is formed by combining the first part of Parent 1 with the second part of Parent 2. Conversely, Offspring 2 is created by taking the first part of Parent 2 and the second part of Parent 1. This mechanism allows for the combination of different feature subsets, enabling the exploration of new potential solutions.

4. Mutation

To maintain genetic diversity and prevent the search from stagnating at a local optimum, a bit-flip mutation operator is applied. Each gene in every new offspring has a low probability (1%) of being flipped from '0' to '1' or vice versa. A mutation rate of 1% was selected as it provides an optimal balance, it is high enough to prevent premature convergence to a local optimum by maintaining genetic diversity. A rate lower than 1% would risk the algorithm stagnating at suboptimal solutions, whereas a rate higher than 1% could lead to the loss of valuable information across generations and disrupt the algorithm's capacity for gradual improvement. This allows the algorithm to explore new feature combinations by occasionally reactivating a previously discarded channel-resolution pair, or deactivating a selected one. Following mutation, the redundancy repair mechanism is applied again to ensure the validity of all chromosomes before they form the new generation.

5. Generational Cycle

With the formation of a completely new population, one generational cycle concludes. The algorithm then iterates, taking this new generation as its input and repeating the entire process, starting again with the fitness evaluation step. This iterative loop of evaluation, selection, crossover, and mutation continues, progressively refining the population toward better solutions.

6. Termination Criterion and Final Model Construction

The evolutionary process is terminated based on a convergence criterion. Specifically, the algorithm concludes its run when the fitness of the best-performing chromosome shows no improvement for 50 consecutive generations. This threshold is chosen in line with standard practice in GA research [42], where a fixed number of stagnating generations is widely used as a practical convergence indicator. Moreover, preliminary experiments in this study confirmed that 50 generations provided a robust balance: smaller values (e.g., 20 or 30) occasionally led to premature termination, while larger values (e.g., 100) unnecessarily increased computation time without significant gains in accuracy. At this point, the process yields the final output of the entire evolutionary search: the single best chromosome identified across all generations. This chromosome, which encodes the most effective combination of multi-channel and multi-resolution features, is then adopted to construct the final proposed model, referred to as MMCF-GA.

In the following figure 2. The diagram illustrates the main stages, including initialization, fitness evaluation, the evolutionary cycle (selection, crossover, mutation), the iterative application of the redundancy repair mechanism, and the convergence-based termination criterion.

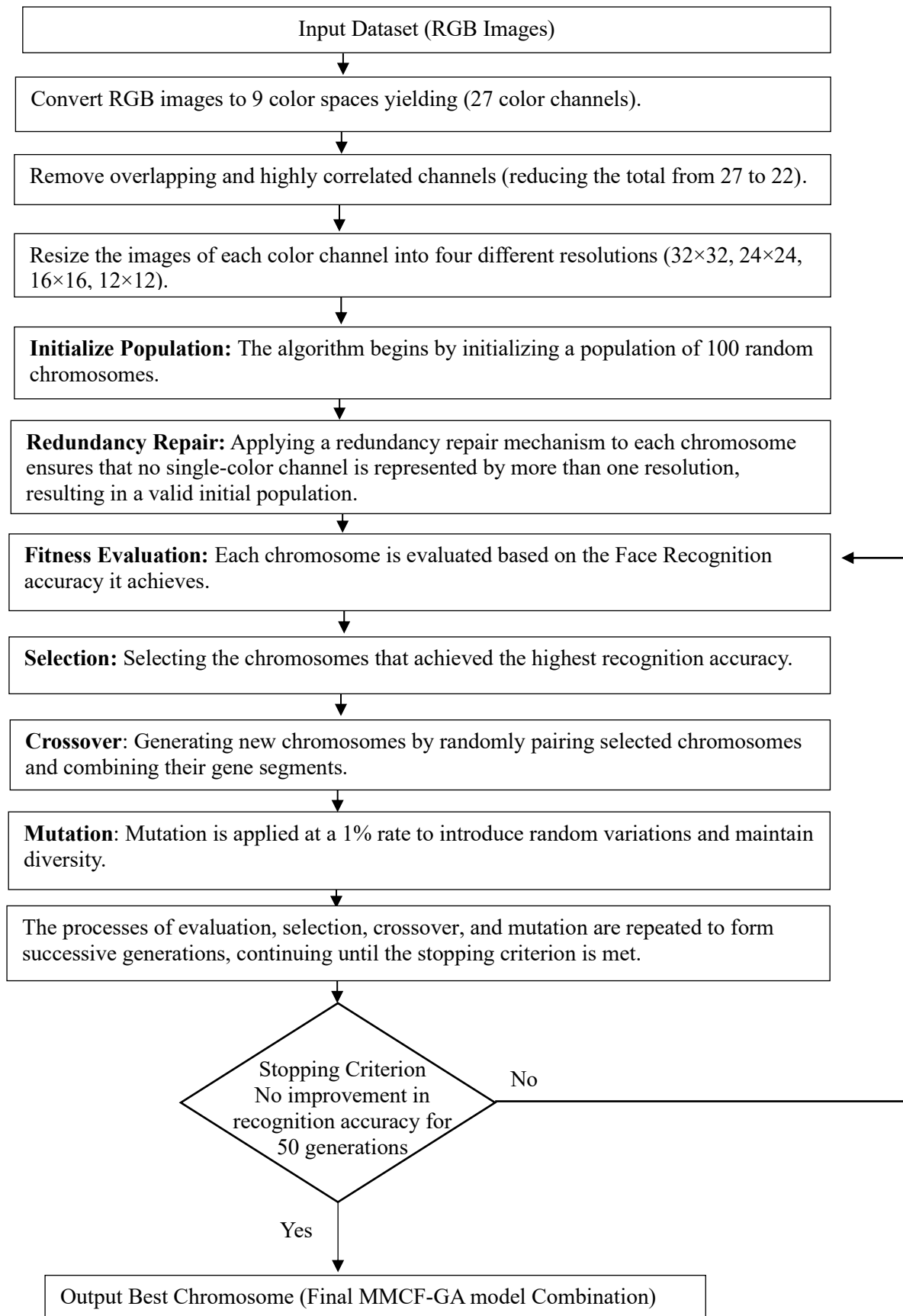


Figure 2: MMCF-GA model training using GA methodology.

3.6 Construction of the MMCF-GA-1 Model

The application of the GA with the SRC classifier, as detailed in the preceding sections, culminated in the construction of a novel feature representation. This final representation, which constitutes the core of the MMCF-GA-1 model, is presented in Table 3. The optimal feature set comprises seven distinct channel-resolution combinations that collectively yielded the highest recognition accuracy during the evolutionary search. This empirically derived representation serves as the primary proposed solution in this research, against which the baseline MMCF model will be compared.

Table 3: The optimal feature representation for the MMCF-GA-1 Model using SRC classifier

Color Channel	V	V	S	nY	nG	nB	I3
Color Space	HSV	YUV	HSV	nXYZ	nRGB	nRGB	I1I2I3
Resolution	24x24	32x32	32x32	24x24	32x32	16x16	16x16

The table shows the seven selected channel-resolution combinations identified by the GA when trained on the combined AR and Curtin databases with the SRC classifier.

3.7 Construction of the MMCF-GA-2 Model

To further validate the robustness and classifier-agnostic nature of the proposed GA-based methodology, a second model, designated MMCF-GA-2, was constructed. The purpose of this experiment is to demonstrate that the GA framework is not merely overfitted to a single classifier but can adapt and construct new optimal feature representations when the underlying classification tool is changed.

The construction of MMCF-GA-2 involved repeating the entire evolutionary process under identical conditions using the same dataset, GA parameters, and termination criterion with one critical modification: the LDA classifier is used as the fitness evaluation tool. The successful generation of a second, unique feature representation, as presented in Table 4, provides strong evidence for the adaptability of the proposed GA framework. A comparison between the SRC-optimized representation Table 3 and the LDA-optimized representation Table 4 reveals that the algorithm does not converge on the same solution. Instead, it discovered a different combination of channels and resolutions, indicating that the GA is insensitive to the change in the classifier and effectively explored the search space to find a new optimum.

Table 4: The optimal feature representation for the MMCF-GA-2 Model using LDA classifier

Color Channel	X	S	U	nG	nY	CB	I	CR
Color Space	XYZ	HSV	YUV	nRGB	nXYZ	YCBCR	YIQ	YCBCR
Resolution	32x32	32x32	32x32	24x24	16x16	16x16	12x12	24x24

This 8-channel combination is identified by the GA when the fitness evaluation is performed using the LDA classifier.

Upon the completion of the training phase, the outcome is a trained and optimized Genetic Model (either MMCF-GA-1 or MMCF-GA-2), which encapsulates the optimal feature representation for Face Recognition. This model can subsequently be employed in various Face Recognition tasks.

It is important to emphasize that the computationally demanding training process comprising the full GA run is executed only once. Thereafter, the resulting trained model can be utilized efficiently and repeatedly across multiple applications, eliminating the need for retraining.

3.8 Handling Pose Variations: Integration with Stacked Progressive Auto Encoder (SPAЕ)

Based on the methodology described in [10], which successfully integrated the MMCF model with the SPAE [16] to address pose variations in Face Recognition, this study adopts the same pipeline. The methodology is extended by applying it on our genetically optimized models, MMCF-GA-1 and MMCF-GA-2, thereby ensuring a fair comparison with the baseline MMCF model while simultaneously evaluating the robustness of the proposed models.

3.8.1 The Principle of Stacked Progressive Auto Encoder (SPAЕ)

SPAЕ is a deep learning framework proposed by [17]. Its core principle is to correct the face images from pose angles to frontal face progressively rather than in a single, complex step. The framework relies on a series of autoencoders stacked in a hierarchy, where each autoencoder is trained to reduce the pose angle by a small, fixed amount (e.g., from $\pm 45^\circ$ to $\pm 30^\circ$, then from $\pm 30^\circ$ to $\pm 15^\circ$, and finally to 0°). This gradual, stepwise transformation has been proven to be more robust and effective than attempting a direct correction of large angles, enabling the generation of high-quality frontal face images.

3.8.2 Training SPAE Models for Color Channels

To achieve better performance than the MMCF model, we use our proposed genetic models (MMCF-GA-1 and MMCF-GA-2) with the SPAE algorithm for pose correction. The training process is as follows:

1. **Data Preparation:** Given an RGB training dataset, all images are resized to 33×33 pixels. Each image is then converted into the specific color channels defined by the selected genetic model (either MMCF-GA-1 as per Table 3, or MMCF-GA-2 as per Table 4).
2. **Network Training:** A separate SPAE network is trained for each color channel individually, using an image size of 33×33 pixels. This results in seven trained SPAE models for the first genetic model and eight trained SPAE models for the second, with each model corresponding to a specific color channel.
3. **Training Dataset:**
A dataset from the Multi-Pie database [43] was employed for training. It consists of 200 subjects, each with 37 images captured at various pose angles, including (0° , $\pm 15^\circ$, $\pm 30^\circ$, $\pm 45^\circ$, $\pm 60^\circ$, $\pm 75^\circ$, and $\pm 90^\circ$).

3.8.3 Pose Correction and Face Recognition

The process of pose correction and Face Recognition using the trained models follows these steps:

1. **Image Preprocessing:** For gallery and query images, all images are converted to the color channels of the selected genetic model. For fair comparison, the image size is set to 33×33 pixels, consistent with the original SPAE work.
2. **Pose Correction:** Each color channel within every face image is corrected to a frontal view using its corresponding trained SPAE network.
3. **Face Recognition:** Each corrected image is transformed into its respective genetic model representation. Subsequently, Face Recognition is performed using a LDA classifier.

The entire process is illustrated in Figure 3 which shows the SPAE pipeline for face pose correction with the genetic model.

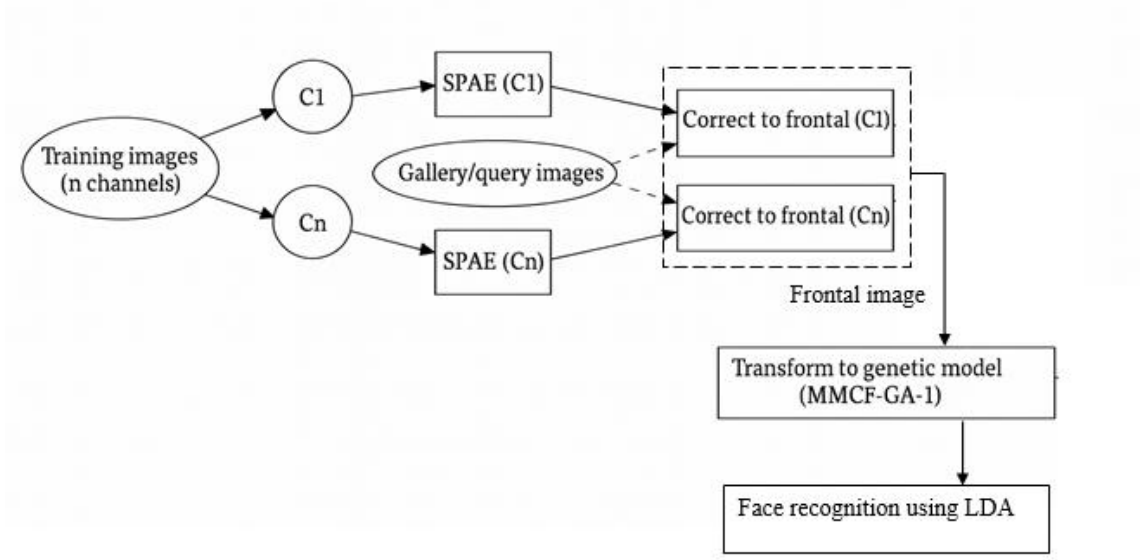


Figure 3: Face pose correction using SPAE and the genetic model.

3.9 Utilizing Advanced Pose Correction Methods for Dataset Creation

Following the application of our pose correction methodology using SPAE and our genetic models, we aim to evaluate the robustness of our models by leveraging state-of-the-art advancements in face frontalization. In this context, the Random Mask Attention Generative Adversarial Network (RMAGAN) [44]. Introduces an advanced, self-supervised GAN framework known as RMAGAN. This framework reframes the pose correction task as an image completion problem, generating high-fidelity and realistic frontal-view images from varied pose angles without the need for paired training data [44]. Accordingly, to evaluate our genetic models on data corrected by state-of-the-art techniques, we will utilize the dataset generated by the RMAGAN methodology [44] on the Labeled Faces in the Wild (LFW) dataset [45] as input for our models.

Specifically, we will take the LFW face images that have been corrected and frontalized by the RMAGAN framework [44] and apply our proposed genetic models MMCF-GA-1 and MMCF-GA-2 to perform Face Recognition. This approach allows us evaluate the robustness and performance of our models on data generated by a state-of-the-art generative algorithm. This investigation directly supports the primary objective of this research: to verify that our proposed genetic models can augment other modern techniques, aiming to improve overall image representation and recognition accuracy.

Moreover, our methodology was not confined solely to addressing the problem of pose

correction. The same frontalized images produced by SPAE and RMAGAN were further utilized as primary inputs to conduct a series of critical experiments under the SSPP scenario. This scenario represents one of the most challenging tasks in Face Recognition, where the system is trained using only a single image per subject. This extension directly supports the primary objective of this research: to verify that our proposed genetic models can augment other modern techniques and construct a robust representation even from minimal information, thereby aiming to improve overall image representation and recognition accuracy.

3.10 Software and Hardware

Software and hardware used for training the models: Computer with Intel Quad-Core i7-8550U CPU, 16GB RAM, running Windows 11 and using MATLAB 2018.

3.11 Summary

This chapter introduced the methodological framework of the study. Building upon the limitations of the original MMCF baseline, a GA was employed to construct enhanced models capable of identifying optimal feature representations. Two GA-based variants, MMCF-GA-1 and MMCF-GA-2, were developed to demonstrate both improved performance and methodological flexibility. Moreover, SPAE and advanced pose-correction technique, the RMAGAN were integrated to handle pose variations and enhance the robustness of the proposed models.

Moreover, the single gallery image problem SSPP is addressed by utilizing enhanced datasets generated through SPAE and RMAGAN to evaluate the models under limited-data conditions.

This foundation sets the stage for the experimental evaluation presented in Chapter 4.

Chapter 4

Experimental Setup and Results

4.1 Introduction

This chapter presents the experimental evaluation of the MMCF-GA-1, MMCF-GA-2 models developed in this thesis. The primary objective is to empirically test the hypothesis that a GA outperforms the Greedy Search for Face Recognition feature representation. In order to rigorously test the models' performance and generalization capabilities under a wide range of challenging real-world conditions, this evaluation is carried out across eight distinct face databases. These challenges include extreme expressions, facial misalignment, uncontrolled environments with significant pose changes, and the difficult SSPP problem. The chapter details the experimental setup, the characteristics of each database, and the evaluation metrics used. Furthermore, it presents a comparative analysis of the proposed models against the original MMCF baseline, thereby quantifying the performance gains achieved through the GA-based optimization.

4.2. Databases Setup

The performance of the proposed genetic models was rigorously evaluated on eight standard face databases to cover a diverse range of acquisition conditions and challenges. These databases include the AR Face Database [40], the Curtin University Face Database [41], the Multi-Pie database [43], the LFW database [45], the Georgia Tech (GT) Face Database [46], the Color FERET database [47], the Psychological Image Collection at Stirling (PICS) [48], and the Face Recognition Grand Challenge (FRGC) [49]. While all images in the databases are in RGB color, they vary significantly in quality, resolution, and the presence of real-world challenges such as facial misalignment, pose, and expression variations. The following subsections provide a detailed description of each database and the specific experimental protocols applied to it.

4.2.1. AR database

The AR face database is a 2D database containing over 2600 colored facial images belong 100 subjects (50 male, 50 female). The images were captured in two sessions, six months apart. Each subject has 10 frontal face images showcasing variations in facial expressions, lighting conditions, and occlusions. Figure 4 shows some sample images of one subject from this database [40].



Figure 4: Displays sample images from the AR database

4.2.2. The Curtin Database

The Curtin database is a 3D color face database containing 5,044 3D face images of 52 subjects (97 images per subject). Furthermore, this database covers varying facial expressions, lighting conditions, pose variations, and occlusion. It was originally designed for 3D Face Recognition. All images were taken in a controlled environment using two cameras, including a low-resolution 3D camera. A variety of challenging condition are covered in this database including: facial expressions, occlusions, pose variations, and illumination and occlusions [42]. Figure 5 shows some sample images of one subject from this database [41].



Figure 5: Displays sample images from the Curtin database.

4.2.3. Multi-Pie Database

The Multi-Pie face database is a foundational benchmark in Face Recognition research, comprising over 750,000 images of 337 subjects. It is distinguished by its comprehensive and systematic variation in pose, captured by 13 head-level cameras at the following angles: 0° , $\pm 15^\circ$, $\pm 30^\circ$, $\pm 45^\circ$, $\pm 60^\circ$, $\pm 75^\circ$, and $\pm 90^\circ$. The database provides a wide range of capture angles, making it highly suitable for evaluating model performance under various challenging conditions [43]. Figure 6 shows some sample images of one subject from this database.



Figure 6: Displays sample images from the Multi-Pie face database.

4.2.4. Labeled Faces in the Wild Face Database

The LFW face database [45] comprises images collected from media articles and websites. It is considered the most challenging database for studying the problem of unconstrained facial recognition due to its coverage of diverse environments and variations in posture, expression, skin color, lighting, facial attire, hair color and style,

image quality, and other factors. The dataset includes 13,233 images representing 5,749 distinct identities. Figure 7 shows some sample images of one subject from this database.



Figure 7: Displays sample images from the LFW face database.

4.2.5. Georgia Tech Face Database

The GT dataset comprises 750 color images, representing 50 individuals with 15 images per person. Captured in the Georgia Tech lab, these facial images showcase a variety of lighting conditions and facial expressions. The database contains frontal images in both upright and tilted orientations [46]. Figure 8 shows some sample images of one subject from this database.



Figure 8: Displays sample images from the GT Face Database.

4.2.6. The Facial Recognition Technology database

The FERET database was assembled at George Mason University and the US Army Research Laboratory as a component of the FERET program [47]. It was established to enhance the development of methods and technologies for automated facial recognition. The database contains roughly 11,338 color images taken in a semi-controlled setting, featuring a variety of poses from almost 994 individuals. Figure 9 shows sample images of one subject from this database [47].



Figure 9: Sample images from the FERET face database.

4.2.7. Psychological Image Collection face databaset

The Aberdeen database is part of the PICs face database. It comprises 261 images across 29 subjects, with each subject represented by 9 images that include minor lighting conditions and some variations in pose [48]. Figure 10 shows some sample images of one subject from this database.



Figure 10: Sample images from the PICs face database

4.2.8. The Face Recognition Grand Challenge Database

The FRGC database, contains over 50,000 facial images taken in both controlled and uncontrolled settings [49]. It is considered as one of the most challenging databases within the Face Recognition research community, comprising full frontal facial images captured under various lighting conditions and with different facial expressions smiling and neutral [49]. Figure 11 shows sample images of one subject from this database.



Figure 11: Displays sample images from the FRGC face database.

4.3. Experimental Evaluation and Results

To evaluate the efficacy of the MMCF-GA-1 and MMCF-GA-2 models, a series of experiments was conducted across the eight distinct face databases is summarized in Table 5. These experiments were designed to test the models' performance under the various challenging conditions outlined in the table. The evaluation was performed using two different classifiers to ensure a comprehensive analysis: first with the LDA classifier, followed by a second set of experiments using the SRC. The subsequent sections present a comparative analysis of the proposed models against the original MMCF baseline, with detailed results for each classifier and database.

Table 5: Databases Used for Face Recognition with MMCF-GA-1 and MMCF-GA-2 models

Database	Images	Subject	Images Subject	Conditions	Gallery images	Query images	Times
AR+ CURTIN	1216	152	8	Extreme expressions/ Open mouth	6	2	1
GT	750	50	15	Expressions/ misalignment	7	8	5
LFW	850	85	10	Uncontrolled / Pose / expressions	9	1	10
FERET	1614	807	2	Expressions / single gallery image	1	1	1
PICS	261	29	9	Expressions /Pose	4	5	5
FRGC	7992	222	36	expressions/ uncontrolled	8	28	5

4.4 Experimental Results Using the LDA Classifier

To begin the evaluation, a first series of experiments was conducted using the LDA classifier. This was done to compare the performance of the MMCF-GA-1 and MMCF-GA-2 models in comparison with the model MMCF across the databases detailed in Table 5. This allowed for a comprehensive analysis of the models' effectiveness under various conditions, and the following sections will present the detailed results of these experiments.

4.4.1 Experiment One

For the purpose of this specific experiment, a subset of images was strategically selected from both the AR and Curtin databases to form a consolidated dataset. While the original databases contain a broader range of variations as described above, this study utilizes a total of 1,216 images from 152 subjects (100 from AR and 52 from Curtin). Specifically, eight images per subject were selected to focus on facial expression variations: two neutral, four with standard expressions, and two with extreme expressions. This selection ensures a balanced evaluation across all subjects. As detailed in Table 5, the first six images constitute the gallery set, while the remaining two serve as the probe query images.

This experiment's result highlights the advantage of MMCF-GA-1 and MMCF-GA-2 model over MMCF model. As shown in Table 6, the MMCF-GA-1 Model achieved a recognition accuracy of 99.01%, while the MMCF-GA-2 Model achieved a recognition accuracy of 99.34%, representing a significant improvement of 0.66% to 1.00% over the MMCF model's accuracy of 98.35%.

Table 6: Recognition Accuracy Applied on AR+CURTIN Database and LDA Classifier.

Model Name	Recognition Rate
MMCF	98.35%
MMCF-GA-1	99.01%
MMCF-GA-2	99.34%

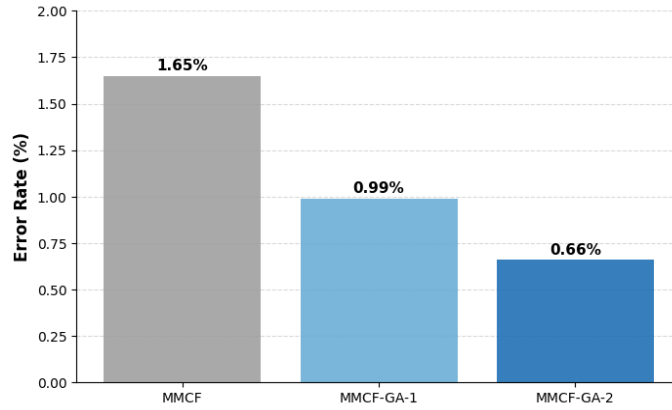


Figure 12: Comparison of Error Rates Among Face Recognition Models on the AR+CURTIN Dataset and LDA classifier.

As shown in Figure 12, the MMCF-GA-1 and MMCF-GA-2 models achieved markedly lower error rates on the AR+CURTIN dataset, reducing misclassification by about 40% and 60%, respectively, compared to the MMCF baseline.

4.4.2 Experiment Two

This experiment evaluated the performance of the MMCF-GA-1 and MMCF-GA-2 models using the GT face database described in Section 4.2.5. As noted, this database features frontal and tilted faces with variations in expression and lighting. We utilized the first seven images from each subject as a gallery set and the remaining eight images as query images. The Face Recognition process was repeated five times, with the results presented in Table 5.

The recognition accuracies on the GT database are presented in Table 7. The results clearly show that both proposed models significantly outperformed the MMCF baseline. The MMCF-GA-1 model achieved an accuracy of 92.25%, marking a substantial improvement of 7.1% over the MMCF baseline's 85.15%. Similarly, the MMCF-GA-2 model reached 91.25% accuracy, with improvement of 6.1% over the MMCF. These findings highlight the ability of the GA-optimized models to better handle the misalignment and expression challenges present in this dataset.

Table 7: Recognition Accuracy Applied the GT Database and LDA Classifier.

Model Name	Recognition Rate
MMCF	85.15%
MMCF-GA-1	92.25%
MMCF-GA-2	91.25%

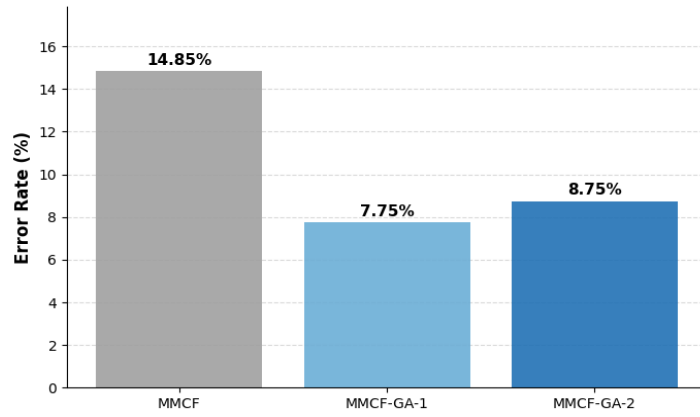


Figure 13: Comparison of Error Rates Among Face Recognition Models on the GT Dataset and LDA classifier.

As shown in Figure 13, the MMCF-GA-1 and MMCF-GA-2 models achieved markedly lower error rates on the GT dataset, reducing misclassification by about 47.8% and 41.1%, respectively, compared to the MMCF baseline.

4.4.3 Experiment Three

In this experimental phase, the proposed models were evaluated using the LFW face database described in Section 4.2.4. While the original database contains over 13,000 images, this study utilizes a strategically selected subset of 850 images representing 85 subjects (10 images per subject). This specific subset was chosen to ensure a fair and direct comparison with the original MMCF model [10], which employed the same data partitioning. Specifically, the first nine images of each subject were designated as the gallery set, while the remaining single image served as the probe (query) set, thereby maintaining statistical consistency across the evaluation. The images were utilized in their raw form, without any pre-processing, to rigorously test the models' robustness in unconstrained environments. To ensure the reliability of the findings, the recognition process was executed over 10 independent trials.

The results, presented in Table 8, indicate the superiority of the proposed models over the MMCF baseline. Specifically, the MMCF-GA-1 model achieved a recognition accuracy of 72.95%, marking a significant improvement of 9.42% over the MMCF model 63.53%. The MMCF-GA-2 model also demonstrated a strong performance, reaching 71.77% accuracy, which represents an 8.24% performance gain.

Table 8: Recognition Accuracy on the LFW Database and LDA Classifier.

Model Name	Recognition Rate
MMCF	63.53%
MMCF-GA-1	72.95%
MMCF-GA-2	71.77%

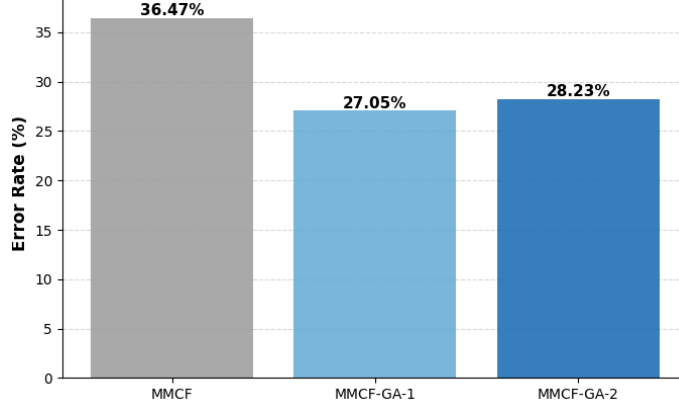


Figure 14: Comparison of Error Rates Among Face Recognition Models on the LFW Dataset and LDA classifier.

As shown in Figure 14, the MMCF-GA-1 and MMCF-GA-2 models achieved markedly lower error rates on the LFW dataset, reducing misclassification by about 25.83% and 22.6%, respectively, compared to the MMCF model.

4.4.4 Experiment Four

In this experimental phase, the proposed models were evaluated using the FERET face database described in Section 4.2.6. While the full database contains over 11,000 images, this study utilizes a specifically selected subset of 1,614 frontal images from 807 subjects (two images per subject). This subset was chosen to ensure a fair and direct comparison with the original MMCF model [10], which employed the same data partitioning to evaluate the SSPP scenario. Specifically, one image per subject was assigned to the gallery set, while the second image served as the probe query set. Consistent with the established protocol for this scenario, the recognition process was conducted once.

The results, presented in Table 9, show that both proposed models improved upon the MMCF model. The MMCF-GA-1 model achieved an accuracy of 98.51%, marking a 2.48% improvement over the MMCF model's accuracy of 96.03%. The MMCF-GA-2

model also demonstrated a performance gain, reaching 97.14% accuracy, which represents a 1.11% improvement.

Table 9: Recognition Accuracy on the FERET and LDA Classifier.

Model Name	Recognition Rate
MMCF	96.03%
MMCF-GA-1	98.51%
MMCF-GA-2	97.14%

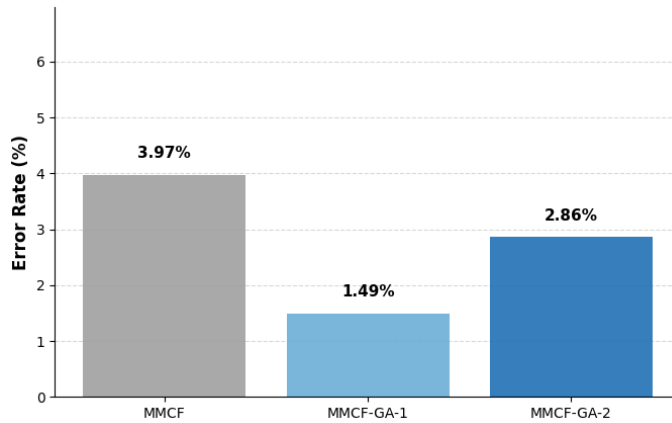


Figure 15: Illustrates the significant error rate reduction on the FERET database and LDA classifier.

The baseline MMCF model recorded an error rate of 3.97%. In stark contrast, the MMCF-GA-1 model reduced this error to just 1.49%, achieving a remarkable 62.5% relative reduction in error. The MMCF-GA-2 model also demonstrated a notable improvement, lowering the error rate to 2.86%, which corresponds to a 28.0% relative error reduction. These results confirm the superior performance of the GA-optimized models in the challenging SSPP scenario.

4.4.5 Experiment Five

In this experimental phase, the proposed models were evaluated using the PICS face database described in Section 4.2.7. Consistent with the database specifications, all 261 images from the 29 subjects were utilized to account for variations in pose and expression. To ensure a fair and direct comparison with the original MMCF model [10], the dataset was partitioned such that the first four images of each subject formed the gallery set, while the remaining five images served as the probe query set. To guarantee the reliability of the results, the recognition process was executed over five independent trials.

The results demonstrated that the MMCF-GA-1 and MMCF-GA-2 models significantly outperformed the MMCF model. Specifically, the MMCF-GA-1 model achieved a recognition accuracy of 99.30%, representing a 0.99% improvement over the MMCF model's accuracy of 98.31%. The MMCF-GA-2 model achieved an even higher accuracy of 99.60%, marking a 1.29% performance gain, as summarized in Table 10.

Table 10: Recognition Accuracy on the PICS Database and LDA Classifier.

Model Name	Recognition Rate
MMCF	98.31%
MMCF-GA-1	99.30%
MMCF-GA-2	99.60%

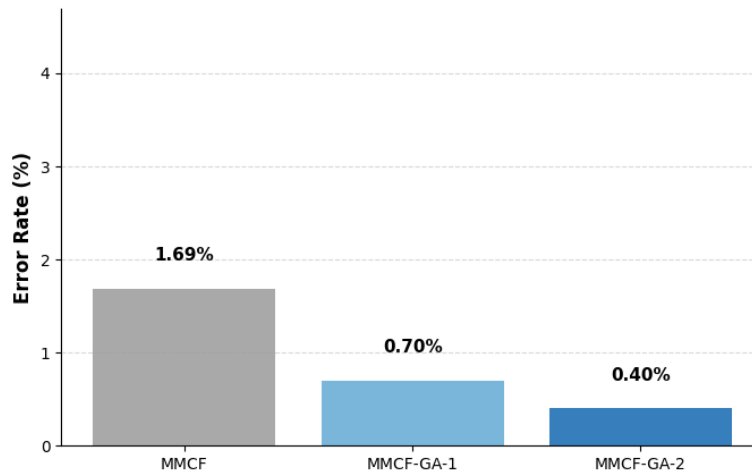


Figure 16: Comparison of Error Rates Among Face Recognition Models on the PICS Dataset and LDA classifier.

Figure 16, clearly demonstrates the superior performance of the MMCF-GA-1 and MMCF-GA-2 models over the traditional MMCF model. The genetic models exhibit significantly lower error rates, with MMCF-GA-1 achieving 0.7% and MMCF-GA-2 reaching 0.4%. In contrast, the MMCF model suffers from a higher error rate of 1.69%. This corresponds to a relative error reduction of 58.6% for MMCF-GA-1 and an impressive 76.3% for MMCF-GA-2, signifying a substantial improvement in Face Recognition accuracy.

4.4.6 Experiment Six

In this experimental phase, the proposed models were evaluated using the FRGC face database described in Section 4.2.8. While the full database contains over 50,000 images,

this study utilizes a specifically selected subset of 7,992 images from 222 subjects (36 images per subject) to capture significant variations in expressions. This subset was chosen to ensure a fair and direct comparison with the original MMCF model [10], which employed the same data partitioning. Specifically, 8 images per subject were designated as the gallery set, while the remaining 28 images served as the probe query set. To ensure the reliability and robustness of the findings, the recognition process was executed over five independent trials.

The results indicated that the MMCF-GA-1 and MMCF-GA-2 models significantly outperformed the MMCF model. The MMCF-GA-1 model achieved a recognition accuracy of 93.13%, representing a 2.32% improvement over the MMCF model's accuracy of 90.81%. The MMCF-GA-2 model achieved a slightly higher accuracy of 93.19%, marking a 2.38% performance gain, as shown in Table 11.

Table 11: Recognition Accuracy on the FRGC Database and LDA Classifier.

Model Name	Recognition Rate
MMCF	90.81%
MMCF-GA-1	93.13%
MMCF-GA-2	93.19%

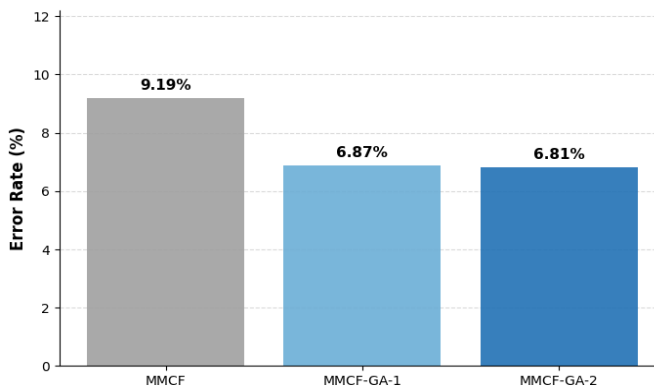


Figure 17: Comparison of Error Rates Among Face Recognition Models on the FRGC Dataset and LDA classifier.

Figure17 illustrates the error rates of the MMCF, MMCF-GA-1, and MMCF-GA-2 models. The MMCF model exhibits an error rate of 9.19%, while the genetic models achieve lower error rates of 6.87% MMCF-GA-1 and 6.81% MMCF-GA-2. This corresponds to a relative error reduction of 25.2% and 25.9% respectively, demonstrating a notable decrease in the Face Recognition error rate compared to the MMCF model.

4.5 Experimental Results Using the SRC Classifier

A second series of experiments was conducted using the SRC classifier to further validate and compare the performance of the MMCF-GA-1 and MMCF-GA-2 models with the MMCF model across the same datasets previously mentioned in Table 5.

4.5.1 Experiment Seven

In this experiment, we combined the AR and Curtin databases, utilizing the same data as referenced in Table 5. The results, as presented in Table 12, show that the proposed models outperformed the MMCF model. The MMCF-GA-1 model achieved the highest recognition rate of 99.99%, representing a 0.99% improvement over the MMCF model's accuracy of 99%. The MMCF-GA-2 model followed with an accuracy of 99.67%, marking a 0.68% performance gain. This demonstrates the superior ability of the GA-optimized models to maintain high accuracy even under challenging conditions such as extreme expressions and open mouths.

Table 12: Recognition Accuracy on the AR + Curtin Dataset Using SRC Classifier.

Model Name	Recognition Rate
MMCF	99.00%
MMCF-GA-1	99.99%
MMCF-GA-2	99.67%

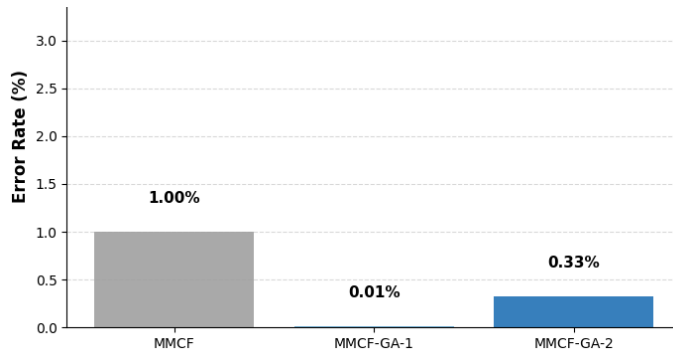


Figure 18: Comparative Analysis of Error Rates Among Models on the AR + Curtin Dataset Using SRC Classifier.

Figure 18 illustrates the comparison of error rates on the AR + Curtin dataset. The analysis reveals a dramatic reduction in errors for the GA-optimized models. While the MMCF model recorded an error rate of 1.00%, the MMCF-GA-1 and MMCF-GA-2 models achieved significantly lower rates of 0.01% and 0.33%, respectively. This corresponds to a relative error reduction of 99.0% for MMCF-GA-1 and 67.0% for MMCF-GA-2, reinforcing the finding that the GA-based approach drastically improves recognition accuracy.

4.5.2 Experiment Eight

In this experiment, we assessed the effectiveness of the MMCF-GA-1 and MMCF-GA-2 models in Face Recognition tasks using the GT face database, with the protocol consistent with that outlined in Table 5.

The findings, as detailed in Table 13, indicate that the MMCF-GA-1 model achieves the highest recognition accuracy at 95.51%, representing a 7.85% improvement over the MMCF model's accuracy of 87.66%. The MMCF-GA-2 model followed with an accuracy of 93.60%, marking a 5.94% performance gain. This demonstrates the models' robustness in handling the expression and misalignment challenges present in this dataset.

Table 13: Recognition Accuracy Applied the GT Database and SRC Classifier.

Model Name	Recognition Rate
MMCF	87.66%
MMCF-GA-1	95.51%
MMCF-GA-2	93.60%

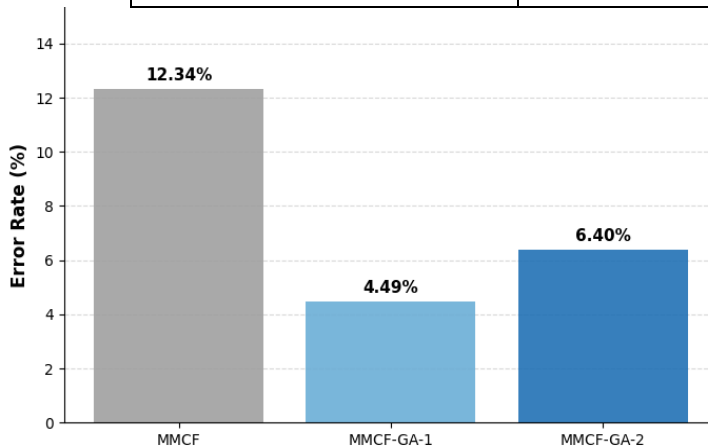


Figure 19: Comparative Analysis of Error Rates Among Models on the GT Dataset Using SRC Classifier.

Figure19 provides a comparison of the error rates on the GT dataset. The analysis highlights a substantial reduction in errors for the GA-optimized models. While the MMCF model recorded an error rate of 12.34%, the MMCF-GA-1 and MMCF-GA-2 models achieved significantly lower rates of 4.49% and 6.40%, respectively. This corresponds to a relative error reduction of 63.6% for MMCF-GA-1 and 48.1% for MMCF-GA-2, confirming the effectiveness of the GA in mitigating misclassifications caused by expression and misalignment challenges.

4.5.3 Experiment Nine

In this experiment, we assessed the effectiveness of the MMCF-GA-1 and MMCF-GA-2 models in Face Recognition tasks using the LFW face database, with the protocol consistent with that outlined in Table 5. The images were used directly, without any pre-processing or applying any pose correction techniques. The results, as illustrated in Table14, highlight the superior performance of the genetic models in this challenging unconstrained environment. The MMCF-GA-1 model achieved the highest recognition rate at 82.36%, representing a substantial 11.77% improvement over the MMCF model's accuracy of 70.59%. The MMCF-GA-2 model also demonstrated a strong performance, reaching 80.01% accuracy, which corresponds to a 9.42% performance gain.

Table 14: Recognition Accuracy Applied the LFW Dataset Using SRC Classifier.

Model Name	Recognition Rate
MMCF	70.59%
MMCF-GA-1	82.36%
MMCF-GA-2	80.01%

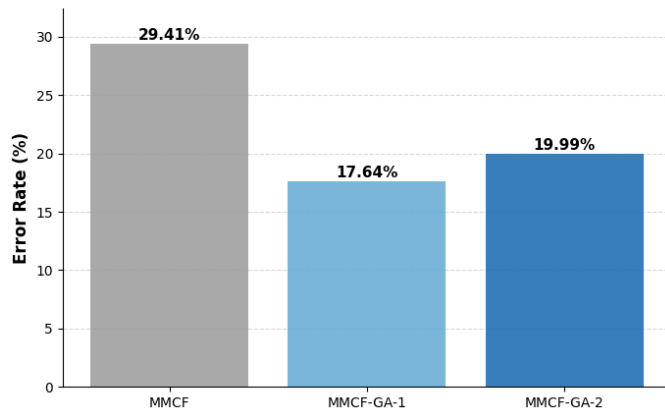


Figure 20: Comparative Analysis of Error Rates Among Models on the LFW Dataset Using SRC Classifier.

Figure20 presents a comparison of the error rates on the challenging LFW dataset. The analysis confirms the significant advantage of the GA-optimized models in unconstrained environments. While the MMCF model recorded a high error rate of 29.41%, the MMCF-GA-1 and MMCF-GA-2 models achieved substantially lower rates of 17.64% and 19.99%, respectively. This corresponds to a relative error reduction of 40.0% for MMCF-GA-1 and 32.0% for MMCF-GA-2, demonstrating the enhanced robustness of the GA approach against real-world variations in pose, lighting, and expression.

4.5.4 Experiment Ten

In this experiment, we evaluated the performance of the models on the FERET database, which tests the challenging SSPP scenario. The protocol used is consistent with that outlined in Table 5. The results, as depicted in Table 15, show that the MMCF-GA-1 model achieved the highest recognition rate at 99.38%, representing a 2.36% improvement over the MMCF model's accuracy of 97.02%. The MMCF-GA-2 model also demonstrated a strong performance, reaching 98.63% accuracy, which corresponds to a 1.61% performance gain. These findings highlight the robustness of the GA-optimized models in the data-scarce SSPP context.

Table 15: Recognition Accuracy Applied the FERET Dataset Using SRC Classifier.

Model Name	Recognition Rate
MMCF	97.02%
MMCF-GA-1	99.38%
MMCF-GA-2	98.63%

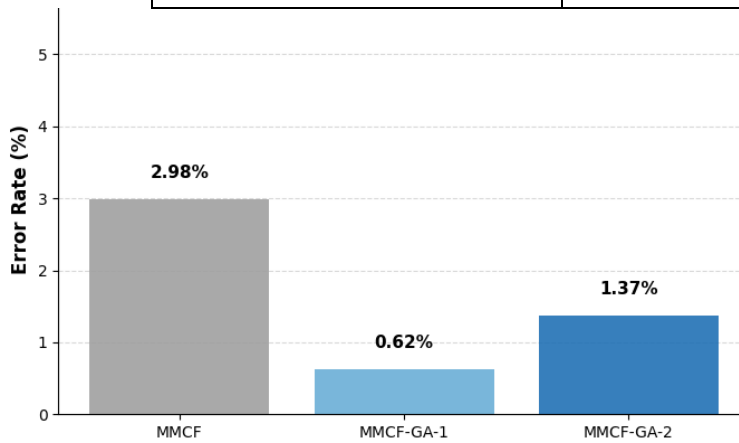


Figure 21: Comparative Analysis of Error Rates Among Models on the FERET Dataset Using SRC Classifier.

Figure 21 provides a comparison of the error rates on the FERET database, highlighting the models' performance in the SSPP scenario. The analysis reveals a dramatic reduction in errors for the GA-optimized models. While the MMCF model recorded an error rate of 2.98%, the MMCF-GA-1 and MMCF-GA-2 models achieved significantly lower rates of 0.62% and 1.37%, respectively. This corresponds to a relative error reduction of 79.2% for MMCF-GA-1 and 54.0% for MMCF-GA-2, confirming the superior robustness of the GA approach in data-scarce conditions.

4.5.5 Experiment Eleven

In this experiment, we evaluated the models' performance on the PICS database, which features slight variations in pose and expression. The protocol used is consistent with that outlined in Table 5.

The results, presented in Table 16, demonstrate the exceptional performance of the GA-optimized models. The MMCF-GA-1 model achieved a near-perfect recognition rate of 99.99%, representing a 1.12% improvement over the MMCF model's accuracy of 98.87%. The MMCF-GA-2 model also performed outstandingly, reaching 99.96% accuracy, which corresponds to a 1.09% performance gain.

Table 16: Recognition Accuracy Applied the PICS Dataset Using SRC Classifier.

Model Name	Recognition Rate
MMCF	98.87%
MMCF-GA-1	99.99%
MMCF-GA-2	99.96%

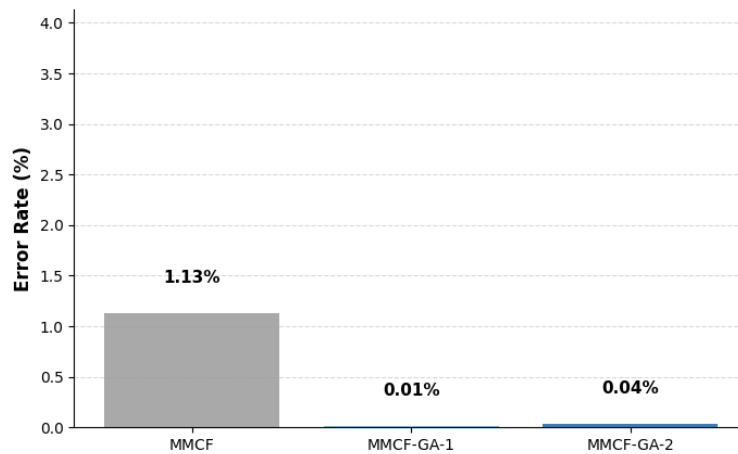


Figure 22: Comparative Analysis of Error Rates Among Models on the PICS Dataset Using SRC Classifier.

Figure22 provides a comparison of the error rates on the PICS database. The analysis reveals a near-total elimination of errors for the GA-optimized models. While the MMCF model recorded an error rate of 1.13%, the MMCF-GA-1 and MMCF-GA-2 models achieved remarkably low rates of 0.01% and 0.04%, respectively. This corresponds to a relative error reduction of 99.1% for MMCF-GA-1 and 96.5% for MMCF-GA-2, confirming the outstanding effectiveness of the GA in achieving near-flawless recognition.

4.5.6 Experiment Twelve

In this experiment, we evaluated the models' performance on the FRGC database, which is known for its challenging expression variations. The protocol used is consistent with that outlined in Table 5.

The results, presented in Table 17, show that the genetic models substantially outperformed the MMCF model. The MMCF-GA-1 model achieved a recognition accuracy of 98.22%, representing a significant 4.13% improvement over the MMCF model's accuracy of 94.09%. The MMCF-GA-2 model also demonstrated a strong performance, reaching 97.21% accuracy, which corresponds to a 3.12% performance gain.

Table 17: Recognition Accuracy Applied the FRGC Dataset Using SRC Classifier.

Model Name	Recognition Rate
MMCF	94.09%
MMCF-GA-1	98.22%
MMCF-GA-2	97.21%

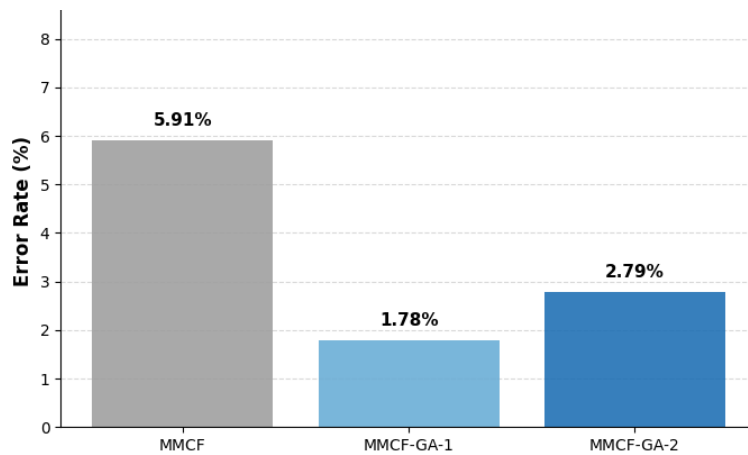


Figure 23: Comparative Analysis of Error Rates Among Models on the FRGC Dataset Using SRC Classifier.

Figure 23 provides a comparison of the error rates on the FRGC database. The analysis highlights the robustness of the GA-optimized models against expression and lighting challenges. While the MMCF model recorded an error rate of 5.91%, the MMCF-GA-1 and MMCF-GA-2 models achieved significantly lower rates of 1.78% and 2.79%, respectively. This corresponds to a relative error reduction of 69.9% for MMCF-GA-1 and 52.8% for MMCF-GA-2, confirming the superior performance of the GA approach.

4.6 Face Recognition Experiment under Pose Variations and Single Sample Per Person (SSPP).

The following subsections present a detailed evaluation of SPAE-integrated genetic models, aiming to address the challenges posed by pose variations and SSPP scenarios.

4.6.1 Experimental Evaluation of SPAE with Genetic Models

In this subsection, the genetic models MMCF-GA-1 and MMCF-GA-2 integrated with SPAE (SPAE-MMCF-GA-1 and SPAE-MMCF-GA-2) are evaluated for the Face Recognition task under pose variations and the SSPP scenario. This evaluation is conducted in comparison with: the original SPAE+MMCF model, SPAE-RGB model, and the grayscale SPAE-G model. To achieve this, an experiment was conducted using the Multi-Pie database, as detailed below.

4.6.2 Evaluation on the Multi-Pie Database

This experiment was conducted using a dataset from Multi-Pie, containing 3699 images corresponding to 137 subjects at angles of 0° , $\pm 15^\circ$, $\pm 30^\circ$, $\pm 45^\circ$, and $\pm 60^\circ$, as described in Section 4.2.3. Three frontal images were used as Gallery Images, while the remaining 24 images (previously corrected to a frontal pose) were used as Query Images. The rLDA algorithm [23] with the NN classifier [50] was employed to ensure a fair comparison with the original SPAE+MMCF model, as well as SPAE (G) and SPAE (RGB).

The results presented In Table18 demonstrate that the genetic models (SPAE-MMCF-GA-1 and SPAE-MMCF-GA-2) significantly outperform SPAE-MMCF, SPAE (RGB) and SPAE (G) particularly when dealing with large pose angles and a single gallery image.

- Under challenging pose variations with only one gallery image, at the angle of ($\pm 15^\circ$), the recognition rates of SPAE-MMCF-GA-1 and SPAE-MMCF-GA-2 reached 83.97% and 83.57%, respectively, at ($+15^\circ$), compared to 80.13% for SPAE-MMCF. This corresponds to an improvement of 3.84% and 3.44%, respectively.
- For larger pose angles with a single gallery image, such as ($+60^\circ$), the genetic models again demonstrated superior performance. SPAE-MMCF-GA-1 and

SPAE-MMCF-GA-2 achieved recognition rates of 56.94% and 56.76%, respectively, compared to only 51.85% for SPAE-MMCF. This represents an improvement of 5.09% and 4.91%, respectively, confirming the effectiveness of the genetic models in addressing the challenges of large pose variations, even with only a single gallery image available.

Under challenging facial pose variations and with a multi-image gallery, the genetic models SPAE-MMCF-GA-1 and SPAE-MMCF-GA-2 consistently outperformed the baseline model SPAE-MMCF, with this advantage being particularly evident in the analysis of error rate reduction.

For instance, at large pose angles such as $+60^\circ$, the SPAE-MMCF model exhibited an error rate of 38.55%, whereas SPAE-MMCF-GA-1 reduced this error to 36.24%, achieving a relative decrease of 5.99%.

At near-frontal angles like -15° , where the baseline model already had high accuracy (99.32%, corresponding to an error of 0.68%), the MMCF-GA-1 model achieved higher accuracy of 99.69% (error 0.31%). Although this improvement in accuracy appears marginal, it translates into a substantial reduction in the error rate by 54.41%, effectively reducing the SPAE-MMCF model's errors by more than half.

Collectively, these results, as detailed in Table 18, confirm the consistent and significant superiority of the SPAE-MMCF-GA-1 and SPAE-MMCF-GA-2 models over the original SPAE-MMCF, as well as the simpler SPAE (RGB) and SPAE (Gray) representations, across all evaluated scenarios.

Table 18: Face Recognition on Multi-Pie Using SPAE-MMCF-GA-1 and SPAE-MMCF-GA-2 Compared with SPAE-MMCF, SPAE-RGB, and SPAE-G.

galleries	Number of	Pose angle	-60	-45	-30	-15	+15	+30	+45	+60	average
1		SPAE(G)	28.97%	32.89%	39.65%	45.15%	45.75%	38.86%	32.67%	28.87%	36.60%
		SPAE-RGB	28.94%	33.05%	40.74%	45.96%	46.32%	42.17%	34.45%	29.24%	37.61%
		SPAE-MMCF	50.99%	70.08%	76.29%	78.68%	80.13%	77.10%	72.64%	51.85%	69.72%
		SPAE-MMCF-GA-1	54.96%	72.99%	79.40%	81.97%	83.97%	80.73%	75.71%	56.94%	75.83%
		SPAE MMCF-GA-2	54.28%	72.52%	79.12%	81.55%	83.57%	80.45%	74.98%	56.76%	75.40%
2		SPAE(G)	43.46%	79.64%	93.17%	98.75%	98.29%	94.59%	82.93%	45.93%	79.60%
		SPAE-RGB	51.38%	84.95%	96.59%	99.30%	98.70%	97.82%	90.75%	52.48%	83.99%
		SPAE-MMCF	61.12%	97.27%	98.76%	99.32%	98.71%	98.50%	97.60%	61.20%	89.06%
		SPAE-MMCF-GA-1	63.30%	98.57%	99.02%	99.58%	99.32%	98.96%	98.57%	63.35%	90.08%
		SPAE MMCF-GA-2	63.25%	98.24%	99.02%	99.53%	99.32%	98.96%	98.54%	63.34%	90.03%
3		SPAE(G)	45.33%	82.87%	94.21%	98.93%	98.69%	95.89%	84.48%	47.43%	80.98%
		SPAE-RGB	52.62%	86.75%	97.93%	99.30%	98.71%	98.72%	95.59%	53.26%	85.36%
		SPAE-MMCF	61.57%	97.95%	98.81%	99.32%	98.83%	98.74%	97.95%	61.45%	89.33%
		SPAE-MMCF-GA-1	63.46%	98.57%	99.21%	99.69%	99.36%	99.30%	98.60%	63.76%	90.25%
		SPAE MMCF-GA-2	63.45%	98.25%	99.19%	99.69%	99.34%	99.27%	98.60%	63.74%	90.20%

4.7 Evaluation of Genetic Models (MMCF-GA-1 and MMCF-GA-2) Using LFW Dataset Corrected by the Random Mask Attention Generative Adversarial Network (RMAGAN).

This section presents the results of a critical experiment designed to evaluate the proposed genetic models in the context of advanced facial pose correction. As outlined in the methodology chapter, the MMCF-GA-1 and MMCF-GA-2 models were applied to the LFW dataset, which consists of 13,233 face images across 5,749 subjects [45]. For this experiment, a subset of 158 subjects was selected, each with exactly 10 images. These images were preprocessed and frontalized using the state-of-the-art RMAGAN [44].

The evaluation was conducted under two scenarios:

- Single-Sample Face Recognition (SSFR): One image per subject was used for training, while the remaining nine images were used for testing.
- Multi-Sample Face Recognition: Three images per subject were used for training, and the remaining seven for testing.

Each experiment was repeated ten times to ensure statistical reliability.

This setup allowed us to rigorously test the robustness and performance of our genetic models on data generated by a modern generative algorithm. The goal was to verify that MMCF-GA-1 and MMCF-GA-2 can effectively augment existing techniques and improve overall image representation and recognition accuracy.

Table 19: Face Recognition Accuracy on LFW Dataset Corrected by the RMAGAN

Number of Gallery Images	RMAGAN (RGB)	MMCF	MMCF-GA-1	MMCF-GA-2
1	93.95%	94.34%	97.85%	97.84%
3	99.82%	99.87%	99.98%	99.96%

The evaluation results underline the progressive improvement achieved across different models. Starting from the baseline RGB representation, which yielded an error rate of 6.05% in the single-sample scenario, the standard MMCF framework already demonstrated its effectiveness by reducing the error to 5.66%. This confirms that MMCF provides a more robust representation than conventional RMAGAN(RGB) features. However, the introduction of genetic optimization produced a remarkable leap in performance. Specifically, MMCF-GA-1 lowered the error rate to 2.15%, corresponding to a 64.5% reduction in error compared to RGB, while MMCF-GA-2 achieved a similar outcome with 2.16%.

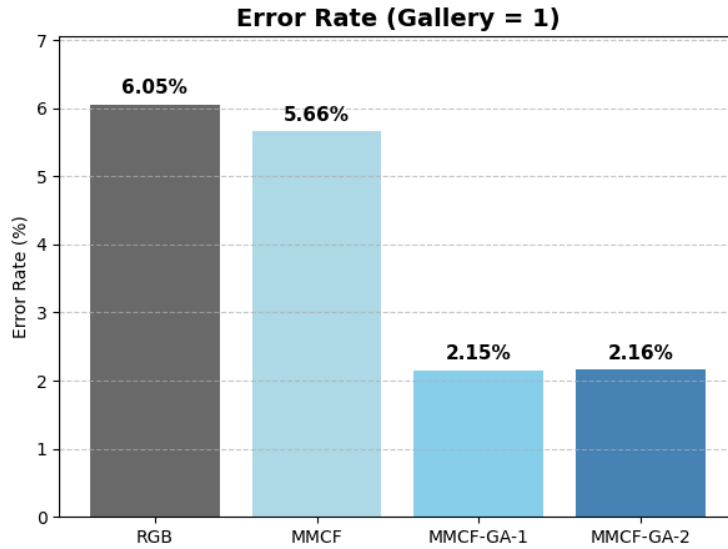


Figure 24: Comparative Analysis of Error Rates Among Models on the Pose-Corrected LFW Dataset (Gallery = 1).

In the multi-sample setting, where the error rates were already very low, MMCF further reduced the baseline error of 0.18% (RGB) to 0.13%. Yet, the genetic extensions again provided a clear advantage: MMCF-GA-1 reached 0.02% (88.9% error reduction) and MMCF-GA-2 attained 0.04%.

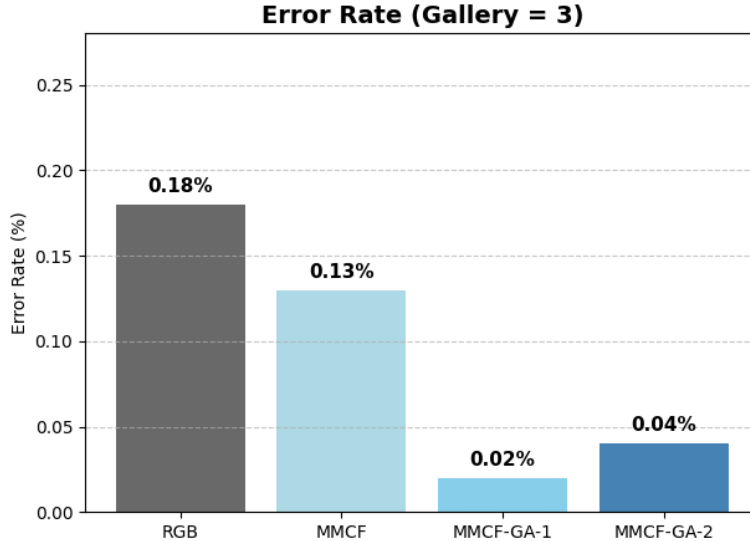


Figure 25: Comparative Analysis of Error Rates Among Models on the Pose-Corrected LFW Dataset (Gallery = 3).

These findings establish a clear hierarchy: RGB provides the weakest baseline, MMCF significantly strengthens representation through color and resolution fusion, and the genetic models (especially MMCF-GA-1) deliver the highest robustness and generalization. Overall, the results validate the effectiveness of combining MMCF with genetic optimization to tackle the challenges of pose-corrected Face Recognition.

4.8 Discussion

The superior performance of the proposed genetic models (MMCF-GA-1 model and MMCF-GA-2 model) resulted in substantial improvements in recognition accuracy and reduction in error rates. Despite the inherent challenges and diverse conditions posed by the employed datasets, the genetic models consistently outperformed their primary counterpart, the MMCF model. This remarkable adaptability was evident across variations in pose, expression, and misalignment, successfully demonstrated through their application across eight distinct facial recognition databases. The AR and Curtin databases, which included frontal faces with their extreme expressions and extremely open mouths, posed a unique challenge, yet the genetic models achieved recognition accuracies 99.99%. Similarly, the GT database demonstrated the models' capacity to handle variations in frontal and tilted faces, as well as differences in facial expressions and pose directions, yielding a significant improvement of 7.85% over the MMCF model. Furthermore, the LFW database, renowned for its unconstrained nature, presented a significant challenge due to extensive variations in facial expression, posture, skin color, conditions, and image quality. Despite these challenges the images were utilized in their original form, without any pre-processing or application of pose correction techniques. The proposed genetic models exhibited superior performance, demonstrating an improvement in recognition accuracy from 9.42% to 11.77% compared to the MMCF model. Similarly, the genetic models demonstrated a significant improvement in accuracy when using a single gallery image, as evidenced in the FERET database.

PICs databases, with their own distinct variations, and FRGC databases, they are considered one of the most challenging due to its variations in facial expressions and other conditions. Further reinforced the genetic models' excellence, consistently outperforming the MMCF model.

Despite the inherent challenges and diverse conditions posed by the employed datasets, the proposed genetic models consistently outperform their primary counterpart, the MMCF model.

Regardless of whether the testing was conducted using the LDA classifier or the SRC classifier, the results were consistently superior. The MMCF-GA-1 and MMCF-GA-2 models exhibit slight differences in performance. MMCF-GA-1 Model, which consisted of 7 color channels, achieved higher recognition accuracy and lower error rates compared to MMCF-GA-2 model, which consisted of 8 color channels.

This performance variance is further elucidated by the specific strengths of each framework. The MMCF-GA-1 model, when integrated with the SRC classifier, demonstrates a strong generalization capability in complex and unconstrained environments, such as the LFW and GT datasets. This robustness is primarily attributed to the adopted robust feature selection strategy, which achieves an effective balance between recognition accuracy and resistance to noise, pose variations, and complex facial expressions commonly encountered in real-world conditions. Conversely, the MMCF-GA-2 model achieves superior performance on more controlled datasets, including AR+Curtin, PICS, and FRGC, particularly when combined with the LDA classifier. This advantage can be explained by the model's emphasis on optimizing color-channel fusion to enhance discriminative power in scenarios characterized by high image quality and more regular intra-class variations.

However, the gap between the two genetic models is minimal, and both significantly outperform the MMCF model, as illustrated in Figures 26 and 27. The results using SRC classifiers are shown in Figure 26, while the results using LDA classifiers are shown in Figure 27.

Furthermore, face pose correction was integrated with the genetic models using the SPAE pose correction algorithm. Experiments on the Multi-Pie dataset demonstrated a notable superiority of the proposed models. In particular, SPAE-MMCF-GA-1 and SPAE-MMCF-GA-2 achieved significantly higher recognition rates compared to RGB representation, the MMCF model, and grayscale representation. The stability of these models was especially evident in cases involving large pose angles and when using only a single gallery image, confirming their ability to maintain high accuracy even under the most challenging testing conditions.

Similarly, the MMCF-GA-1, MMCF-GA-2, and MMCF models were evaluated on the LFW dataset corrected using the RMAGAN method, all showing improvements over the RMAGAN (RGB) representation. In the single-sample scenario, the error rate decreased from 6.05% RMAGAN (RGB) to 5.66% (MMCF), and further to .15% and 2.16% for the genetic variants. In the multi-sample scenario, the error rate dropped from 0.18% to 0.13% (MMCF), and then to 0.02% and 0.04% for the genetic models.

Such strong performance demonstrates the capability of the proposed models to effectively support and complement other face recognition methodologies, stemming from their core emphasis on enhancing the underlying image representation.

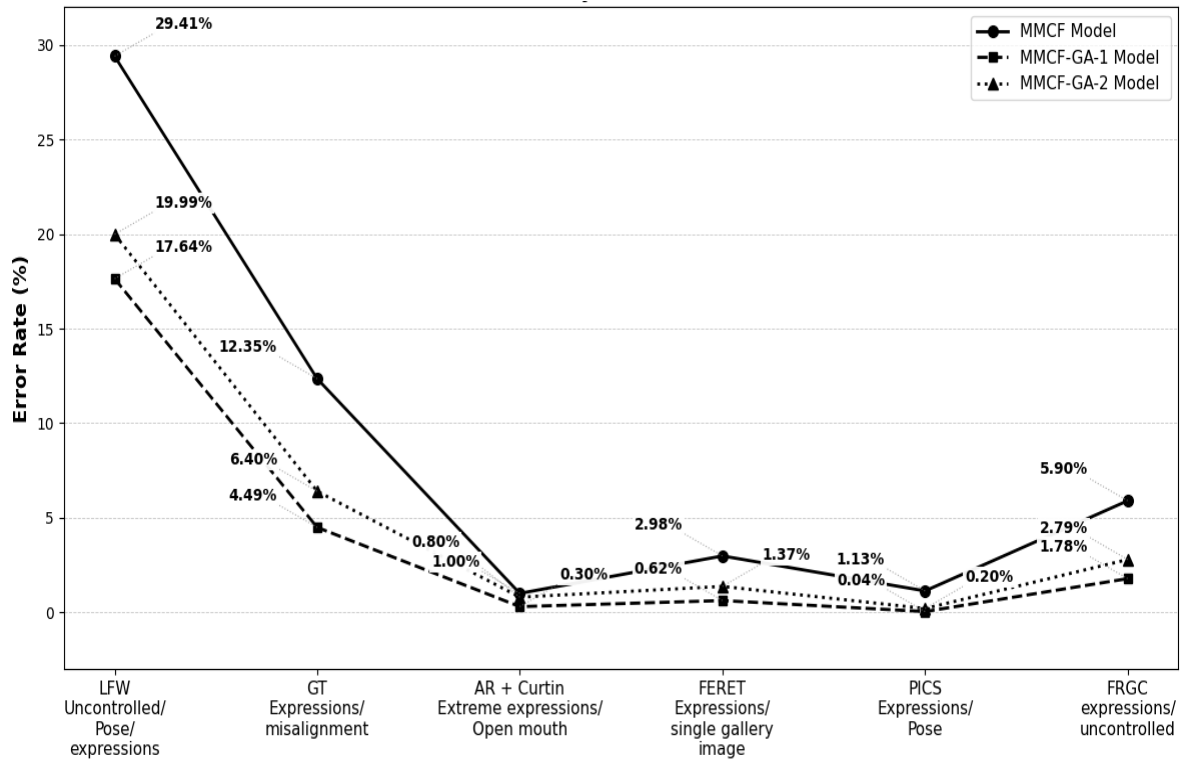


Figure 26: Comparison of Error Rates Among MMCF-GA-1, MMCF-GA-2 models and MMCF Model Under Various Database Conditions Using SRC Classifiers.

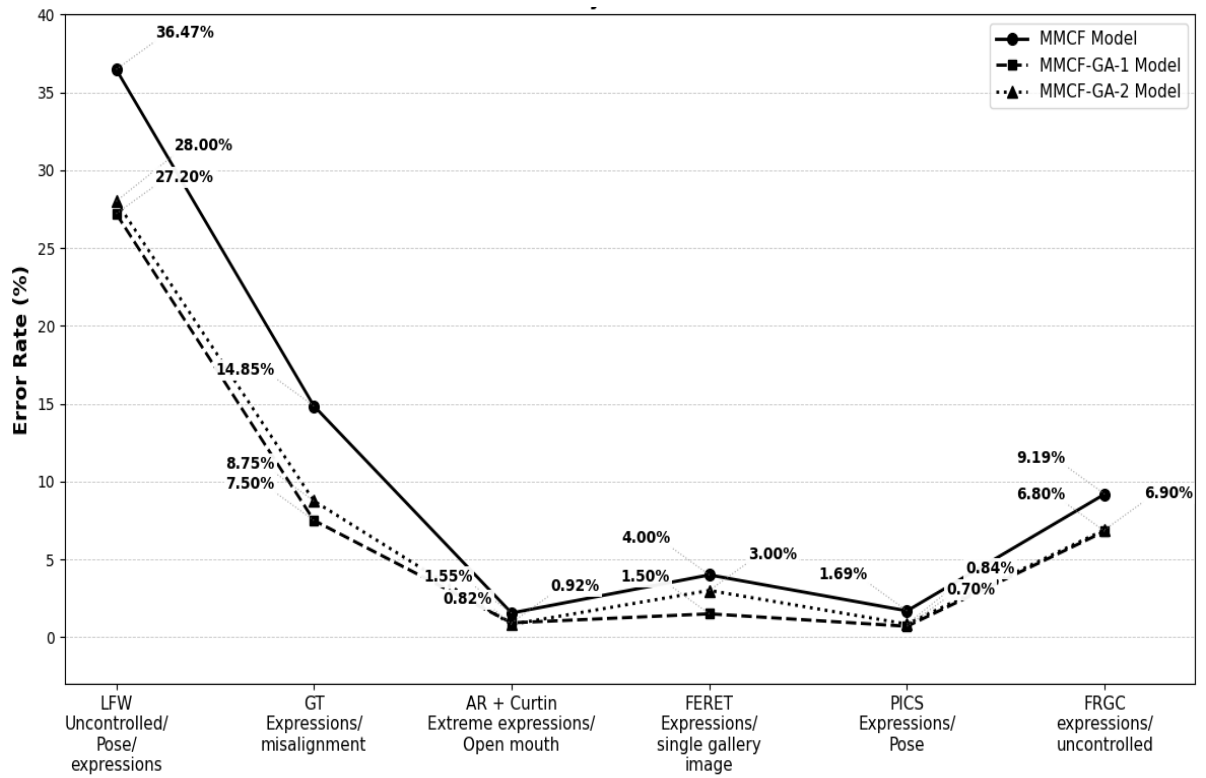


Figure 27: Comparison of Error Rates Among MMCF-GA-1, MMCF-GA-2 models, and MMCF Model Under Various Database Conditions Using LDA Classifiers.

Chapter 5

Conclusion and recommendations for Future Works

5.1 Conclusion

The primary objective of this research was to develop and enhance image representation techniques to improve the accuracy and robustness of Face Recognition systems. This goal was achieved by designing a novel framework that employs a GA to optimize the selection of multi-resolution multiple color channels, replacing the sub-optimal greedy search used in previous MMCF models.

Through this enhanced search strategy, two innovative models MMCF-GA-1 and MMCF-GA-2 were developed and evaluated across a wide range of benchmark Face Recognition datasets, encompassing variations in pose, facial expression, and image quality. The experimental results consistently demonstrated that the proposed genetic models outperform the baseline MMCF, representations while maintaining stable performance under challenging and unconstrained conditions.

Furthermore, integrating the genetic models with advanced pose correction techniques, such as SPAE and RMAGAN, further improved their adaptability and resilience to large pose variations. This confirms that combining multi-resolution multiple color information with robust optimization and pose correction strategies can significantly strengthen Face Recognition systems. Additionally, the genetic models demonstrated high strength and robustness in the highly challenging SSPP scenario, which proves their effectiveness and ability to operate with minimal data.

In conclusion, this research successfully delivered on its objective by introducing a more effective search methodology for constructing multi-channel, multi-resolution color models. The proposed approach provides a solid foundation for developing intelligent and reliable Face Recognition systems.

5.2 Future Works

With significant potential for future enhancement, the proposed framework could be extended by integrating deep learning techniques or additional optimization strategies to further improve accuracy and robustness. Moreover, developing a hybrid algorithm that combines greedy search with the GA may yield better performance and faster convergence. Such advancements could strengthen the adaptability and resilience of Face Recognition systems in complex real-world scenarios.

References

References

- [1] Q. Liu, T. Chen, Y. Zhang, and Y. Wang, "Facial Recognition technology: A comprehensive overview," *Academic Journal of Computing & Information Science*, vol. 6, no. 7, pp. 15–26, 2023.
- [2] L. Li, X. Mu, S. Li, and H. Peng, "A review of Face Recognition technology," *IEEE Access*, vol. 8, pp. 139110–139120, 2020.
- [3] H. L. Gururaj, B. C. Soundarya, S. Priya, J. Shreyas, and F. Flammini, "A comprehensive review of Face Recognition techniques, trends and challenges," *IEEE Access*, vol. 12, pp. 107903–107926, 2024.
- [4] Z. Cheng, X. Zhu, and S. Gong, "Face re-identification challenge: Are Face Recognition models good enough?," *Pattern Recognit.*, vol. 107, Art. no. 107422, Nov. 2020.
- [5] F. Liu, D. Chen, F. Wang, Z. Li, and F. Xu, "Deep learning based single sample Face Recognition: A survey," *Artif. Intell. Rev.*, vol. 56, no. 3, pp. 2723–2748, Mar. 2023, doi: 10.1007/s10462-022-10240-2.
- [6] A. Bhatta, D. Mery, H. Wu, J. Annan, M. C. King, and K. W. Bowyer, "What's color got to do with it? Face recognition in grayscale," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, 2024.
- [7] M. M. Alrjebi, "Robust Face Recognition via multi-channel models and feature selection," Ph.D. dissertation, Curtin University, 2017.
- [8] B. Li, S. An, W. Liu, and A. Krishna, "The MCF model: Utilizing multiple colors for Face Recognition," in *Proc. 6th Int. Conf. Image Graph. (ICIG)*, 2011, pp. 1029–1034.
- [9] M. M. Alrjebi, W. Liu, and L. Li, "Two directional multiple color fusion for Face Recognition," in *Proc. Int. Conf. Comput., Commun., Syst. (ICCCS)*, Nov. 2015, pp. 171–177.
- [10] M. M. Alrjebi, W. Liu, and L. Li, "Face Recognition against pose variations using multi-resolution multiple color fusion," *Int. J. Mach. Intell. Sens. Signal Process.*, vol. 1, no. 4, pp. 304–320, 2017.
- [11] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*, 3rd ed. Cambridge, MA, USA: MIT Press, 2022.
- [12] R. E. Neapolitan and K. Naimipour, *Foundations of Algorithms*. Boston, MA, USA: Jones & Bartlett, 2015.
- [13] M. Turk and A. Pentland, "Eigenfaces for Recognition," *J. Cognit. Neurosci.*, vol. 3, no. 1, pp. 71–86, 1991.
- [14] K. Etemad and R. Chellappa, "Discriminant analysis for Recognition of human face images," *J. Opt. Soc. Am. A*, vol. 14, no. 8, pp. 1724–1733, Aug. 1997.
- [15] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust Face Recognition via sparse representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 2, pp. 210–227, Feb. 2009.
- [16] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 4th ed. New York, NY, USA: Pearson, 2018.

- [17] M. Kan, S. Shan, H. Chang, and X. Chen, "Stacked progressive auto-encoders (SPAEE) for Face Recognition across poses," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2014, pp. 1883–1890.
- [18] I. Goodfellow *et al.*, "Generative Adversarial Nets," in *Adv. Neural Inf. Process. Syst.*, vol. 27, 2014.
- [19] X. Huang, L. Wu, and Y. Ye, "A review on dimensionality reduction techniques," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 33, no. 10, p. 1950017, Sep. 2019.
- [20] G. T. Reddy *et al.*, "Analysis of dimensionality reduction techniques on big data," *IEEE Access*, vol. 8, pp. 54776–54788, 2020.
- [21] H. Zhao, Z. Lai, H. Leung, and X. Zhang, "Principal component analysis: In feature learning and understanding algorithms and applications," in *Information Fusion and Data Science*. Cham, Switzerland: Springer, 2020, pp. 31–52, doi: 10.1007/978-3-030-40794-0.
- [22] H. Zhao, Z. Lai, H. Leung, and X. Zhang, "Linear discriminant analysis: In feature learning and understanding algorithms and applications," in *Information Fusion and Data Science*. Cham, Switzerland: Springer, 2020, pp. 71–84, doi: 10.1007/978-3-030-40794-0.
- [23] J. Lu, K. N. Plataniotis, and A. N. Venetsanopoulos, "Regularization studies of linear discriminant analysis in small sample size scenarios with application to Face Recognition," *Pattern Recognit. Lett.*, vol. 26, no. 2, pp. 181–191, Jan. 2005.
- [24] M. M. Alrjebi, N. Pathirage, W. Liu, and L. Li, "Face Recognition against occlusions via I fusion using 2D-MCF model and SRC," *Pattern Recognit. Lett.*, vol. 95, pp. 14–21, 2017.
- [25] J. Yang, D. Chu, and J.-Y. Yang, "Sparse representation classifier steered discriminative projection," in *Proc. Int. Conf. Pattern Recognit. (ICPR)*, Aug. 2010, pp. 694–697.
- [26] K. De and M. Pedersen, "Impact of I on robustness of deep neural networks," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2021, pp. 21–30.
- [27] A. Humeau-Heurtier, "Color texture analysis: A survey," *IEEE Access*, vol. 10, pp. 107993–108003, 2022.
- [28] G. Padmashree and K. A. Kotegar, "Skin segmentation-based disguised Face Recognition using deep learning," *IEEE Access*, vol. 12, pp. 51056–51072, 2024.
- [29] J. Yang, C. Liu, and L. Zhang, "Color space normalization: Enhancing the discriminating power of color spaces for Face Recognition," *Pattern Recognit.*, vol. 43, no. 4, pp. 1454–1466, 2010.
- [30] J. C. Russ, *Introduction to Image Processing and Analysis*. Boca Raton, FL, USA: CRC Press, 2017, doi: 10.1201/9781315221939.
- [31] M. Li *et al.*, "A better color space conversion based on learned variances for image compression," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2019, pp. 0–7.
- [32] Y. Ohta, T. Kanade, and T. Sakai, "Color information for region segmentation," *Comput. Graph. Image Process.*, vol. 13, no. 3, pp. 222–241, 1980.

- [33] S. K. Ranade and S. Anand, "Color Face Recognition using normalized-discriminant hybrid color space and quaternion moment vector features," *Multimedia Tools Appl.*, vol. 80, no. 7, pp. 10797–10820, 2021.
- [34] J. Yang and C. Liu, "A discriminant color space method for face representation and verification on a large-scale database," in *Proc. IEEE Int. Conf. Pattern Recognit. (ICPR)*, 2008, pp. 1–4.
- [35] B. Y. Li, W. Liu, S. An, A. Krishna, and T. Xu, "Face Recognition using various scales of discriminant color space transform," *Neurocomputing*, vol. 94, pp. 68–76, Oct. 2012.
- [36] C. Liu, "Learning the uncorrelated, independent, and discriminating color spaces for Face Recognition," *IEEE Trans. Inf. Forensics Security*, vol. 3, no. 2, pp. 213–222, Jun. 2008.
- [37] S. Mirjalili, J. S. Dong, A. S. Sadiq, and H. Faris, "Genetic algorithm: Theory, literature review, and application in image reconstruction," in *Studies in Computational Intelligence*. Cham, Switzerland: Springer, 2020, pp. 69–85.
- [38] S. Katoch, S. S. Chauhan, and V. Kumar, "A comprehensive review on genetic algorithm: Past, present, and future," *Multimedia Tools Appl.*, vol. 80, no. 5, pp. 8091–8126, Feb. 2021, doi: 10.1007/s11042-020-10139-6.
- [39] S. J. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 4th ed. Hoboken, NJ, USA: Pearson, 2020.
- [40] A. M. Martinez and R. Benavente, "The AR face database," CVC Tech. Rep., Jun. 1998.
- [41] B. Y. Li, W. Liu, S. An, and A. Krishna, "Robust Face Recognition by utilizing color information and sparse representation," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 28, no. 03, p. 1456004, May 2014.
- [42] J. Brownlee, "Genetic algorithm afternoon: A practical guide for software developers," AlgorithmAfternoon.com, 2024, ch. 6.
- [43] R. Gross, I. Matthews, J. Cohn, T. Kanade, and S. Baker, "Multi-PIE," *Image Vis. Comput.*, vol. 28, no. 5, pp. 807–813, 2010.
- [44] J. Liao, T. Guha, and V. Sanchez, "Self-supervised random mask attention GAN in tackling pose-invariant Face Recognition," *Pattern Recognit.*, vol. 159, p. 111112, Mar. 2025, doi: 10.1016/j.patcog.2024.111112.
- [45] G. B. Huang, M. Mattar, T. Berg, and E. Learned-Miller, "Labeled faces in the wild: A database for studying Face Recognition in unconstrained environments," in *Proc. Workshop Faces Real-Life Images: Detect., Align., Recognit.*, 2008.
- [46] A. V. Nefian, "Georgia tech face database," Center for Signal and Image Processing, Georgia Institute of Technology, Atlanta, GA, USA, Tech. Rep., 1999.
- [47] P. J. Phillips, H. Wechsler, J. Huang, and P. J. Rauss, "The FERET database and evaluation procedure for face-recognition algorithms," *Image Vis. Comput.*, vol. 16, no. 5, pp. 295–306, 1998.
- [48] P. Hancock, The Psychological Image Collection at Stirling (PICS), 2008. [Online]. Available: <http://pics.psych.stir.ac.uk/>

- [49] P. J. Phillips *et al.*, "Overview of the Face Recognition grand challenge," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2005.
- [50] T. M. Cover and P. E. Hart, "Nearest neighbor pattern classification," *IEEE Trans. Inf. Theory*, vol. 13, no. 1, pp. 21–27, Jan. 1967.

الملخص

يُعدّ التعرف على الوجوه أحد المجالات المحورية في الرؤية الحاسوبية، نظرًا لتطبيقاته الواسعة في الأمن، والتحقق من الهوية، والتفاعل بين الإنسان والحاسوب. وقد أظهرت الأبحاث الحديثة أن الاستفادة من مساحات الألوان المتعددة والتمثيلات متعددة الدقة يمكن أن يعزز دقة التعرف على الوجوه بشكل ملحوظ مقارنة بالأساليب التقليدية المعتمدة على تمثيل RGB أو التدرج الرمادي. ومع ذلك، فإن النماذج متعددة القنوات الألوان الحالية، مثل نموذج دمج قنوات متعدد الألوان ومتعدد الاحجام (MMCF)، تعتمد على خوارزمية البحث الشره لاختيار أفضل التركيبات المثلى من القنوات اللونية، وهي خوارزمية عرضة للوقوع في الحلول المحلية وغير قادرة على استكشاف مساحة البحث بشكل كافٍ.

تقترح هذه الرسالة إطارًا جديدًا للتحسين وذلك باستبدال خوارزمية البحث الشره بخوارزمية التحسين الجينية، بهدف تحديد أكثر التركيبات فعالية للقنوات اللونية متعددة الألوان متعدد الاحجام في التعرف على الوجوه. وقد تم تطوير نموذجين معتمدين على الخوارزمية الجينية، هما MMCF-GA-1 و MMCF-GA-2، باستخدام مصنفين مختلفين هما التصنيف القائم على التمثيل المتناثر (SRC) وتحليل التباين الخطي (LDA) وتم تقييم أدائهما على عدة قواعد بيانات قياسية تحت ظروف متنوعة وصعبة، بما في ذلك اختلافات الوضعية، وتعبيرات الوجه، والبيئات غير المقيدة. أظهرت النتائج التجريبية أن النماذج المقترحة المعتمدة على الخوارزمية الجينية تتفوق باستمرار على كل من تمثيلات التدرج الرمادي ونموذج RGB و نموذج MMCF، محققة دقة أعلى ومثانة أكبر في جميع السيناريوهات المختبرة. علاوةً على ذلك فإن دمج هذه النماذج الجينية مع تقنيات متقدمة لتصحيح وضعية الوجوه في الصور، مثل المُشَفِّر التلقائي التراكمي التدريجي (SPAЕ) وشبكة الخصومة التوليدية المعتمدة على انتباه القناع العشوائي (RMAGAN) قد عزز من قدرتها على التكيف مع التغيرات الكبيرة في وضعيات الوجه.

وقد تجلّت هذه المثانة والقوة بشكل خاص في سيناريو العينة الواحدة لكل شخص (SSPP) شديد الصعوبة، حيث حافظت النماذج على دقة تعرف عالية رغم الحد الأدنى من بيانات التدريب. وتؤكد هذه النتائج أن دمج المعلومات القنوات اللونية المتعدد مع الاحجام المتعددة واستراتيجيات التحسين القوية يمكن أن يعزز بشكل كبير من قوة وكفاءة أنظمة التعرف على الوجوه.



التعرف على الوجه باستخدام نموذج الصور متعدد الألوان والأحجام اعتماداً على الخوارزمية الجينية

إعداد الطالبة: بشرى علي عمر أبو السلوقي

تحت إشراف: د. مصطفى المختار الرجبي

الدرجة العلمية: أستاذ مساعد

قدمت هذه الرسالة استكمالاً لمتطلبات الحصول على درجة الإجازة العالية

(الماجستير) في علوم الحاسب الآلي

بتاريخ 1 / 1 / 2026م