A NOVEL FACE RECOGNITION SYSTEM IN UNCONSTRAINED ENVIRONMENTS USING A CONVOLUTIONAL NEURAL NETWORK

by

Muhtahir Oluwaseyi Oloyede

Submitted in partial fulfillment of the requirements for the degree
Doctor of Philosophy (Computer Engineering)

in the

Department of Electrical, Electronic and Computer Engineering
Faculty of Engineering, Built Environment, and Information Technology

UNIVERSITY OF PRETORIA

November 2018

DEDICATION

This Ph.D. research work is dedicated to Almighty Allah the beneficent, the merciful for making this dream a reality; and keeping me strong and focused throughout the research journey.

ACKNOWLEDGEMENT

This doctoral program would not have been possible without the blessings and support of Almighty Allah, and the lovely people that surround me. I will not be able to mention everyone, but I really appreciate all those who have contributed to the success of my Ph.D. degree.

Firstly, I would like to thank my supervisor, Professor Gerhard P. Hancke who has given me this great opportunity of undergoing the Ph.D. research under his supervision. I have gained so much from working with you, and I am grateful for the opportunity. You are indeed a role model and an academic father.

I would like to thank my co-supervisor, Dr. Herman C. Myburgh who has been a strong pillar of support upon which the Ph.D. degree is achieved. I am sincerely grateful for all your contribution, encouragement and guidance towards the successful completion of the Ph.D. degree.

I would like to appreciate the Dean, Faculty of Engineering, Built Environment and Information Technology; Professor S. Maharaj; the Head, Department of Electrical Electronic and Computer Engineering; Professor. J. Joubert, the Head of Advanced Sensors Networks Research Group; and all staffs who have contributed to the successful completion of the Ph.D. program.

To all my friends and colleagues who have in one way or the other shown support and encouragements during the Ph.D. programme, thank you and may Almighty GOD bless you all. Also, I wish to appreciate Mrs. Mari Ferreira for her continuous service throughout the Ph.D. journey. Thank you for all your assistance.

I would like to appreciate the following institutions: Telkom SA, Council for the Scientific and Industrial Research (CSIR); and Tertiary Education Trust Fund (TETFUND) for their financial support towards the course of my Ph.D. programme.

To my wonderful parents, Professor and Hajia H.O.B. Oloyede, I cannot thank you enough. All your prayers, efforts and support with kind words of encouragements have made me who I am today. May you live long to reap the fruit of your labour, Amin. To my siblings, Muyideen, Sherifdeen, Mubarak, and Najeeb, thank you for the prayers and words

of encouragement during the Ph.D. journey. Also, I thank the families of Dr. and Dr. (Mrs) A.O. Folorunsho for their support and words of encouragement shown throughout the Ph.D journey.

Finally, a special appreciation to my lovely wife Waliyat and beautiful kids Taheerah, Marzuq and Muqtadir Oloyede for the patience, love, and perseverance shown towards the success of the Ph.D. programme. I do not take this sacrifice for granted and, I pray Almighty Allah to grant me the opportunity to pay you back in multiple folds.

SUMMARY

A NOVEL FACE RECOGNITION SYSTEM IN UNCONSTRAINED ENVIRONMENTS USING A CONVOLUTIONAL NEURAL NETWORK

By

Muhtahir Oluwaseyi Oloyede

Promotor: Prof. G.P. Hancke

Co-promotor: Dr. H.C. Myburgh

Department: Electrical, Electronic, and Computer Engineering

University: University of Pretoria

Degree: Doctor of Philosophy (Computer Engineering)

Keywords: Biometric system, convolutional neural network, expression, face recognition system, pre-processing, image enhancement, unconstrained environments.

The performance of most face recognition systems (FRS) in unconstrained environments is widely noted to be sub-optimal. One reason for this poor performance may be due to the lack of highly effective image pre-processing approaches, which are typically required before the feature extraction and classification stages. Furthermore, it is noted that only minimal face recognition issues are typically considered in most FRS, thus limiting the wide applicability of most FRS in real-life scenarios. Thus, it is envisaged that developing more effective pre-processing techniques, in addition to selecting the correct features for classification, will significantly improve the performance of FRS.

The thesis investigates different research works on FRS, its techniques and challenges in unconstrained environments. The thesis proposes a novel image enhancement technique as a pre-processing approach for FRS. The proposed enhancement technique improves on the

overall FRS model resulting into an increased recognition performance. Also, a selection of novel hybrid features has been presented that is extracted from the enhanced facial images within the dataset to improve recognition performance.

The thesis proposes a novel evaluation function as a component within the image enhancement technique to improve face recognition in unconstrained environments. Also, a defined scale mechanism was designed within the evaluation function to evaluate the enhanced images such that extreme values depict too dark or too bright images. The proposed algorithm enables the system to automatically select the most appropriate enhanced face image without human intervention. Evaluation of the proposed algorithm was done using standard parameters, where it is demonstrated to outperform existing image enhancement techniques both quantitatively and qualitatively.

The thesis confirms the effectiveness of the proposed image enhancement technique towards face recognition in unconstrained environments using the convolutional neural network. Furthermore, the thesis presents a selection of hybrid features from the enhanced image that results in effective image classification. Different face datasets were selected where each face image was enhanced using the proposed and existing image enhancement technique prior to the selection of features and classification task. Experiments on the different face datasets showed increased and better performance using the proposed approach.

The thesis shows that putting an effective image enhancement technique as a pre-processing approach can improve the performance of FRS as compared to using unenhanced face images. Also, the right features to be extracted from the enhanced face dataset as been shown to be an important factor for the improvement of FRS. The thesis made use of standard face datasets to confirm the effectiveness of the proposed method. On the LFW face dataset, an improved performance recognition rate was obtained when considering all the facial conditions within the face dataset.

LIST OF ABBREVIATIONS

AHE	Adaptive histogram equalization
AMBE	Absolute mean brightness error
ANN	Artificial neural network
ATM	Automated teller machine
BPD	Brightness preserving dynamic fuzzy histogram equalization
BS	Biometric systems
CHE	Contrast limited adaptive histogram equalization
CNN	Convolutional neural network
CSO	Cuckoo search optimization
DB	Database
DLA	Dynamic link architecture
DT	Decision tree
EBGM	Elastic bunch graph matching
EF	Evaluation function
EHD	Edge histogram descriptor
EM	Entropic measure
FBRS	Face biometric recognition system
FERET	Face recognition technology
FR	Face recognition
FRGC	Face recognition grand challenge
FRS	Face recognition system
GA	Genetic algorithm
GF	Gabor filter

GFM	Geometric feature matching
HE	Histogram equalization
ICA	Independent component analysis
IG	Information gain
IIA	Image intensity adjustment
ILSVRC	ImageNet Large Scale Visual Recognition Competition
IMG	Image
JAFFE	Japanese facial expression
LBP	Local binary patterns
LCS	Linear contrast stretching
LDA	Linear discriminant analysis
LLIE	Low light image enhancement
MA	Metaheuristic algorithm
MATLAB	Matrix laboratory
MBS	Multimodal biometric system
MCSO	Munteanu with cuckoo search optimization algorithm
MG	Munteanu with genetic algorithm
MP	Munteanu with particle swarm optimization
NB	Naïve Bayes
ORL	Olivetti research laboratory
PCA	Principal component analysis
PHOG	Pyramid histogram of gradients
PSNR	Peak signal-to-noise ratio
PSO	Particle swarm optimization
ResNet	Residual network
RF	Random forest
SCface	Surveillance cameras face

SVM	Support vector machine
TF	Transfer function
UBS	Unimodal biometric system
YeCSO	Ye with cuckoo search optimization algorithm
YF	Yale face dataset

TABLE OF CONTENTS

CHAPTER 1 INTRODUCTION 1

1.1 BACKGROUND 1

1.2 PROBLEM STATEMENT 4

1.2.1 Context of the problem 4

1.2.2 Research gap 5

1.3 RESEARCH GOAL 6

1.4 RESEARCH QUESTIONS AND OBJECTIVES 6

1.5 APPROACH 7

1.6 RESEARCH CONTRIBUTIONS 8

1.7 RESEARCH OUTPUTS 9

1.8 OVERVIEW OF STUDY 9

CHAPTER 2 A REVIEW OF FACE RECOGNITION SYSTEMS 11

2.1 CHAPTER OBJECTIVES 11

2.2 AN OVERVIEW OF FACE BIOMETRIC RECOGNITION SYSTEMS 11

2.3 FACE RECOGNITION IN UNCONSTRAINED ENVIRONMENTS 13

2.3.1 Illumination 14

2.3.2 Pose variation 16

2.3.3 Expression 21

2.3.4 Plastic surgery 22

2.3.5 Aging 25

2.3.6 Occlusion 27

2.3.7 Low resolution 28

2.4 REVIEW OF FACE RECOGNITION TECHNIQUES 30

2.4.1 Pre-processing techniques 30

2.4.2 Feature extraction techniques 34
2.4.3 Feature classification techniques 39
2.5 FACE RECOGNITION BENCHMARK DATASETS 45
2.5.1 AR face dataset 45
2.5.2 Yale face dataset 46
2.5.3 Yale face dataset B 46
2.5.4 Labelled faces in the wild 46
2.5.5 Facial recognition technology 47
2.5.6 Olivetti research laboratory 48
2.5.7 Plastic surgery face dataset 48
2.5.8 MS-Celeb-1M 48
2.6 CHAPTER SUMMARY 49

CHAPTER 3 MATERIALS AND METHODS 51
3.1 COMPARATIVE ANALYSIS OF FACE RECOGNITION TECHNIQUES IN UNCONSTRAINED ENVIRONMENTS 51
3.2 FEATURE EXTRACTION TECHNIQUES 52
3.2.1 Local binary patterns 52
3.2.2 Gabor filter 52
3.2.3 Edge histogram descriptors 53
3.2.4 Pyramid histogram of gradients 53
3.3 FEATURE CLASSIFICATION TECHNIQUES 53
3.3.1 Random forest 54
3.3.2 Decision tree 54
3.3.3 Support vector machine 55
3.3.4 Naïve Bayes 55
3.3.5 Convolutional neural network 55
3.4 PROPOSED FACE IMAGE ENHANCEMENT METHOD IN UNCONSTRAINED ENVIRONMENTS 56
3.4.1 Transformation function 57
3.4.2 Evaluation function 58
3.4.3 Metaheuristic algorithms 61
3.4.4 Summary of the process involved in the proposed face image enhancement technique 64

3.5 PERFORMANCE EVALUATION AND DATA SAMPLES 66
3.5.1 Number of edges 67
3.5.2 Number of pixels in the foreground 67
3.5.3 Entropic measure 67
3.5.4 Peak-signal-to-noise ratio 67
3.5.5 Absolute mean brightness error 68
3.6 IMPROVING FACE RECOGNITION SYSTEMS BY USING A NEW IMAGE ENHANCEMENT TECHNIQUE, HYBRID FEATURES, AND THE CONVOLUTIONAL NEURAL NETWORK 68
3.6.1 Selected hybrid features 70
3.6.2 Feature classification technique 70
3.7 CHAPTER SUMMARY 71

CHAPTER 4 RESULTS 73
4.1 COMPARATIVE ANALYSIS OF FACE RECOGNITION TECHNIQUES IN UNCONSTRAINED ENVIRONMENTS 73
4.1.1 Recognition performance on the Yale face dataset 73
4.1.2 Recognition performance based on illumination, on the Yale face dataset . 74
4.1.3 Recognition performance based on expression on the Yale face dataset 75
4.1.4 Recognition performance on the AR face dataset 75
4.1.5 Recognition performance based on lighting, on the AR face dataset 76
4.1.6 Recognition performance based on expression, on the AR face dataset 76
4.1.7 Recognition performance based on upper face occlusion, on the AR face dataset 77
4.1.8 Recognition performance based on lower face occlusion on the AR face dataset 78
4.2 A NEW EVALUATION FUNCTION FOR FACE IMAGE ENHANCEMENT IN UNCONSTRAINED ENVIRONMENTS USING METAHEURISTIC ALGORITHMS. 78
4.2.1 Choice of CSO parameters 79
4.2.2 Evaluation of different metaheuristic algorithms 80
4.2.3 Comparison of different evaluation functions 85
4.2.4 Comparison of different image enhancement methods 87

4.3 IMPROVING FACE RECOGNITION SYSTEMS USING A NEW IMAGE ENHANCEMENT TECHNIQUE, HYBRID FEATURES, AND THE CONVOLUTIONAL NEURAL NETWORK 100
4.3.1 Choice of proposed hybrid features 101
4.3.2 Recognition based on constraints 103
4.3.3 Recognition performance on the LFW dataset 108
4.4 CHAPTER SUMMARY 110

CHAPTER 5 CONCLUSION AND FUTURE WORK 112

REFERENCES 117

LIST OF FIGURES

Figure 2.1. Stages involved in the FBRS 13
Figure 2.2. Facial changes due to illumination. 15
Figure 2.3. Facial **c**hanges due to pose variation. 19
Figure 2.4. Facial changes due to expression. 21
Figure 2.5. Facial changes due to plastic surgery. 23
Figure 2.6. Facial changes due to age difference 26
Figure 2.7. Samples of occluded face images 27
Figure 2.8. Samples of constructed Eigenface 35
Figure 2.9. Samples of Fisherface. 36
Figure 2.10. Local binary pattern operator. 37
Figure 2.11. The EBGM technique 38
Figure 2.12. A CNN architecture for face recognition. 41
Figure 3.1. Process involved in the comparative analysis of the different FRS. 51
Figure 3.2. Process involved in the proposed image enhancement of facial images 57
Figure 3.3. Graphical flowchart representation of the proposed image enhancement method. 69
Figure 3.4. The framework of proposed face recognition system 70
Figure 4.1. Choice of CSO parameter 79
Figure 4.2. Performance and convergence analysis of the different metaheuristic algorithms on Image 1 80
Figure 4.3. Performance and convergence analysis of the different metaheuristic algorithms on Image 2 81
Figure 4.4. Performance and convergence analysis of the different metaheuristic algorithms on Image 3 82
Figure 4.5. Qualitative comparison of the different metaheuristic algorithm with the proposed evaluation function on Images 1 (a) original, (b) CSO + proposed evaluation method, (c) PSO + proposed evaluation method and (d) GA + proposed evaluation method. 84
Figure 4.6. Qualitative comparison of the different metaheuristic algorithm with the proposed evaluation function on Images 2 (a) original, (b) CSO + proposed evaluation method, (c) PSO + proposed evaluation method and (d) GA + proposed evaluation method. 84

Figure 4.7. Qualitative comparison of the different metaheuristic algorithm with the proposed evaluation function on Images 3 (a) original, (b) CSO + proposed evaluation method, (c) PSO + proposed evaluation method and (d) GA + proposed evaluation method. 85
Figure 4.8. Qualitative comparison of the different evaluation functions with the CSO algorithm on Images 1: (a) original, (b) MCSO, (c) YeCSO, (d) Proposed + CSO..... 86
Figure 4.9. Qualitative comparison of the different evaluation functions with the CSO algorithm on Images 2: (a) original, (b) MCSO, (c) YeCSO, (d) Proposed + CSO..... 87
Figure 4.10. Qualitative comparison of the different evaluation functions with the CSO algorithm on Images 3: (a) original, (b) MCSO, (c) YeCSO, (d) Proposed + CSO..... 87
Figure 4.11. Qualitative comparison of the different image enhancement algorithms on the AR face dataset where Figures 1 – 6 represent images of different subjects respectively; and a – j denotes the methods labelled as (a) original, (b) LCS (c) HE (d) IIA (e) MP (f) MG (g) BPD (h) CHE (i) LLIE (j) Proposed..... 97
Figure 4.12. Qualitative comparison of the different image enhancement algorithms on the Yale face dataset where Figures 1 – 6 represent images of different subjects respectively; and a – j denotes the methods labelled as (a) original, (b) LCS (c) HE (d) IIA (e) MP (f) MG (g) BPD (h) CHE (i) LLIE (j) Proposed..... 98
Figure 4.13. Qualitative comparison of the different image enhancement algorithms on the ORL face dataset where Figures 1 – 6 represent images of different subjects respectively; and a – j denotes the methods labelled as (a) original, (b) LCS (c) HE (d) IIA (e) MP (f) MG (g) BPD (h) CHE (i) LLIE (j) Proposed..... 99
Figure 4.14. Average recognition performance based on lighting. 104
Figure 4.15. Average recognition performance based on expression..... 105
Figure 4.16. Average recognition performance based on occlusion. 106
Figure 4.17. Average recognition performance based on lighting. 107
Figure 4.18. Average recognition performance based on expression..... 108
Figure 4.19. Qualitative comparison of the different image enhancement algorithms on the LFW face dataset where Figures 1 – 4 represent images of different subjects respectively; and a – h denote the methods labelled as (a) original, (b) LCS (c) HE (d) AHE (e) IIA (f) MP (g) MG (h) Proposed..... 110

LIST OF TABLES

Table 2.1. A summary of recent face recognition systems for addressing the issue of illumination..... 18
Table 2.2. A summary of recent face recognition systems to address the issue of pose variation. 20
Table 2.3. A summary of recent face recognition systems for addressing the issue of expression. 24
Table 2.4. A summary of recent face recognition systems for addressing the issue of plastic surgery. 25
Table 2.5. A summary of recent face recognition systems for addressing the issue of aging 26
Table 2.6. A summary of recent face recognition systems for addressing the issue of occlusion. 28
Table 2.7. A summary of recent face recognition systems for addressing the issue of low resolution. 29
Table 2.8. Face dataset and conditions in uncontrolled environment. 49
Table 4.1. Recognition performance on Yale. 74
Table 4.2. Recognition performance on Yale, based on illumination. 74
Table 4.3. Recognition performance on Yale, based on Expression. 75
Table 4.4. Recognition performance on AR 76
Table 4.5. Recognition performance on AR, based on lighting..... 76
Table 4.6. Recognition performance on AR, based on expression. 77
Table 4.7. Recognition performance on AR, based on upper face occlusion 77
Table 4.8. Recognition performance on AR, based on lower face occlusion 78
Table 4.9. Comparison of the different metaheuristic algorithms with the proposed evaluation function based on all the performance evaluation metrics..... 83
Table 4.10. Quantitative Comparison for different evaluation functions 85
Table 4.11. Number of pixels in the foreground value obtained with different enhancement methods on the AR face dataset. 88
Table 4.12. Number of pixels in the foreground value achieved by using different enhancement methods on the Yale face dataset. 89
Table 4.13. Number of pixels in the foreground value when different enhancement methods on the ORL face dataset were used. 89

Table 4.14. Number of edges obtained by using different enhancement methods on the AR face dataset 90
Table 4.15. Number of edge values obtained by using different enhancement methods on the Yale face dataset. 90
Table 4.16. Number of edge values obtained by using different enhancement methods on the ORL face dataset..... 91
Table 4.17. PSNR values achieved by using different enhancement methods on the AR face datasets. 91
Table 4.18. PSNR values achieved by using different enhancement methods on the YF face datasets. 92
Table 4.19. PSNR values achieved by using different enhancement methods on the ORL face datasets. 92
Table 4.20. Entropic measure values achieved by using different enhancement methods on the AR face dataset. 92
Table 4.21. Entropic measure values achieved by using different enhancement methods on the Yale face dataset. 93
Table 4.22. Entropic measures values achieved by using different enhancement methods on the ORL face dataset..... 93
Table 4.23. AMBE values achieved by using different enhancement methods on AR face datasets. 93
Table 4.24. AMBE values achieved by using different enhancement methods on Yale face datasets. 94
Table 4.25. AMBE values achieved by using different enhancement methods on ORL face datasets. 94
Table 4.26. Fitness value achieved by using different enhancement methods on the AR face datasets. 95
Table 4.27. Fitness value achieved by using different enhancement methods on the YF face datasets. 95
Table 4.28. Fitness value achieved by using different enhancement methods on the ORL face datasets. 96
Table 4.29. Performance of the proposed hybrid feature on the AR dataset based on different image enhancement methods. 101
Table 4.30. Performance of the proposed hybrid feature on the Yale dataset based on different image enhancement methods. 103

Table 4.31. Recognition performance of the 18-Layer ResNet CNN architecture on the LFW dataset using different features and image enhancement methods. 109

CHAPTER 1 INTRODUCTION

1.1 BACKGROUND

Biometric systems (BS) have been in existence for several decades and have been used successfully for identification and verification [1]. The process involved in the use of BS involves registering a biometric trait on a dataset, which is used for further recognition [1]. A BS relates a biometric trait with a registered biometric profile associated with a particular subject that seeks to find a match. An identification system associates biometric features in order to identify an unknown person or a biometric entity in an existing dataset. A BS achieves identification when it makes a one-to-many (1:N) search while verification is made based on a one-to-one (1:1) search [2, 3].

The BS uses a subject's human body characteristics that tend to be consistent for a long period. The characteristics of a reliable biometric system should have the following properties: acceptability, uniqueness, permanence, and universality[4]. There are different modalities of the biometric systems that are based on the physiological and behavioural features of human beings that do not usually change over time. The behavioural biometric types include voice, gait and hand-written recognition systems. The physiological biometric system types also include fingerprint, hand geometry, ear, vein, iris, retina and face recognition systems (FRS) [5-7]. A major advantage of the BS for human identification and verification rests on the fact that the biometric characters exhibit an effective link with a subject based on the uniqueness, which can hardly be shared or forged. BS prevents wrong identity claims from being made by use of genuine subjects to participate in the use of the system [8].

Patterns such as face, fingerprints and iris are applied in areas such as border control, mobile phones authentication; and time management and attendance systems [9]. A BS delivers two vital functionalities: verification and identification. In verification, the system authenticates a subject to gain approval [10]. For instance, a subject claims an identity based on a private identification number while using a computer system. The verification system conducts a 1:1 matching because it uses the subject's biometric data and compares it with a certain name in the dataset. The claim will be authentic if the input from the subject is reasonably similar to the identity's registered records; however the system terms the subject as an impostor if the claim is denied [11]. Therefore, during verification, a biometric system indicates that ''you are whom you say you are''. However, in identification, a subject tries to claim that he/she is one of the registered subjects in the dataset. To identify, the system validates the subject's input biometric traits to find the subject with the highest similarity index [12]. Furthermore, when the system confirms that the input and the registered template are different, it rejects a decision that is below a given threshold to identify an impostor. Thus, identification involves a one-to-many matching process. Consequently, an identification system indicates that a subject is ''someone who is known to the system'' [13].

Biometric systems are classified as either unimodal or multimodal BS. A unimodal biometric system (UBS) uses only one biometric trait [14] or one biometric data source for identification and verification. UBSs have become more accurate and reliable; however, they are limited by noisy data at the enrolment process [15]. A typical instance is an issue with FRS that deal with facial images, where faces of the same individual can be seen differently by the system due to real-world situations. On the other hand, Multimodal biometric systems (MBS) describe a system that combines more than one biometric feature for identification and verification.

The FRS has been shown to be more efficient and effective as compared to other biometric traits because of its unique structure and characteristics that are made up of unique features located at different parts of the human face [16]. These characteristics present in the face are different and unique amongst individuals and are further used for identifying and

verifying individuals. The growth and the existence of the FRS have been studied by various researchers in both academia and industry in an effort to improve on the system which is seen as both an image processing and computer vision tasks [17]. These techniques that have been developed and designed over time are proven to be effective in normal face recognition (FR) conditions. Normal FR conditions or scenarios occur when the properties of the facial image registered in the dataset are the same as the probe image, for examples pixel size, position regarding pose, background and facial expression. However, most of the face recognition techniques developed do not perform optimally in real-life scenarios also termed unconstrained environments [18].

These are situations that make the facial image of an individual stored in the dataset look entirely different from the test face image. The following are the factors that cause FRS to fail, which are also seen as the issues of FRS, i.e., illumination, expression, pose variation, low resolution, occlusion, aging, and plastic surgery [19]. Though recent studies in the area of FR have shown great interest in trying to improve algorithms in such scenarios, yet optimal performance in these conditions have not yet been attained.

There have been developments in various face datasets that include these face constraints for testing different algorithms developed by researchers [19]. Like other biometric modalities, crucial stages of the FRS include the pre-processing, feature extraction and classification stages. It is in these stages that the various algorithms or techniques developed are utilized. A significant concern is that the fact that an appropriate pre-processing approach is not put in place in most FRS models. Also, appropriate feature extraction techniques and classification algorithms are not adequately selected to give optimal performance in FR in real-life scenarios.

The pre-processing stage of the FRS is a significant and vital block of the FRS model where image enhancement which is overlooked in the literature supposedly plays a vital role. Hence, having a proper face image enhancement technique in place increases the FRS performance in unconstrained environments, which is the primary focus of this thesis.

In this regard, and firstly in this thesis, a detailed comparative analysis of different FRS was carried out to confirm their performance in unconstrained environments. This approach enables us to observe which features perform better with a classification technique to produce a better recognition performance in selected unconstrained environments. This involves the combination of different feature and classification methods that make up the FRS. The results obtained allowed for the proper selection of methods used in this thesis.

Secondly, a new image enhancement technique based on a metaheuristic algorithm is presented as a pre-processing approach to facial images in unconstrained environments. The proposed image enhancement techniques consist of a new evaluation function that allows for the selection of the most appropriate enhanced face image without human intervention. This results in the addition of more features to the facial images affected by different conditions of the unconstrained environments. The proposed method outperforms several image enhancement methods both quantitatively and qualitatively.

Furthermore, a FRS model is presented. The FRS model comprises a new image enhancement method at the pre-processing stage and a new selection of hybrid features that can extract features from the enhanced facial image effectively. Results have shown that putting these approaches in place has increased the performance of the FRS using the convolutional neural network (CNN) classification method.

1.2 PROBLEM STATEMENT

1.2.1 Context of the problem

It is envisaged that having an effective image enhancement technique by proposing a new evaluation function at the pre-processing stage of the FRS model can improve its performance in unconstrained environments. Although researches and studies have been carried out on designing FRS, most do not perform efficiently in unconstrained environments or real-life conditions. These conditions include lighting, occlusion,

expression, and plastic surgery amongst others [26]. A primary reason for not achieving satisfactory performance in such conditions is failure to put in place an effective pre-processing approach where image enhancement plays a vital role.

Current image enhancement techniques do not automatically select the appropriate enhanced face image, and therefore requiring human intervention. Also, existing image enhancement techniques for face images produce either too dark or too bright images due to over-enhancement, which further affects the performance of FRS [169].

In this thesis, the problem statement, is to identify and offer a solution to the problem of putting in place an effective image enhancement technique as a pre-processing approach that can assist in enhancing recognition performance in unconstrained environments. In this thesis, it is conjectured that knowledge of the right features to be identified and extracted from the enhanced face dataset will result in improved recognition performance. Hence, the term ‘an effective pre-processing approach for face biometric recognition systems in unconstrained environments,’ which clearly defines and describes the problem, is the emphasis of this thesis.

1.2.2 Research gap

FRS have performed satisfactorily in normal conditions. However, FR in unconstrained environments or real-life scenarios remains a lingering issue. This has resulted in the failure of different FR applications such as surveillance and, security, which promotes terrorist attacks [24]. Due to this, algorithms and techniques have been proposed in an attempt to solve these issues faced by the biometric recognition system.

Most FR approaches focus more on the feature extraction and classification stages, and consequently neglect certain pre-processing steps that can enhance FRS. Face alignment has in recent times gained attention due to its comprehensive application in automatic face analysis, although, it has been shown to be extremely challenging in unconstrained environments [27]. Also, face image enhancement is another pre-processing step that

improves the performance of FRS. A major challenge for the process is the ability to come up with an effective image enhancement method for face images in unconstrained environments [176]. In addition, extracting the right features at the feature extraction stage for effective classification is regarded as an issue [28].

1.3 RESEARCH GOAL

The aim of this thesis is to propose an effective and novel image enhancement technique to enhance the performance of FRS in unconstrained environments by using the convolutional neural network. Effective methods are also presented to extract features from the enhanced facial images that can lead to improved recognition performance.

1.4 RESEARCH QUESTIONS AND OBJECTIVES

The objectives and questions of this research accomplished the research goal stated in Section 1.3 are:

Research Questions

[1] Which feature methods and classification technique can perform well in the selected unconstrained environments?

[2] Can a new image enhancement technique as a pre-processing approach in FRS be developed?

[3] Can parameters be identified to confirm the performance of the image enhancement technique at the pre-processing stage of FRS?

[4] Which features can be extracted from the enhanced face image to improve recognition performance?

[5] Does introducing image enhancement as a pre-processing technique improve the performance of FRS?

Research Objectives

[1] To conduct a comparative analysis of different FRS to enable us to confirm their respective performances in unconstrained environments.

[2] To propose a novel image enhancement technique as a pre-processing approach for FRS.

[3] To identify different parameters that can confirm the performance of image enhancement technique at the pre-processing stage of FRS.

[4] To propose the right features to be extracted from the enhanced face image so as to improve recognition performance.

[5] To confirm if designing and presenting an image enhancement technique at the pre-processing stage, will improve the recognition performance of FRS.

1.5 APPROACH

FR in unconstrained environments remains a research concern where efforts have been made by various researchers in both academia and industry. Hence, many algorithms and techniques have been developed [21]. However, it is still unknown which of these techniques perform optimally in certain unconstrained conditions. The thesis has followed the standard approach for deploying technical research in the area of computer engineering, particularly in image-processing and computer vision. Also, to enable the deployment of a FRS that can perform optimally regardless of the facial constraints, this work is structured as follows:

- **State-of-the-art review**: The research started with an in-depth review and study of biometric systems, and then the various modalities of biometrics were studied, and various research gaps were identified. An in-depth survey of FRS was also done, where all current techniques were critically reviewed; the issues with FRS were highlighted while the face datasets containing these constraints were also considered. The objectives of this survey were to have a good and sound foundation for the subject area, to identify research gaps to provide a focus and direction for the thesis. The outcome from this stage was a substantial amount of information which is detailed in the second chapter of the thesis.

- **Simulation and numerical analysis**: For the first technical part of this thesis, a comparative analysis of different features and classification methods was carried out. For the implementation of the proposed image enhancement technique, we used the Matrix laboratory (MATLAB) software. Finally, simulations were done to confirm the effectiveness of the proposed image enhancement technique and the selection of new hybrid features, thus, we came up with a novel approach that can improve the performance of FRS in unconstrained environments.
- **Verification and validation of results**: Results obtained from the simulations were validated by numerical analysis. These results were further validated and verified by comparing them with results from related work in the literature.

1.6 RESEARCH CONTRIBUTIONS

The thesis makes the following contributions to the body of scientific knowledge:

[1] A comparative analysis of various FR techniques has been carried out; to observe the performance of these techniques in unconstrained environments.

[2] A new face image enhancement technique based on metaheuristic algorithms is presented as a pre-processing approach to facial images in unconstrained environments.

[3] A novel evaluation function is proposed as a component within the image enhancement technique, such that appropriately enhanced face images are selected and produced. These face images are stored and then used for feature extraction and classification purpose.

[4] A defined scale mechanism is presented for evaluation of the enhanced images, where extreme values depict either too dark or too bright images.

[5] A set of new selected hybrid features to be extracted from the enhanced face images is presented to improve recognition performance.

[6] The proposed image enhancement technique can be of benefits to other image processing applications such as satellite and medical imaging.

1.7 RESEARCH OUTPUTS

[1]. M.O. Oloyede, and G.P. Hancke, "Unimodal and Multimodal Biometric Systems: A Review", *IEEE Access*, vol. 4, pp. 7532 – 7555, 2016.

[2]. M.O. Oloyede, and G.P. Hancke, "A Robust Face Recognition System for Solving Occlusion", *SATNAC Conference*, Western Cape, South Africa, 2016.

[3]. M.O. Oloyede, G.P. Hancke, and N. Kapileswar, "Evaluating the Effect of Occlusion in Face Recognition Systems", in *Proc. of the IEEE Africon Conference*, Cape Town, SA, September 2017, pp. 18 – 20.

[4]. M.O. Oloyede, G.P. Hancke, and H.C. Myburgh, "Improving Face Recognition Systems Using a New Image Enhancement Technique, Hybrid Features, and the Convolutional Neural Network", *IEEE Access*, vol. 6, pp. 75181 – 75191, 2018.

[5]. M.O. Oloyede, G.P. Hancke, H.C. Myburgh and A.J. Onumanyi, "A New Evaluation Function for Face Image Enhancement in Unconstrained Environments Using Metaheuristic Algorithms" *EURASIP Journal on image and video processing*, vol. 27, pp. 1 – 18, 2019.

[6]. M.O. Oloyede, G.P. Hancke, and H.C. Myburgh. A Review on Face Recognition Systems: Techniques and Challenges. (**Under review**).

1.8 OVERVIEW OF STUDY

The rest of the thesis is structured as follows:

Chapter 2 details the foundational knowledge of the research area through a well-presented review of FRS in unconstrained environments. In this chapter, recent and related works on FRS in unconstrained environments are critically examined and analyzed. The review work presents and discusses the various face recognition techniques, while also stating their pros and cons. The issues faced by the face recognition techniques that do not allow satisfactory performance, i.e., real-life conditions or unconstrained environments, are also discussed. The review has helped in identifying the gap in the pre-processing stage of the FRS model.

It also assisted in ordering and aligning the structure of the thesis as the remainder of the studies conducted are presented in subsequent Chapters.

In Chapter 3, an extensive description is presented of the materials, methods, and approach used to achieve the various objectives of this thesis. First, the approach to compare different FRS in a real-world scenario is presented. Secondly, the approach to designing the proposed face image enhancement method is described. This also includes the approach to designing an intelligent scaling mechanism that guarantees the output of an enhanced face image. Finally, the approach to improving the FRS with the use of the image enhancement method, new selection of hybrid features and the convolutional neural network is described.

Chapter 4 presents and discusses the results from the different experiments carried out in the thesis. Also, the main observations from the thesis are provided. The discussion begins with a comparative analysis of the recognition performance of different face recognition techniques in unconstrained environments. Then, the proposed face image enhancement technique in unconstrained environments is discussed. In this chapter, the proposed FRS model in unconstrained environments is also discussed.

Finally, Chapter 5 concludes this thesis and highlights possible areas for future work.

CHAPTER 2 A REVIEW OF FACE RECOGNITION SYSTEMS

2.1 CHAPTER OBJECTIVES

For an FRS to be considered effective and offering optimal performance, it must be able to perform the recognition process regardless of any constraints, i.e., in real-life scenarios or unconstrained environments [23]. Therefore, a substantial amount of research work has been done into investigating and designing models with various techniques for FRS. However, the efforts put forward by the researchers have not yet been able to solve the issues faced by FRS. As a way of addressing these problems, this chapter presents a broad study of the various techniques developed for FRS in unconstrained environments. The major issues of FRS also termed unconstrained environments, are discussed and analyzed. It also presents the different approaches researchers have used in designing FRS in unconstrained environments. The different standard and benchmark face datasets that represent the different facial conditions in unconstrained environments are also presented and analyzed. Hence, the study gives the right guidance for the further approach to designing FRS in unconstrained environments.

2.2 AN OVERVIEW OF FACE BIOMETRIC RECOGNITION SYSTEMS

The Face recognition biometric system (FBRS) has been a familiar part of biometric systems since its inception decades ago. This has brought about research outputs in the form of technical reports, scholarly articles and book chapters published over time. Hence, it is assumed that many of the readers of this thesis are familiar with the basics of FBRS;

however, it is also reasonable to assume that there are some readers who are not well grounded in the concept. Furthermore, it is essential to highlight the reasons why the subject remains an active research area. This section will benefit different categories of readers because it provides an extensive and up- to-date overview of the concept involved in the FBRS.

FR is an active research area due to the various applications where it can be applied, such as border security, surveillance, law enforcement and access control [20-23]. Recently, other applications of the FR system have included computer graphics, neural networks, and psychology as it has lately been generating more multidisciplinary interest. As in other biometric systems, as shown in Figure 2.1, the stages involved in the FR process are face detection, pre-processing of the face image, extraction of facial features and lastly feature classification [24, 25]. The first stage, which is the detection of the face, is the process whereby the FRS verifies the face in an image or video. After detection, the facial image is pre-processed to identify the region of interest and improve the image quality. Normalization is a type of pre-processing technique where the face images of different scales are transformed and mapped into the same scale. Face alignment is another pre-processing approach in which the fiducial points such as mouth, eyes chin and nose are localized. This approach is seen to improve FRS, though it remains a challenge and an open problem in unconstrained environments [26-28].

Image enhancement is also a pre-processing approach that is overlooked in the literature. Its primary objective is to come up with an enhanced face image from the original, which is supposed to improve the overall performance of the FRS [29-31]. Feature extraction is the next stage of the FRS model, and its purpose is to simplify the number of resources that describe a large set of data. Unique features are also extracted to minimize noise and irrelevant information present in the original face image, and a feature vector that is sufficient to describe the face is extracted from the face image. Many feature extraction methods exist, however, selecting the right features for various FRS in unconstrained environments remains a challenge [32-35]. Most common classifiers used in the field of FR are the minimum distance classifier, nearest and k-nearest neighbour classifier [36].

The minimum distance classifier places the label of a testing sample as the class with whose mean it is associated. The nearest neighbour classifier places testing samples in the class with which its nearest neighbour is associated, while the k – nearest neighbour classifier places the testing samples with the class that has the most nearest neighbours by first searching for the k-nearest neighbour [37].

In recent times, machine learning algorithms have been the preferred and most used classification techniques, for example, a convolutional neural network. The feature classification stage that leads to the recognition of face images involves both identification and authentication. Identification compares a face image with other face images to be able to come up with an identity of the face among several possibilities, while authentication occurs when a face is compared with another to approve the requested identity [38]. In both scenarios, face images of known individuals are registered in the system known as a gallery. After that, face images are used as probes that can either be those of the registered or unregistered individuals are used.

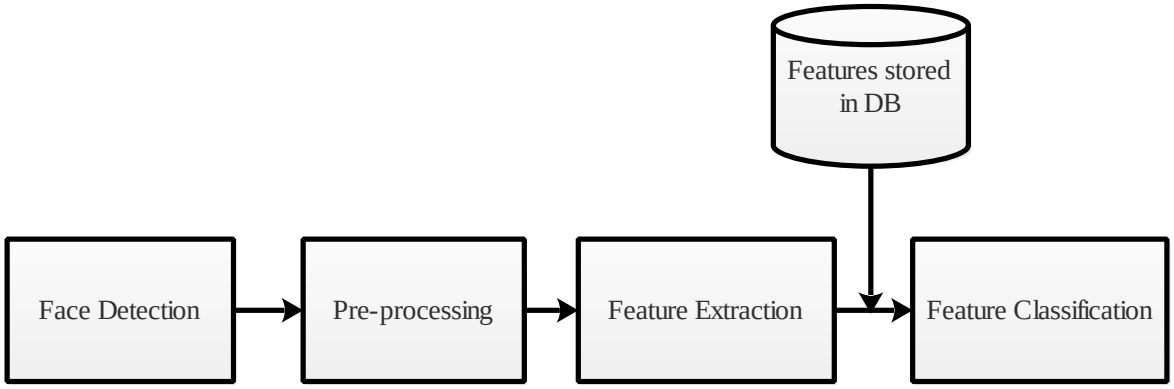


Figure 2.1. Stages involved in the FBRS

2.3 FACE RECOGNITION IN UNCONSTRAINED ENVIRONMENTS

FR is regarded as a complex biometric system because of its constraints caused by changes in the appearance of facial images [26, 39, 40]. The various face recognition techniques have recorded remarkable success in recognition performance in well-controlled

environments. However, these techniques tend to fail as a result of the non-stable structure of facial images and also in realistic scenarios as these make the facial images of the same person look different [24]. At present, most face recognition techniques still have to achieve optimal recognition due to these issues, as the face acquisition procedure varies widely [41, 42]. The issues of the FRS include illumination, pose variation, expression, plastic surgery, aging and most importantly occlusion [43]. This chapter further describes the different issues which are also seen as the FRS in unconstrained environments or FR in the wild. A detailed presentation of a recent and different approach to each issue is also presented.

2.3.1 Illumination

Illumination is referred to as lighting variations where a face image can appear differently due to change in lighting [44-46]. As shown in Figure 2.2, this change in lighting appears greater than the differences amongst individuals, hence causing FRS to fail by misclassifying faces when comparing images [47, 48]. Illumination remains an issue for robust FRS that involve extracting illumination invariant features. Illumination has great effects on an image which results in change of location, shadow shape, and reversal of contrast gradients [49]. It is normal for humans to still recognize face images regardless of changes in lighting; however, this task remains an issue for FRS. Earlier research using various image representations such as grey-scale, edge maps, filtered images with a Gabor filter showed that changing the sides of lighting conditions may cause greater variations in images than changing the identity of the face [50].

Furthermore, testing or training is also seen to be sensitive under varying illumination conditions [49, 51]. This has caused the FRS to fail in an illumination scenario and has over the past decade attracted researchers in the field to come up with various algorithms to solve this problem. Algorithms proposed to handle the issue of illumination follow different approaches. First, they use image-processing techniques to normalize face images affected by various lighting environments [47]. Consequently, Gamma intensity correlation, logarithm transform or histogram equalization in the form of enhancement

perform well under various lighting environments. Secondly, the 3D-face approach has been used for the issue of illumination [52].

Frontal face images under various lighting conditions result in an illumination cone formed on the subspace. It is quite easy to identify this in the low dimensional subspace by applying a generative model to the training data. However, the 3D-face model approach needs more training samples, and the light source may have to be specified, which is not an ideal process in a real-life scenario. The third approach involves processes where features of the parts where the illumination occurs are extracted, which are passed on the recognition stage [53]. To date, algorithms have been proposed for these approaches; however, research is still being done to achieve optimal performance. Table 2.1 shows a summary of recent approaches to the issue of illumination.

Figure 2.2. Facial changes due to illumination (Taken from [54], with permission).

In [55], two methods based on local histogram specifications to pre-process face images affected by different lighting conditions were proposed. The idea behind the approach was to remove illumination with high and low-frequency sections present in the facial image effectively, and to enhance facial features present in the lower-frequency section.

First, a high pass filter to remove low-frequency illumination in the face image was applied. Local statistics of the histogram to the entire face image was applied and these are learnt from normal lighting images. In the second approach, it is assumed that the regions have contain a higher-frequency lighting and that local histogram statistics know weak facial features.

Then the local histogram specification is applied to these sections to reduce high-frequency illumination and enhance weak facial features. For the experiment, three standard face datasets were used, i.e., CMU PIE dataset, CAS-PEARL dataset, and the Extended Yale B dataset. To further confirm the effectiveness of the methods various distance metrics were used such as distance transformation, correlation, cosine, and histogram distance. Considering CMU PIE dataset, the recognition rate was satisfactory using the cosine distance metrics, correlation, histogram distance, and distance transformation.

In [52], an adaptively weighted ULBP_MPHOG and WSRC method was proposed. First, the face images were normalized from which a uniform local binary pattern and multiple histograms of oriented features were extracted in each block. Information entropy was then used to achieve the adaptively weighted ULBP_MHOG features. Experiments were carried out using the convolutional neural network (CNN) on the ORL, CMU PIE, Yale and Extended Yale face datasets based on the varying number of blocks, types of feature and different classifiers. Experiments on the ORL dataset were carried out by using 4 x 4 blocks and six training samples for each class. On the Yale dataset, a recognition rate of 87.68% was achieved by using 10 x 10 blocks and 3, 4 and 5 training samples per class. Finally, on the Yale B dataset, the recognition rate was shown to be better when 16 x 16 blocks were used. This, achieved a satisfactory recognition rate.

In [56], an extraction model to solve the illumination issue based on robust discriminative multi-layer illumination features was proposed. A linear combination to decompose the multi-layer features into smaller robust features was utilized. Weights were assigned to each layer to fine-tune its relevance in order to use this relevant information in large-scale features for FR. Experiments conducted on the FRGC dataset demonstrated satisfactory recognition rates. Table 2.1 presents a summary of the FRS to address illumination issues.

2.3.2 Pose variation

The performance of a FRS also tends to fail when there is pose variation in the input images as shown in Figure 2.3. In practices such as passport control, face images are

mostly from a frontal view. However, in uncontrolled environments where face images can appear at various angles due to the rotation, this can cause the FRS to fail [57-61].

Pose variation refers to face images presented in different poses; humans might find it a simple task to recognize them, however; it remains an issue for computers, most especially computer vision applications such as surveillance [54]. With the increase in the rate of terrorist attacks all around the globe, most airports are equipped with surveillance cameras. Face images of terrorists are captured and stored and later compared with the face images of genuine travellers as their faces will be scanned. These images of the terrorist must have been captured in a different pose at an earlier time to simplify recognition when simple face recognition techniques are used. In real-life situations where the facial images across various poses are unavailable in the dataset, the system will not perform optimally as the different poses of the same individual will appear as a non-match [52].

When pose tolerance is absent from a FRS, the system becomes non-passive and non-intrusive. Hence, there is the desired need for FRS to work optimally under different poses which may be different from the face image stored in the dataset. A major issue faced by the various face recognition techniques in a pose variation situation is to acquire unique features free from pose variation.

In situations where the gallery of images stores face images in different poses, a particular face image will be recognized better even in a different pose [65]. This has been shown in the literature, where different face recognition techniques performed better when multiple face images per individual were stored in the dataset as compared to just one face image. This improvement in performance results from the fact that the techniques can tolerate minor pose variations, i.e., as there is an increase in the number of gallery images, there is a high probability that the probe pose will be close to an image in the gallery [66]. Adding more side view images to the gallery provides more required information of the human facial structures, which leads to models with better reconstruction than single gallery images.

Table 2.1. A summary of recent face recognition systems for addressing the issue of illumination.

Technique	Research work	Face dataset	Comments
Local histogram specification	Liu *et al.*, [55]	Extended Yale	Removal of both high and low-frequency sections of illumination present on the face image and enhancing facial features present in the lower frequency section.
Adaptively weighted ULBP_MPHOG and WSRC method with Convolutional neural network	Wang *et al.*,.[62]	CMU-PIE	Face images are normalized from which uniform LBP and multiple histograms of oriented features are extracted in each block. Information entropy is then used to achieve the adaptively weighted features produced by the ULBP_MHOG.
Discriminative multi-layer illumination robust feature extraction	Yu *et al.*,.[56]	Extended Yale	Large-scale features are sectioned into smaller parts, and weights are given to every layer to fine-tune its significance.
Sparse error of robust PCA	Luan *et al.*,.[63]	AR	The sparse error component displays more critical information beneficial to face identification. Weighted and ratio-based methods are utilized to distinguish facial images.

Figure 2.3. Facial **c**hanges due to pose variation (Taken from [64], with permission).

Most FR datasets contain a few gallery images such as police mug shot images that have a frontal image. Hence, a gallery with various pose images restricts the applicability of the techniques, and the ideal situation is meant to identify a test image from a single gallery image with random pose [67]. This is seen to be more challenging than a multiple gallery view. Various face recognition techniques have been proposed and used on benchmark face datasets where there are pose variation constraints. Benchmark face datasets that have pose variation constraints include the FERET face dataset, CMU-PIE face dataset, ORL face dataset, WVU dataset, MIT face dataset, Yale B face dataset. Of these listed datasets, FERET and CMU-PIE have either more face images or pose variations than other face datasets [68]. The various face recognition techniques have not performed optimally on these benchmark datasets.

In [67], a local binary pattern-like feature extraction method that adjusts the code rule Huffman was proposed. Also, a strategy based on divide and rule is applied to both the face representation and classification with the purpose of improving recognition performance with different pose positions. Experiments carried out on the CMU PIE dataset showed a recognition rate of 84.85%, while those on FERET showed a recognition rate of 85.17%.

In [68], an approach that uses landmarks of facial images and depth warping for effective cross-pose FR was proposed. Different from current 3-D reconstruction, the spontaneously identified broad facial landmarks to change the computationally expensive 3-D reconstruction procedure was utilized. Experiments were carried out on the PIE and Multi-PIE datasets and the results were satisfactory.

In [32], a highly effective pose-invariant FRS was proposed. An homography based normalization technique was used, with the approach covering the full range of pose variations within 90 degrees of yaw. Experiments carried out on public datasets showed satisfactory results in constrained and unconstrained environments. In [69], a convolutional neural network approach for head pose estimation was proposed, where the performance of different network architectures was measured. The adaptive gradient methods were utilized and led to satisfactory performance on wild datasets. Table 2.2 presents a summary of the FRS for addressing the issues of pose variation.

Table 2.2. A summary of recent face recognition systems to address the issue of pose variation.

Technique	**Research work**	**Face dataset**	**Comments**
Local binary pattern enhanced by divide and rule strategy	Zhou *et al.*,.[67]	FERET	The local binary pattern features change the code rule of Huffman. A divide and rule strategy is applied to the face classification for recognition.
3-D reconstruction procedure based on facial landmarks and sparse regression	Hsu *et al.*, [68]	Multi-PIE	Broad facial landmarks are used to change the computationally expensive 3-D reconstruction process. For matching purpose, registered depth-warped faces in the dataset are rotated to match the position of the probe image.
Pose-invariant FR with homography-based normalization	Ding *et al.*, [32]	Multi-PIE	Homography based normalization approach was used to design an effective pose invariant FR. The proposed approach covers the full range of pose variation within 90 degrees of yaw.
Convolutional neural network	Patacchiola *et al.*, [69]	Life in the wild	Fusing dropout and adaptive gradients increases recognition performance.

2.3.3 Expression

Human beings can continually display various facial expressions unless the face image is in a static mode [70-72]. These expressions are used to represent different emotions and mental states of individuals as shown in Figure 2.4. Also, the issue of expression in FR is not only seen as an identity or verification problem but also used for medical applications where a particular facial expression can be linked to a particular ailment [73].

Figure 2.4. Facial changes due to expression (Taken from [75], with permission).

It has been found that 7% of all the information an individual expresses is passed through language, speech represents 38%, and facial expression represents 55% [74]. From this, we can deduce that a substantial amount of valuable information can be gathered to detect an individual's consciousness and mental activities. Facial expression results in a change in the facial appearance and geometry leading to reduced recognition performance. This has motivated researchers to present different models for trying to solve this problem.

In [76], a biologically-based disparity energy model that produces 3D-disparity maps to identify and verify different facial expressions was proposed. The disparity energy model estimates local disparities by using two neuronal populations namely the encoding population and higher-level decoding population. Experiments were conducted using the BU-3DFE dataset and showed satisfactory recognition performance.

In [77], a pose-invariant FRS was proposed by enhancing the modified decision based unsymmetrically trimmed median filter to eradicate noise from the face image. Also, proposed was a dominant gradient local ternary pattern descriptor for feature extraction with the support vector machine classifier for classification. Experiments were carried out

on the Japanese facial expression (JAFFE) dataset that showed the effectiveness of their approach.

In [78], a 3D FRS that uses specific curves for multiple subjects which are unaffected by intra-subject differences was proposed. The bone structure with sharp differences in the face convex regions to extract curves with convex crests was closely linked, and the exact location of the symmetry plane and noise tip to detect the curves of the nose was utilized.

Because of their unique ability to change expression, the different curves are joined with appropriate weights at the feature level and used for matching 3D faces with the iterative closest point algorithm. An experiment carried out on the GavabDB, and BU-3DFE 3D public datasets performed satisfactorily performance. To address the issue of not enough data in the public dataset for facial expression recognition, in [79], applied specific pre-processing approaches that could extract only expression-specific features from the face image. Then, a convolutional neural network architecture was used as the classifier on the JAFFE dataset for recognition. Table 2.3 presents a summary of the FRS for addressing the issues of expression.

2.3.4 Plastic surgery

It is observed that face recognition techniques fail to identify an individual's face after plastic surgery. This is a situation where the face image of an individual is entirely changed, into that of an extremely different individual [80-82]. The plastic surgery process on the human face introduces change in skin texture between images of an individual, examples are shown in Figure 2.5, thus making the FR task hard [83, 84]. A typical scenario of plastic surgery and how it can cause a change to the face image is rhytidectomy. This is a procedure where the face is completely changed. This can enhance the change of aged skin texture, thus making the skin texture appear younger. Eye lift, reshaping of the nose, enhancement of the jaw are major changes done on the face to cause a change in facial appearance [85-88].

Face recognition techniques have been developed such that they are robust to the issues of plastic surgery. The dataset used to evaluate FRS in this issue include the plastic surgery face dataset. In [89], a method known as entropy based volume SIFT for recognizing faces after plastic surgery was proposed. Their approach extracts important regions and volumes of the scale-space that represent the features for facial information. The corresponding features are applied to the SVM for classification. Experiments were carried out for different plastic surgery conditions on the plastic surgery dataset. It was reported that on blepharoplasty, a recognition rate of 88% was achieved, while on brow lift and liposhaving, satisfactory recognition rates of 87% and 86%, respectively.

Figure 2.5. Facial changes due to plastic surgery (Taken from [85], with permission).

In [85], an edge-based Gabor feature representation technique to recognize surgically changed faces was utilized. The information gathered from the edge is used to address the textual differences caused by the plastic surgery. Also, in [88], the duo of GIST global descriptor and local binary patterns that were feature-based and texture based respectively were combined. First, the local binary pattern was applied on important regions instead of the whole face image. The concept assumes that only local binary patterns such as corners and edges are used to recognize faces after plastic surgery. Their experiments on the plastic surgery dataset reported 81% recognition accuracy. Table 2.4 presents a summary of the FRS for addressing the issues of plastic surgery.

Table 2.3. A summary of recent face recognition systems for addressing the issue of expression.

Technique	Research work	Face dataset	Comments
3-D disparity maps based on biological disparity model	Martins *et al.*, [76]	BU-3DFE	The disparity energy model uses two neuronal populations namely encoding population and higher-level decoding population to estimate local disparities. Neurons adjust to a broad variety of constraints like orientations, frequencies and horizontal differences.
Local tenary pattern descriptor with support vector machine	Revina *et al.*, [77]	JAFFE	A decision-based asymmetrically trimmed median filter is used to eradicate noise from the face image; and a dominant gradient local ternary pattern descriptor is used for feature extraction with the support vector machine classifier for classification.
3-D technique using multiple subject-specific curves	Li *et al.*, [78]	GavabDB	The different curves are joined with suitable weights at the feature level and used to match 3-D faces.

Table 2.4. A summary of recent face recognition systems for addressing the issue of plastic surgery.

Technique	Research work	Face dataset	Comments
Shift-invariant Fourier transform based on entropy with support vector machine.	Sable *et al.*, [89]	Plastic surgery	Extracts important regions and volumes of the scale-space that represent the features for facial information.
Gabor feature representation based on edges	Chude *et al.*, [85]	LFW	The information generated from the edge is used to solve the difference in texture caused by the plastic surgery.
GIST global descriptor with local binary patterns	Ali *et al.*, [83]	Plastic surgery	LBP features with information such as corners and edges are useful; while the GIST descriptor is used to obtaining a basic and subordinate description of related dimension.

2.3.5 Aging

Aging is seen as a type of within-class appearance variation in the faces of humans. It occurs where much time difference exists within the target face and the query image of the same individual [90, 91]. Aging variation can have a significant effect on the overall facial structure of individuals as shown in Figure 2.6. Unique facial features can also change due to changes in the age. Hence, face recognition techniques tend to fail in such situations.

In literature, it is observed that low performance occurs mostly when there is a significant age gap between the target image and the query image [92]. Datasets used to evaluate face recognition techniques where the issue of aging is involved include FG-net and MORPH. These datasets contain some images and other uncontrolled issues such as illumination and pose.

Figure 2.6. Facial changes due to age difference (Taken from [93], with permission).

In [94], a model that uses appearance-age labels obtained from age-subspace learning was proposed. The aging subspace using the labels to determine the subspace of the identity using the expectation-maximization algorithm was trained. It was observed that the trained aging subspace was unaffected by the face images used. Experiments were carried out on the FGNET, and the MORPH dataset achieved satisfactory recognition rates.

In [95], motivated by the fact that age difference is a nonlinear but smooth transform, and by the strength of autoencoder networks to learn hidden patterns from inputs, a neural network labelled coupled auto-encoder networks for FR was presented. Experiments carried out on the FGNET dataset showed a recognition rate of 86.5%. Table 2.5 presents a summary of the FRS for addressing the issues of aging.

Table 2.5. A summary of recent face recognition systems for addressing the issue of aging

Technique	Related work	Face dataset	Comment
Identity inference model based on age-subspace learning	Zhou *et al.*,.[94]	FGNET	Appearance age label to model the human aging and identity variables was utilized. The aging subspace using the labels to determine the subspace of the identity using the expectation-maximization algorithm were trained.
Couple a-encoder networks based on neural networks	Xu *et al.*, [95]	FGNET	Non-linear factor analysis is presented to decompose a face image into an identity feature that is age-invariant.

2.3.6 Occlusion

Occlusion can be defined as the hindrance or blockage of a section of an image or object. Hence, occlusion of the face image can be described as intentional or unintentional covering or obstruction of part of the face, which results in a reduction in the performance of the FRS [96-98]. Crimes have always been associated with activities such as terrorist attack, theft at ATMs and burglary. These activities are mostly conducted by individuals who have their faces occluded such that even installed cameras do not help in identifying them. The blockage of the face image with scarves, sunglasses or hats can impede the performance of the FRS as quite substantial information of the face image will be lost. Also, occlusion results in changes in the appearance of the face image [99]. Hence, it is important that FRS have to be designed in a way that they are robust to the issue of occlusion.

Figure 2.7. Samples of occluded face images (Taken from [100], with permission).

In [101], a pixel-level occlusion detection based on sparse representation was proposed. The approach was introduced to represent the query image and to obtain the residuals of each class, where each pixel was estimated as either occluded or not in each class's residual. A dilation to remove the isolated occlusion estimation pixel was performed, since occlusion was contiguous. They reported a recognition rate of 87.6% on the AR dataset. In [102], an efficient locally-constrained occlusion coding technique that improves the sparse error correction with Markov random field algorithm was proposed. Their approach removed occluded sections by locally constrained coding and reduced running time for recognition. Experiments conducted using the Yale B face dataset reported a satisfactory recognition rate.

In [103], a double occlusion scenario where occluded face images appear both in the training and test data was presented. A fuzzy max-pooling approach based on the

convolutional neural network was proposed. Experiments carried out on the AR dataset reported a satisfactory recognition rate. Table 2.6 presents a summary of the FRS for addressing the issues of occlusion.

2.3.7 Low resolution

The issue of low resolution in FRS occur when the test face image has been degraded drastically. This results in loss of important information from the face image across different individuals. This issue affects FRS especially in applications such as surveillance [104]. Hence, it remains an issue in FRS as compared with images in high resolution.

Table 2.6. A summary of recent face recognition systems for addressing the issue of occlusion.

Techniques	Related work	Face dataset	Comments
Sparse representation classification	Zhao *et al.*, [101]	AR	Sparse representation classification is applied to represent the query image and obtain the residuals of each class, where each pixel is estimated as either occluded or not in each class's residual.
Sparse error correction with Markov random field algorithm	Fu *et al.*, [102]	Yale B	Removes occluded sections by locally constrained coding and reduces running time for recognition.
Convolutional neural network fuzzy max-pooling method	Long *et al.*, [103]	AR	A fuzzy max-pooling approach based on the CNN was proposed. An average pooling was then used to improve the effectiveness of their approach.

In [105], a low-resolution FR with a sample per individual was investigated, and proposed a cluster-based regularized simultaneous discriminant analysis technique . Their approach regularized the between-class and within-class matrices with intercluster and intracluster matrices. The cluster-based scatter matrices were estimated from unsupervised clustering. Experiments were conducted using the SCface dataset, and they reported a recognition rate of 82.36%.

In [106], an effective coupled distance metric learning algorithm label coupled marginal discriminant mappings was proposed. Their approach makes the data points in the original low, and high-resolution features extend into a unified space, where classification is carried out. Experiments were done on the AR face dataset, and satisfactory recognition rate was reported. Table 2.7 presents a summary of the FRS for addressing the issues of low resolution.

Table 2.7. A summary of recent face recognition systems for addressing the issue of low resolution.

Technique	Related work	Face dataset	Comments
Cluster-based discriminant analysis technique.	Chu *et al.*, [105]	SCface	Their approach regularizes the between class and within class matrices with intercluster and intracluster matrices. The cluster-based scatter matrices are estimated from unsupervised clustering.
Coupled marginal discriminant mappings.	Zhang *et al.*, [106]	AR	Data points in the original low and high-resolution features extend into a unified space, where classification is carried out.

2.4 REVIEW OF FACE RECOGNITION TECHNIQUES

Many approaches at the different stages of FRS has been implemented because the subject area has been a fast-growing and tasking research topic in image processing and computer vision [43]. These techniques have shown satisfactory performance, especially in well-controlled conditions, while there is a drop in performance in real-life conditions. In this section, review on different approaches in the stages of FRS as shown in Figure 2.1 are discussed. The pros and cons of these approaches are highlighted where necessary.

2.4.1 Pre-processing techniques

Pre-processing in FRS involves that certain approaches be made on facial images to improve its quality thereby enhancing the performance of the FRS [24]. Furthermore, the pre-processing stage amends distorted facial images and acquires regions of interest from the image for onward feature extraction [185]. The following sub-sections discusses major and important pre-processing approaches been used in FRS.

2.4.1.1 Face Alignment

Alignment of faces or localizing fiducial facial point as a pre-processing approach has received great attention from researchers because it is beneficial for important tasks such tracking of the face, facial expression recognition and head pose estimation [26]. Face alignment refers to the predefined landmarks on a face image, that are usually pointed at the facial components such as nose, eyes and mouth. It is seen as a task of looking over a facial image for the pre-defined facial points that starts from an initial shape and goes on to refining the shape estimate until convergence [27]. The facial appearance and the shape information are two sources of information used during the process of looking over the facial image. The information derived from the shape tends to relate the locations of facial points to make sure that the calculated points can construct a face shape [28].

Face alignment is of benefits to facial attribute computing because facial attributes like the shape of the nose and eye glasses are closely related to specific longitudinal locations of a

face [26]. Also face alignment has been of benefit for expression recognition where the configurations of facial points are reliable indicative of facial changes caused by expressions. The analysis from this will show the specific type of expression that might have led to such change. This idea has been used by different researchers for expression recognition [186].

Most importantly, in face recognition, face alignment has widely been used to improve different algorithms in unconstrained environments. For instance, in registering the facial image, the initial phase is locating major facial points and making use of them as anchor points for specific warping [187]. Other face recognition algorithms rely on an effective alignment of the face to construct the correspondence among local features such as the mouth, nose and eyes [188]. Face alignment is effective on FRS in controlled environments, however it remains a challenge in unconstrained environments [27].

The different face alignment algorithms for facial images in unconstrained environments can be grouped into two categories such as the generative and discriminative methods [28]. These methods are based on the modelling principles in pattern recognition [28]. The generative method models for both the facial appearance and face shape. Using these methods, face alignment is seen as an optimization problem to find parameters of the shape and appearance that can come up with an effective appearance model to test the face [187]. The appearance of the face can be denoted using the entire facial image or, local image patches centred at the facial points. Examples of the generative methods includes the regression-based fitting, gradient decent-based fitting and the part-based generative deformable model [186].

The discriminative method generates the target location from the appearance of the face. This is carried out by learning independent local detector for every individual point on the face and using a whole shape model to standardize the predictions [186]. It could also be carried out by learning a vectoral regression function to generate the global shape of the face, where the shape constraint is indirectly determined. Examples of the discriminative

methods includes the PCA shape model, ensemble regression voting and deep neural networks [187].

2.4.1.2 Image enhancement

Image enhancement is a major pre-processing approach that helps increase the performance of various face recognition algorithms, however most face recognition utilizes a less or rather no efficient image enhancement approach before undergoing the recognition process [185]. Image enhancement is considered as an important field of digital image processing used to improve on the quality of face images thereby increasing the number of unique features present in the image [181]. Also, image enhancement is the process of applying certain transformations to an input image to achieve a more detailed or less noisy output image [175]. For an image enhancement technique to be effective it must be self-adaptive, i.e., adjustment of parameters for enhancement technique should be made automatically for different image types. It must be able to adjust the entire image accordingly without the issue of over-stretching [185]. Also, the enhancement technique must be able to undergo the process without resulting in an unnecessary change in brightness or loosing of important features and should be made simple with low computation complexity [189].

The quality of most facial images in unconstrained environments are degraded for reasons such as lighting conditions, i.e., dark or too bright environment. Also, various facial expressions tend to change the appearance of face images by hiding some important features of the face image [27]. In the literature, there have been some image enhancement techniques developed for various image processing tasks. However, minimal work has been carried out in aspect relating to face recognition where various real-life scenarios are considered [190]. Also, most image enhancement techniques result in less enhanced outputs or unnatural effects and over enhancement, which all have a negative effect on the image. For these reasons, there is the need for enhancement of facial images as it is a major challenge and an important piece at the pre-processing stage of the FRS [68].

Histogram equalization (HE) is an image enhancement approach that is considered as a simple and an efficient contrast enhancement technique for enhancing images [191]. The

HE is a technique in image processing of contrast adjustment by increasing the global contrast of the image, especially when the usable data of the image is depicted by close contrast values [190]. It operates by spreading the intensities of image pixels based on the information from the whole image. This results in situations where low occurring intensities are transformed to become fused with neighbouring high occurring intensities, thus leading to over enhancement [189]. Also, mean shift issues tend to arise in such situations, hence, maintaining the brightness is not guaranteed when using histogram equalization. In attempting to improve on this problem, the bi-histogram equalization was introduced, which displayed better performance while also maintaining the quality of the original image [175]. However, it does not always produce the preferred output image when the image pixel distribution does not follow the symmetric distribution. Also, enhancement methods such as gamma correction and logarithm transforms have been used and observed to be of low computational complexity. However, they are not able to manage with complex illumination differences [185]. Furthermore, an extension of other histogram equalization approaches such as block-based histogram equalization, oriented local histogram equalization and adaptive histogram equalization (AHE) has been proposed in the literature but underperform on face recognition tasks in a complex illumination situation. This poor performance occurs because these methods rescind the entire distribution, which may contain important image characteristics [189].

Linear contrast stretching (LCS) is an image enhancement technique that uses linear transformation to increase the dynamic range of grey levels present in the original image [175]. LCS improves the contrast grade of the image, though the threshold must be set manually. When the threshold is not suitable, the supposed enhanced image might be of lower quality as compared to the original image. Since no universal standard exist for image quality assessment, it becomes difficult to improve on an image by simply stretching its histogram or utilizing simple gray level transformations [181].

Thus, to solve these general issues, particularly when computers need to decide autonomously how good an enhanced image is, researchers have recently proposed methods based on evolutionary computation and metaheuristic optimization algorithms

[168]. Metaheuristic algorithms are generally used to finding solutions involving non-linear optimization tasks. The process involved in using metaheuristic algorithms requires that an evaluation function selects automatically the optimal enhancement parameters of a transformation function to appropriately enhance a facial image [168].

2.4.2 Feature extraction techniques

The feature extraction stage if the FRS involves the simplification of the amount of resources that describes a large set of data. Feature extraction is used to minimize the original face dataset by getting some properties that can be used to classify and get patters that are present in the input facial images. The different methods of features that can be extracted from the facial images have been presented in the following sub-sections.

2.4.2.1 Principal component analysis

Principal component analysis (PCA) is a known face recognition technique also called the Eigenface or Karhunen-Loeve expansion [107, 108]. Sirovich and Kirby represented face images adequately by using PCA after stating that face images could be modified through a minimum aggregation of weights for every face and a standard face image. Subsequently, Turk and Pentland in 1991 introduced the Eigenfaces method for FR [109].

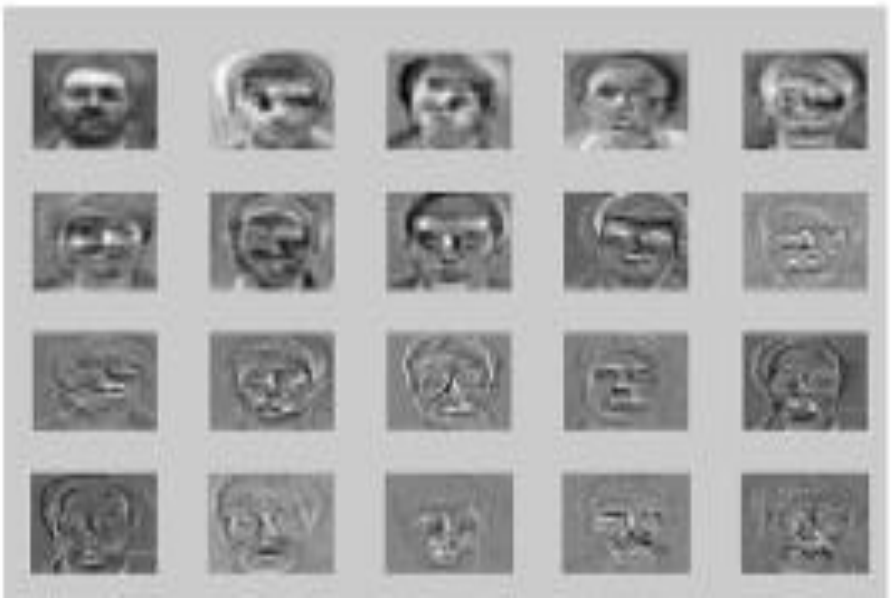


Figure 2.8. Samples of constructed Eigenface(Taken from [111], with permission).

The PCA follows a given principle: given a set of training images each of a uniform pixel size, the standard features are obtained from the facial image so that the unique features remain. This is done by subtracting the average face vector from the face vector [108]. After that, the eigenvectors will be calculated from a covariance with reduced dimensionality which is arranged to correspond to the larger Eigenvalues [110]. Figure 2.8 shows an example of Eigenfaces constructed.

2.4.2.2 Linear discriminant analysis

Linear discriminant analysis (LDA) is a commonly used approach for FR. Unlike the PCA, the LDA constructs a subspace to differentiate the faces of different people [43]. LDA enables evaluation of the important information in various parts of the face to recognize the human face. The LDA classifies the face image into multiple groups using features that describe the face. Its process entails - i) collecting images and classifying them into classes, ii) calculating the vector of the classes where the goal is to reduce the within-class matrix and increase the between-class matrix, iii) calculating Eigenvector and iv) obtaining the Fisher faces [112]. Figure 2.9 shows some Fisher's faces. The majority of the linear discriminant analysis FRS approaches suffer from the fact that there is no relationship between their optimality criteria and the feature representation obtained [113].

Figure 2.9. Samples of Fisherface (Taken from [111], with permission).

2.4.2.3 Independent component analysis

Like the PCA, independent component analysis (ICA) is also a known and used subspace method that projects data from a high-dimensional space to its lower form. ICA is a feature extraction method that has been considered as a generalization of the PCA that is mainly used to solve problems related to signal to process [43]. ICA is seen as a method that can be employed in the FR task where information amongst pixels are in the high order relationship [114]. PCA considers images as variables that are random with a Gaussian distribution and minimized second-order statistics. Hence, for a non-Gaussian distribution, large variances will not be matched to PCA basis vectors. The ICA method finds statistically dependent characteristics in the input data and reduces both second and higher-order dependencies [115, 116].

2.4.2.4 Local binary patterns

The local binary pattern (LBP) operator converts an image into an array of integer labels. The LBP operator was initially designed for the description of texture where every pixel is assigned a label by an operator. This is done by using binary images from a greyscale image processed using the three-by-three neighbourhood of each pixel [117, 118].

The LBP registers a point as the centre point and then calculates the difference between that point and the points around it. If the difference happens to be greater than zero, a one is assigned, otherwise it remains zero. Figure 2.10 shows the illustration of a generalized LBP operator. Furthermore, it can generate a texture descriptor by using the histogram of labels.

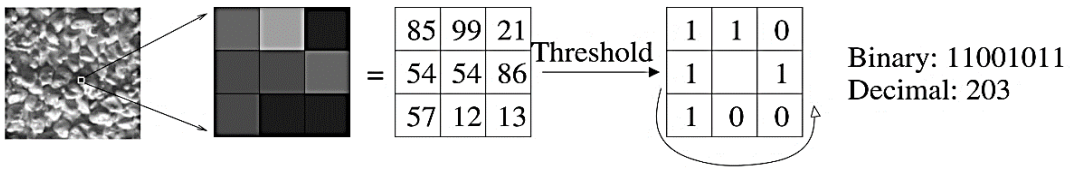


Figure 2.10. Local binary pattern operator(Taken from [119], with permission).

2.4.2.5 Dynamic link architecture

The dynamic link architecture (DLA) is a neural information processing concept that was first proposed in 1981 as an attempt to solve issues encountered by the artificial neural network (ANN) such as the expression of syntactical relationships [120]. The basic idea of the DLA is the use of synaptic plasticity already present on the time scale of information processing and not only for the acquisition of memory that allows it to immediately group neurons into higher units [121]. The ability of the DLA is best used in issues of FR such as facial expression and invariant object recognition. A DLA based on multiscale morphological dilation-erosion was presented for FR to yield a feature vector to verify faces of individuals from a given test set. Their experiment showed that the method presented performed better than the dynamic link-matching based on Gabor wavelets [122].

2.4.2.6 Elastic bunch graph matching

The elastic bunch graph matching (EBGM) technique compares images where the algorithms first identify landmark locations on images that have look-alike features of the face. These features are described by using the Gabor wavelet convolution which is known as a Gabor jet when all the values are at a single point. Furthermore, the face graph is used

to depict each image where the nodes are placed at the landmark locations with each containing a Gabor jet extracted from that location [123].

There is difficulty in identifying faces from a large dataset of images due to image variation in terms of size, position, pose and expression. The EBGM is regarded as one of the most successful FR techniques as it has been applied to some FR tasks. However, the modern FRS is required to be automated-based without any intervention on the part of humans. Hence, the drawback of the EBGM lies in the fact that landmark selection of the face image is made manually at the initial stage of the recognition process [124].

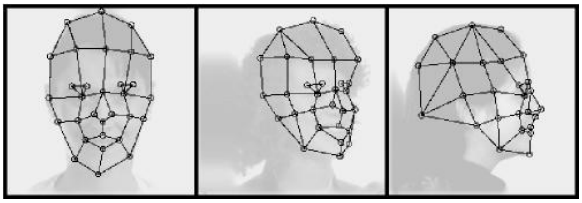


Figure 2.11. The EBGM technique (Taken from [125], with permission).

2.4.2.7 Geometric feature matching

Geometric feature matching (GFM) methods were one of the earliest ways to recognize individuals by using their faces. The face image can be recognized even though the details of its major features are not resolved. The remaining information is purely geometrical and represents what is left at a very coarse resolution (i.e., loss of image pixels). However, the idea is to get the informative features such as chin and other important facial parts. This method uses training face images to locate the eye position in the image and the correlation coefficients are calculated. It then compares this with the test image and searches for the maximum values [126, 127].

The geometric feature-matching method starts by calculating a set of geometrical features in the facial image. What makes FR acquirable at a coarse resolution below 8 x 6 pixels even when the features of the face are not easily shown, indicates that the geometric pattern of the facial features is enough for FR processes. A vector depicting the main features of the face through its size and its physical outline can interpret this geometric

pattern. Geometric feature matching is seen as a technique that performs well when searching for a common match in large datasets. However, it relies mostly on the correct feature location algorithms [128].

2.4.3 Feature classification techniques

The feature classification stage of the FRS involves the process where the acquired facial feature data are categorized into different and given number of classes depending on a particular task. The different methods of feature classification techniques are presented in the following sub-sections.

2.4.3.1 Artificial neural network

Artificial neural network (ANN) is a system of interconnected artificial neurons that can share messages amongst one another and also learn from experience. They are inspired by the biological neural system where the computational unit of the brain is a set of interconnected neurons with synapses [129]. The ANN is seen as an efficient and robust classification algorithm that is used to predict both known and unknown data.

The interconnected neurons consist of numeric weights that are adjusted during the training stage in such a way that a well-trained network will efficiently respond when given a recognition task [130, 131]. It also includes multiple layers of feature detecting neurons, where every layer has various neurons that respond to combinations of inputs from the previous layer. These layers are built in a way that the foremost layer detects a set of given inputs, and subsequent layers detect “patterns of patterns” of the previous layers.

2.4.3.2 Convolutional neural network

The convolutional neural network (CNN) is seen to be inspired by the biological evidence found in the visual cortex of the mammalian brain. An experiment by Hubel and Wisel displayed that neuronal cells present in the brain were triggered when they noticed edges of particular orientation. For instance, certain neurons trigger the vertical edges while others

trigger to the diagonal or horizontal edges, where the neurons are placed together to execute visual perception. This concept is the basis of the CNN [132-134].

The CNN was first introduced in the 1980s and the 90s but was neglected due to its complexity in real-world applications. However, interestingly, it has lately been revived and has since outperformed most computer vision techniques while still growing at a fast pace [135]. The CNN architecture consists of various layers such as the convolutional layer, pooling layer, non-linear layers, and the fully connected layers. Each layer has its impact responsible for the performance of the CNN network in various recognition tasks [136]. The convolution layer and the pooling layer act as a neural network that extracts features while the fully connected layers act as the classification neural network that functions based on the image features and comes up with an output [137]. Figure 2.12 shows a generic architecture of the CNN approach to FR.

The CNN is an approach that attempts to solve the shortcomings of the NN. These shortcomings include the fact that when the input dimension is high, as regards the number of connections and images, this will make the number of free parameters high, as there will be a connection between each hidden unit and the input layer. Hence, the training samples might be smaller than the pattern dimension, i.e. NN would have high complexity and could potentially lead to overfitting of data. Thus, the CNN automatically learns local feature extractors and implement the weight-sharing principle enabling a reduction in the number of free parameters, thereby increasing the performance capacity as compared to the NN architectures [137]. The CNN is regarded as a class of feed-forward ANN that has shown great achievements in the field of image recognition [135].

One of the early CNN LeNet5 introduced the field of deep learning done by Yann LeCunn in the 1990s. The LeNet architecture was mainly used for recognition tasks such as digits and zip codes (90). Moreover, there are several other CNN architectures such as AlexNet (2012), ZF Net (2013), GoogLeNet (2014), VGGNet (2014) and Residual networks (ResNet) (2015).

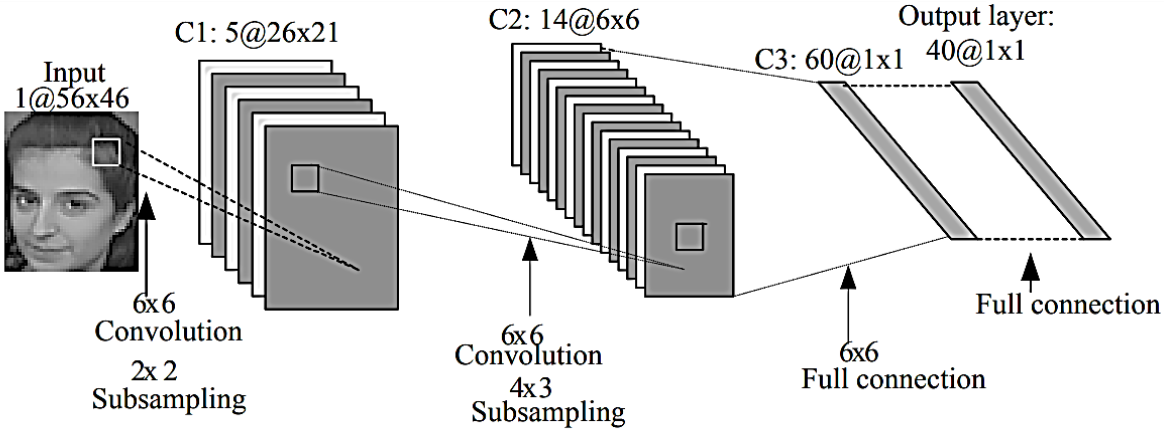


Figure 2.12. A CNN architecture for face recognition (Taken from [137], with permission).

AlexNet was the first modern approach of the CNN designed by Alex Krizhevski [138]. This approach was identical to the LeNet CNN architecture but was deeper and had convolution layers arranged on top of one another. It was submitted to the ImageNet ILSVRC challenge and performed well [139]. ZF Net was developed by Mathew Zeiler and Rob Fergus and won the ILSVRC challenge in 2013. It was an extended work of the AlexNet, where they adjusted the parameters of the CNN design by increasing the size of the centred convolution layers and reducing the stride and size of the filter [137]. GoogLeNet also became the winners of the ILSVRC challenge in 2014 and was developed by Szegedy *et al.*, [140], an employee of Google. Its CNN was designed in such a way that the numbers of network parameters were reduced. Also, the approach made use of average pooling rather than fully connected layers above the ConvNet thereby eliminating a large number of parameters that were not important [135]. VGGNet developed by Karen Simonyan, and Andrew Zisserman [141] was the runner-up in the ILSVRC challenge in 2014. The CNN architecture showed that the depth of the network is an important component for good performance. Their model can be used in Caffe. However, it is expensive to evaluate and make use of large memory and parameters (140M). ResNet designed by Kaiming *et al.*, [142] won the ILSVRC challenge in 2015. Their CNN approach is currently seen as one of the most used CNN models in recent times [135].

A major advantage of the CNN is the fact that it presumes that inputs are images that enable coding of certain properties into the architecture, thus making the forward function easier to implement while also reducing the number of parameters in the network [143]. Moreover, the CNN has embedded, in its architecture, the feature extractor in its training process, unlike traditional techniques where the feature extraction stage is manually designed, thus leading to an increase in execution time and computational cost. The CNN is a supervised multilayer ANN that has the capability to search for the most effective way of extracting features during the training stage, i.e., it gets important features from face images that can be used for classification purposes [144]. For tasks that involve pattern and image recognition applications, the CNN has shown that it can achieve the highest detection rates, as compared to other detection algorithms when it comes to FR in unconstrained environments (FRUE) [145]. CNN has been used in quite some image-processing and computer vision tasks where it has shown excellent performance.

2.4.3.3 Random forest

The random forest (RF) machine learning technique is one that can be used for both classification and regression tasks. It is a supervised classification technique where the approach is created like a forest with some trees as the name implies. The machine learning technique was proposed by Leo Breiman [146, 147], and it is an amalgamation of tree predictors in a way that an individual tree is dependent on the values of a random vector tested separately using similar distribution. The ability to have random inputs and random features makes RF technique robust, thus capable of dealing with large feature spaces.

RF has been used for various tasks like object detection, handwritten signature and FR due to its low computation complexity, accuracy, its ability to be easily implemented and capability to handle training datasets that are large [148]. RF is a collective classifier that includes some decision trees. An individual tree is made recursively by allocating a binary test to the individual non-leaf node according to the training samples. For classification purposes, the results of the decision tree are combined to select the highest nominated class [149]. The RF classifier could be defined as shown in (2.1):

$$H_{(P)} = \arg\max_q \sum_{i=1}^{k} I(h_i(p)) = Q \tag{2.1}$$

where $H_{(P)}$ is the is the eventual joined classifier, k is the number of decision trees, $h_i(p)$ depicts a decision tree, Q represents the class label, and $I(h_i(p) = Q$ depicts that p belongs to a class Q.

2.4.3.4 Support vector machine

The support vector machine (SVM) has been widely studied over the years and applied to classification and regression tasks where it has shown satisfactory results in detecting faces [150]. The SVM is an effective approach for most pattern recognition tasks due to its ability to perform well without the need for added knowledge [151]. For two classes having a set of points, the SVM locates the hyperplane to maximize the distance from any classes. It is a method used to train polynomial neural networks functions classifiers [152, 153].

These learning techniques often use the Structure Risk Minimization (SRM) in which the best generalization capabilities minimize the boundary of the generalization error. The decision surface is a result of the weighted combination of elements called the support vectors. Assuming the input to an SVM algorithm is defined (p_i, p_i of labelled training data, in which pi represents the data, and qi equals -1 or 1 and represents the label. Hence, the output of the algorithm depicts a set of support vectors vi; class labels q_i, coefficient weights αi and a constant k [154]. Hence, the linear decision surface is given in (2.2) and (2.3):

$$w.z + k = 0 \tag{2.2}$$

In which case

$$w = \sum_{i=1}^{N_S} \alpha_i v_i q_i \tag{2.3}$$

The SVM can further be extended to a non-linear decision surface by using a kernel C that must satisfy the condition of Mercers. Hence the non-linear decision surface is defined in (2.4) as:

$$\sum_{i=1}^{N_S} \alpha_i q_i c_i \left(v_i . z\right) + k = 0 \tag{2.4}$$

2.4.3.5 Decision tree

The decision tree (DT) is a technique seen as a predictive modeling tool that carries out the classification task where it does not take into consideration that the attributes are independent. The decision tree presents a model that depicts correlations amongst pairs of attributes in an organized hierarchy[155]. Decisions are made by this approach in such a way that the attributes are calculated within the entire level of the tree rather than on a specific sub-tree. Results are further shown in an organized table instead of a tree-like form [156, 157].

Zhou *et al.*, [158] showed how a DT approach to spatial continuity of occlusion could be linked into the computation of a sparse representation of the probe image. Their experiment indicates that their algorithm also discovers regions that are not useful and removes them from the sparse representation, thus tolerating fractions and different types of facial occlusion.

2.4.3.6 Naïve Bayes

The naïve bayes (NB) is a technique based on the probabilities of conditions that use the Bayes' theorem. The Bayes theorem finds the probability by finding the number of frequency values and combinations of values in the dataset. It identifies the probability of an act happening, assuming the probability of another act is known [159].

Let Y be a training set of multiple elements and corresponding class labels, where each element depicts an n-dimensional attribute vector, $V = (V_1.V_2.V_3.....V_n)$, representing n measurement made on the multiple elements from n attributes, respectively,

$B_1, B_2, B_3.....B_n$ Assuming there are n classes, $I_1, I_2, I_3.....I_n$. Given a tuple, Z, the classifier will envisage that Z is appropriate to be in a class that has the highest posterior probability, as regard Z. Thus, the NB envisages that tuple Z is of the class I_1 only if $P(I_1 \mid Z) > P(I_j \mid Z)$ for $1 < j < n$ is not equal to i. The class I_1 for which $P(I_1 \mid Z)$ is maximized is known as the maximum posterior hypothesis [160, 161]. Hence by Bayes theorem as given in (2.5):

$$P(I_1 \mid Z) = \frac{P(I_1 \mid Z)P(I_i)}{P(Z)} \tag{2.5}$$

Yan *et al.*, [162] proposed an optimized Naïve Bayes algorithm with the application on FR. First, the algorithm, at each grey level, estimates the probability of each pixel. Then Laplace smoothing is performed to get the zero-probability issue. Lastly, maximum filtering is used to optimize the probability distribution matrix for classification. Their approach performed satisfactorily performance when evaluated on the Labelled Face in the Wild face dataset.

2.5 FACE RECOGNITION BENCHMARK DATASETS

In this section, the major FR face datasets that have been developed are discussed. These face datasets include one or more constraints that depict real-life scenarios which researchers use to evaluate their various face recognition techniques or approaches. A summary showing the constraints in the various face datasets is presented in Table 2.8.

2.5.1 AR face dataset

The Aleix Martinez (AR) face dataset is a collection of over 4000 coloured face images of 126 individuals (56 women and 70 men) obtained by the Computer Vision Centre in Barcelona, Spain in 1998. During the collection of these images, the recording and the imaging conditions (camera distance, camera parameters, and illumination settings) were carefully and strictly selected to ensure that the settings were the same for different

subjects [100]. These images were recorded twice, at a two-week interval with a pixel size of 768 x 576.

2.5.2 Yale face dataset

The Yale face dataset was constructed by the Yale Centre for Computation Vision and Control and includes a collection of 165 images involving 15 individuals each providing 11 images. These images are made up of a variety of conditions, such as demonstration of variations in lighting conditions (right-light, centre-light and left-right), facial expression (normal, sad, happy, surprised, sleepy and wink) and with or without glasses [163]. Minimal changes in the facial expression and head position are visible since all the 64 facial images for a particular pose were captured in a time frame of about two seconds. The Yale face dataset has been separated into four different sections with respect to the angle between the camera axis and the light source (120, 250, 500, and 770). The hand-labelled locations of both the eyes and the centre of the mouth are distributed along with the dataset.

2.5.3 Yale face dataset B

Because of the need for proper evaluation of FRS under great pose and lighting variations, Yale face dataset B was collected. This dataset contains subjects imaged inside a geodesic dome using 64 computer-controlled xenon strobes. The dataset contains 5760 faces gotten from 10 different subjects under 64 illumination situations and 9 different poses.

2.5.4 Labelled faces in the wild

The Labelled Faces in the Wild (LFW) dataset is a collection of 13233 of faces collected from the web and labelled with each person's name. The dataset consists of 5749 different individuals with 1680 of the individuals having at least two or more distinct images in the dataset [164]. The remaining 4069 have only a single image in the dataset. Most of the images in the LFW dataset are coloured and available in 250 x 250 pixels JPEG with only

a few greyscale images. The LFW dataset is divided into two; one for algorithm development and the other for performance reporting.

2.5.5 Facial recognition technology

The Facial Recognition Technology (FERET) dataset is a collection of 24 facial image categories, obtained from the George Mason University and the US Army Research Laboratory facility as part of the FERET program which was sponsored by the US Department of Defence Counterdrug Technology and Development Program [163]. FERET has been used extensively in past works, for example, FERET and facial recognition vest test (FRVT) 2000 datasets have been exclusively evaluated and are available for use by both commercial facial recognition systems and research algorithms. In the FERET dataset, the list of images used for training, gallery, and probing is distributed in line with the dataset, thereby ensuring that adequate comparison of proposed recognition algorithms can be extensively carried out. Records obtained show that the FERET dataset has been distributed to over 460 research groups.

The facial images in the FERET dataset were recorded between August 1993 and July 1996 over 15 sessions. A 35mm camera was used for the image recording before the images were digitized and transformed to 8-bit grey-scale images of 256 X 384 pixel in size. The FERET dataset contains some variations between recording sessions as a result of the recording equipment reassembled per session. The ground-truth information, together with the date the recording was carried out and information that indicates if the subject is wearing glasses, are supplied for each image in the dataset. Additionally, the locations of the right and left eye and mouth centre that have been manually determined are available for 3816 images. Adjustment has been made to the new version of the FERET dataset, as NIST has made higher-resolution colour images (512 X 768) of the initial grey-scale images.

2.5.6 Olivetti research laboratory

The Olivetti research laboratory (ORL) dataset is a collection of facial images containing slight differences in face expression (open/closed eye, smiling/not smiling), illumination and face details (glasses/non-glasses) collected from 1992 to 1994 at the ORL in Cambridge. The ORL facial dataset contains images of 40 different individuals with 10 images per subject. Each facial image has been normalized to present a 23 x 28-pixel array with a grey level range between 0 and 225 [163].

2.5.7 Plastic surgery face dataset

The plastic surgery face dataset is a collection of faces that depicts a real-life scenario of the surgery issue to FRS. The facial dataset contains 1800 before and after surgery face images that belong to 900 different individuals. 519 of these individuals represent the situation of local surgeries, i.e. changes in some part of the facial image and 381 individuals represent cases of global surgery, i.e. total change of the face image [165]. To assist researchers to come up with robust face recognition techniques for various scenarios of plastic surgery, the dataset contains different types of plastic surgery, i.e. rhytidectomy (face lift), rhinoplasty (surgery of the nose) and blepharoplasty (surgery of the eyelid).

2.5.8 MS-Celeb-1M

The face dataset is a collection of one million celebrities selected from freebase. The sizes of face images are manually labelled using well-designed procedures. The dataset has its facial images aligned and cropped [166]. Table 2.8 further presents the different face datasets and its conditions in uncontrolled environment.

Table 2.8. Face dataset and conditions in uncontrolled environment.

Face dataset	Scenarios in Uncontrolled Environment						
	Illumination	**Pose**	**Expression**	**Plastic surgery**	**Aging**	**Occlusion**	**Low resolution**
AR Face	✓		✓			✓	✓
Yale B	✓	✓					
Life in the Wild	✓	✓	✓			✓	✓
FERET	✓		✓				
ORL		✓	✓				
FG-net	✓	✓			✓		
Plastic Surgery	✓			✓			
MS-Celeb-1M	✓			✓		✓	

It is seen from Table 2.8 that most of the face datasets do not simultaneously contain the different facial conditions in a real-life scenario. Life in the wild face dataset has the most scenarios in unconstrained environment with five different facial conditions such as illumination, pose, expression, occlusion and low resolution. This is followed by the AR face dataset, with four different facial conditions such as illumination, expression, occlusion and low resolution. Also, this is followed by the FG-net and MS-celeb-1face datasets with three facial conditions each. From this, we can deduce that there is the need to come up with face datasets that represent more facial scenarios in unconstrained environments.

2.6 CHAPTER SUMMARY

In this chapter, literature on face recognition techniques that have been developed over time are reviewed and provided. Then, recently published literatures between 2012 to date were also reviewed to appreciate the level of work that has been done in the field of FRS. The survey has provided an in depth understanding of the research area which is going to be of benefits for both new and existing researchers within the field. The literature has also

acted as a motivation to implement an effective pre-processing approach for FRS in unconstrained environments.

In Section 2.2, face biometric recognition systems were presented. This highlights the different stages involved in a typical FRS i.e., the face detection, pre-processing, feature extraction and feature classification stages. It has been found that the pre-processing stage is less talked about in the literature, and as a result affects the performance of existing FRS.

In Section 2.3, a broad review of FRS in unconstrained environments is presented. This includes scenarios such as illumination, pose variation, facial expression, plastic surgery, aging, occlusion and low resolution. Various approaches by researchers to solve these issues are also presented. However, it is found that most face recognition techniques or approaches tend to solve minimal issues faced by FRS in unconstrained environments.

In Section 2.4, the different face recognition techniques are discussed and presented. Finally, in Section 2.5, the different face recognition benchmark datasets that depict unconstrained environments are discussed and presented. It is found that there is a need for easy access to FR datasets and it is also suggested, that a face dataset can be developed such that most of the major constraints are designed within it.

Chapter 3 follows giving details of the materials and methodology used in the thesis. The first section presents the experimental process for the comparative analysis of different face recognition techniques in unconstrained environments. This is followed by the second section that presents the techniques and tools used to model the new image enhancement method as a pre-processing approach for facial images in unconstrained environments. The third section focuses on the approach used to improve FRS in unconstrained environments by using the new image enhancement technique, the new selection of hybrid features and the convolutional neural network.

CHAPTER 3 MATERIALS AND METHODS

In this chapter, the first section presents the experimental process for the comparative analysis of different face recognition techniques in unconstrained environments. Experiments to determine the performance of the different FRS in unconstrained environments are determined. The second section of this chapter presents the techniques and designs used to model the new image enhancement method as a pre-processing approach for FRS in unconstrained environments. The techniques used for the comparative analysis of the different meta-heuristic algorithm-based image enhancement method is also presented in this section. The third section focuses on the approach used to improve FRS in unconstrained environments by using the convolutional neural network. Figure 3.1 further describes the approach used in carrying out the comparative analysis of the different FRS.

3.1 COMPARATIVE ANALYSIS OF FACE RECOGNITION TECHNIQUES IN UNCONSTRAINED ENVIRONMENTS

This section describes the approach and methods used for the comparative analysis of the different FRS in unconstrained environments. The FRS are a combination of feature extraction and classification methods. The different feature and classification techniques used in this work are further described. Figure 3.1 shows the approach followed in carrying out the comparative analysis of the different FRS.

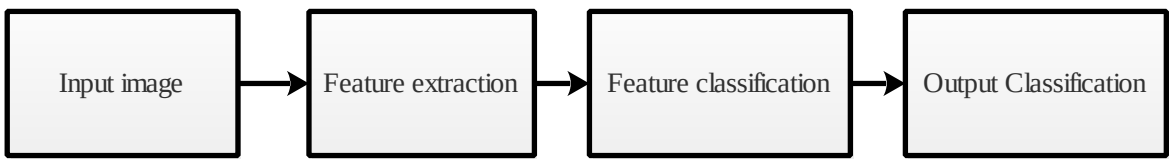


Figure 3.1. Process involved in the comparative analysis of the different FRS.

3.2 FEATURE EXTRACTION TECHNIQUES

The feature extraction stage is a significant step in the FRS, and its major purpose is to simplify the number of resources that describe a large set of data. The feature extraction stage is the stage where the unique features of the face image are extracted. The feature extraction methods are also used to minimize noise and eradicate unnecessary information in the image data. In this thesis, the feature extraction methods selected for the extraction of features include local binary patterns, gabor filter, and edge histogram descriptors that have been described in section 2.4. These features have been seen to be effective, and aid performance for various classification tasks. The following sub-sections 3.2.1 to 3.2.4 further describe how the selected features have been adopted in this thesis.

3.2.1 Local binary patterns

The LBP operator was used to transform each image considered in each respective dataset into an array of integer labels that describe the small-scale appearance of each image. The LBP operator typically described the texture associated with each image where a label was given to every pixel. This was achieved by creating a binary image from the greyscale image considering the three by three neighbourhood of each pixel. The LBP operator was used to label centre points, and then the difference between each centre point and the points in the neighbourhood was calculated. For differences greater than zero, a value of one was assigned, while a zero value was assigned for differences of less than zero.

3.2.2 Gabor filter

The GF was used to extract features for edge detection considering the spatial and frequency domain of each image. A Gaussian function multiplies a harmonic function in a GF. An appropriate selection of the filter parameters was carried out. Following some empirical tests conducted using test images and different parameter values, the following values were used: ksize = 11 x 11 pixels, Sigma = 4, Lambda = 10, Gamma = 0.5, psi = 0 and Theta = 0. Each parameter is defined as follows: ksize refers to the size of the Gabor

kernel, which is preferably odd, and the kernel is a square. Sigma is the standard deviation used in the Gabor filter. Gamma represents the spatial aspect ratio and Lambda is the wavelength. Psi refers to the phase offset, and Theta changes the sensitivity to edge and texture orientations.

3.2.3 Edge histogram descriptors

The EHD was used to identify features describing the edges in each image. The EHD can describe both texture and shape features that are vital components for face image recognition. In this case, each image was converted into its corresponding grey image. The grey image was then divided into 4 x 4 blocks. The local edge histogram was calculated, and the percentage of pixels that corresponded to an edge histogram was found. The same procedure was used to find the pixels that corresponded to the global edge histogram bin. Both the local and global histogram values were then saved in a feature vector that described each image.

3.2.4 Pyramid histogram of gradients

The pyramid histogram of gradients (PHOG) that is an active and effective feature extraction technique was used to extract features from the face images. This was done by first applying the canny edge detector on the face images, which was further divided into spatial grids at all pyramid levels. Secondly, a 3x3-Sobel mask was applied to the edge contours to estimate the orientation gradients. The gradients of each grid were then fused together at each pyramid level.

3.3 FEATURE CLASSIFICATION TECHNIQUES

The classification techniques are used as classifiers at the face recognition stage, where the overall performance of the system is determined. This stage involves both the verification and identification processes, where a face is stored in the dataset, and a probe image is used to compare. Hence, the classifier used in a FRS needs to be effective. The various

selected classification techniques and how they have been adopted in this thesis are further described in the following sub-sections 3.3.1 to 3.3.5.

3.3.1 Random forest

The RF classifier was considered in the comparative study in a bid to determine the most appropriate classifier to be used. The RF classifier was used as follows: A number of bootstrap samples were drawn from the original data. For each of the bootstrap samples, an unpruned classification tree was grown. At each node, rather than choosing the best split among all predictors, a random sample of the predictors was obtained, and the best split was chosen from among those variables. New data was then predicted by aggregating the predictions of the sample trees, that is, the majority votes for classification were used.

The two important parameters of the RF are the forest size and the test size. The forest size specifies the number of random trees to grow. Usually, increasing the number of trees improves the generalization accuracy. In this work, the forests were built in the range of 100 – 1000 trees at intervals of 100 trees. This enabled proper identification of the most effective range of the RF classifier. The test size parameter specified the number of variables used to determine the best split at each node. In this regard, a test size of between 20 to 150 was considered to grow 20 trees, and then the test size from which the minimum out-of-bag rate was obtained was chosen.

3.3.2 Decision tree

The DT was also examined as a potential to classify for face recognition. The decision tree algorithm was derived from Algorithm C4.5, which used an information-theoretical approach, entropy, for building the decision tree. In this thesis, the DT algorithm was used to determine the feature that most discriminated the dataset. This was then used to dichotomize (split) the data into classes categorized by this feature. The next significant feature of each of the subsets was then used to partition the category further and the

process was repeated recursively until each of the subsets contained only one kind of labelled data (which belonged to a specific face image).

3.3.3 Support vector machine

The SVM is a well-known classifier considered for FR. In this work, the SVM was trained with the labelled class information. After learning, the discrimination functions between each pair of classes, that is, the face image to be identified and other face images in the dataset were obtained. These functions were represented by several support vectors together and their combination coefficients. In using the SVM, first, the incoming face image was projected unto the features learned during the training stage. The projection coefficients were normalized, and the mean and standard deviation were obtained from the training data. These were used by the SVM to classify the images, and the algorithm's performance was analyzed.

3.3.4 Naïve Bayes

The NB classifier was also studied. It is applies the Bayes theorem with the assumption of naïve independence. Essentially, the model parameters are approximated by using the relative frequencies from the training set. The features were discretized, and a small-sample correction was incorporated in all probability estimates to ensure that no probability was set to zero. The classification outcome of the NB classifier was then analyzed.

3.3.5 Convolutional neural network

There are several CNN architectures in the literature. However, in this work, the state-of-the-art CNN architecture proposed in [167] was considered. Essentially, it consists of six layers including the input and the output layers. Other layers include the Convolutional, Rectified Linear Unit (RELU), Pooling, and the Fully-Connected layers. An array of numbers representing the input features was considered in the input layer while producing

another set of arrays of numbers as output. The Input layer holds the [1 x Q] raw feature values of the face images used in this work. The number of features, Q, typically changes based on the types of feature used. The Convolutional layer consists of 12 filters, which results in a volume of [1 x Q x 12]. The RELU layer uses a max (0, x) thresholding at zero, thus leaving the size of the volume unchanged as [1 x Q x 12]. Down-sampling operations along spatial dimensions are performed in the Pooling layer resulting in a volume size of [1 x Q/2 x 12]. The Fully-Connected layer consists of a [1 x 1 x N] volume size, where N denotes the number of class scores associated with the different image datasets considered. In this way, the CNN architecture typically transforms the input features in a layer-by-layer process to the final class scores. The classification outcome of the CNN classifier is then analyzed.

3.4 PROPOSED FACE IMAGE ENHANCEMENT METHOD IN UNCONSTRAINED ENVIRONMENTS

In this section, the proposed image enhancement method for facial images in unconstrained environments is presented. Image enhancement is defined as the process of improving the quality of a digitally saved image by operating on the image with computer software, such that it becomes suitable for further processing and analysis. Image enhancement belongs to the family of pre-processing methods where it has been effective. This has been shown in different kinds of literature as it is being used for various imaging tasks such as medical and satellite imaging, where it is supposed to add features to the original input image, such that it becomes enhanced both qualitatively and quantitatively. Hence, it is envisaged that image enhancement as a pre-processing approach can improve face images in unconstrained environments before the extraction and classification of features towards the recognition process. The processes involved in the proposed image enhancement method are shown in Figure 3.2. These comprise an input image being transformed to an enhanced image via a transformation function, a new evaluation function and a metaheuristic algorithm. Each process is described in more details in the following subsections.

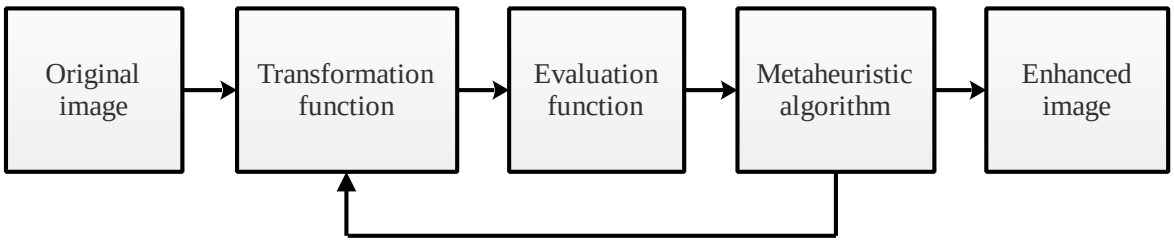


Figure 3.2. Process involved in the proposed image enhancement of facial images.

3.4.1 Transformation function

Typically, the image enhancement process is carried out on a spatial domain using the transformation function, which gives each pixel within the original image a new intensity value in order to produce the supposed enhanced image. In this thesis, the transformation function from the work of Munteanu *et al.*, [168] is used. In their work; a function applicable to each pixel at a location *(i, j),* a transformation T that used the grey level intensity of the pixel in the input image *f (i, j)* and converted it to the value *g (i, j)* i.e. the grey level intensity in the output image was proposed. The horizontal and vertical sizes of the image were depicted as H_{size} and V_{size} respectively. Hence, the transformation function T is defined in (3.1) as:

$$g(i,j) = T(f(i,j)) = k\left(\frac{M}{\sigma(i,j)+b}\right).[f(i,j) - c.m(i,j)] + m(i,j)^a \quad (3.1)$$

where, $m(i,j)$ and $\sigma(i,j)$ represent the grey level mean, and standard deviation generated for the pixels present in the neighbourhood centred at (i,j). The global mean of the image can be defined in (3.2) as:

$$M = \sum_{i=0}^{H_{size}-1} \sum_{j=0}^{V_{size}-1} f(i,j) \quad (3.2)$$

The parameters *a, b, c,* and *k,* in (3.1) have the following effects: Parameter *a* introduces a brightening bias to the output image based on the last term *m (i, j)* in (3.1). It also enables further control over the amount of smoothening effect required in the output image. Parameter *b* ensures that a zero-standard deviation value in the local neighbourhood pixels does not have a huge whitening effect on the final output image. Thus, by its introduction,

the denominator component in (3.1) typically remains nonzero. Parameter *c* allows for only a fraction of the mean *m(i,j)* to be subtracted from the original pixels of the input image. The parameter *c* controls the degree of darkening introduced in the output image. Parameter *k* is introduced to create a fair balance between pixels existing in the mid-range boundaries of the greyscale. Essentially these pixels are controlled from being made either too dark or too white.

The main challenge in using the transformation function is to determine the best parameter values [173]. The optimal parameter values are usually obtained by using optimization algorithms based on the boundary values of each parameter. Typical boundary values used are $2 \leq a \leq 2.5$; $0.3 \leq b \leq 0.5$; $0 \leq c \leq 3$ and $3 \leq k \leq 4$. These values were found to be highly effective based on an extensive empirical parameter-tuning exercise conducted in this thesis.

3.4.2 Evaluation function

In this section, the new evaluation function is presented. The Evaluation function (EF) is used to assess the quality of an enhanced image, $g(i, j)$ without human involvement. It is used to select the optimal parameter values of the transformation function automatically. To develop an evaluation function, the qualities of a well-enhanced image should be quantified [169]. These qualities include having more edge pixels in the enhanced image than the original image.

Furthermore, an enhanced image is expected to have a greater measure of information than the original image. This measure of information can be quantified by using an entropic metric such as having the histogram of the image approach a uniform distribution. This is similar to the information measure used in the histogram equalization technique. Similarly, more pixels belonging to the foreground objects in an enhanced image should be revealed better than in the original image. It is also noted that an enhanced image should contain fewer alien artefacts than the original image.

Based on the qualities of an enhanced image mentioned in this thesis, a new EF is proposed that comprises of different performance metrics. These metrics that are used to measure the qualities mentioned include the number of edge pixels, number of foreground pixels, entropic measure (EM), and the Peak to Signal Noise Ratio (PSNR). The computation of these metrics about the new EF is presented as follows:

First, the number of edge pixels, N_g, in the enhanced image is computed. To achieve this, a Sobel threshold, T_f, is automatically computed from the original image, $f(i, j)$, by using Sobel's edge detector. This threshold, T_f, is then used in the Sobel edge detector to obtain the edge intensities, $E_g(i, j)$ of the enhanced image. In addition to being invariant, T_f was considered for computing $E_g(i, j)$ to ensure a fair comparison between the original image and the different instances of the enhanced image. Thus, the number of edge pixels N_g, in the enhanced image is obtained in (3.3) as:

$$N_g = \sum_{i=1}^{H} \sum_{j=1}^{V} E_g(i, j) \tag{3.3}$$

Secondly, the number of pixels, ϕ_g, belonging to the foreground objects in $g(i, j)$ is computed. To achieve this, the variance, $\vartheta_g(i, j)$, of $g(i, j)$, and the variance, $\vartheta_f(i, j)$ of $f(i, j)$ are computed within a neighbourhood (window) having $n \times n$ pixels. A threshold value, η_f, is automatically computed for $\vartheta_f(i, j)$ by using Otsu's threshold algorithm. A representation, $D_g(i, j)$, revealing pixels belonging to the foreground objects in the enhanced image is obtained in (3.4):

$$D_g(i, j) = \begin{cases} 1 & \text{if } \vartheta_g(i, j) \geq \eta_f \\ 0 & \text{if otherwise} \end{cases} \quad \text{for } i = 1, 2, ..., H;\ j = 1, 2, ..., V \tag{3.4}$$

Thus, ϕ_g, is obtained in (3.5) as:

$$\phi_g = \sum_{i=1}^{H}\sum_{j=1}^{V} D_g(i,j) \tag{3.5}$$

Thirdly, an EM, β_g, of $g(i,j)$ is computed as shown in (3.6) as:

$$\beta_g = \begin{cases} -\sum_m \Omega_m \log(\Omega_m) & \text{for } \Omega_m \neq 0 \\ 0 & \text{for } \Omega_m \neq 0 \end{cases} \tag{3.6}$$

where Ω_m is the frequency of pixels having grey levels in the histogram bin, $m = 1,\ldots,256$

The PSNR, ρ_g, of $g(i,j)$ in (3.1) is obtained in (3.7) as:

$$\rho_g = 10\log_{10}\left[\frac{(L-1)^2}{MSE}\right] \tag{3.7}$$

where L is the maximum pixel intensity value in $g(i,j)$ and MSE is given in (3.8) as:

$$MSE = \frac{1}{H \times V}\sum_{i=1}^{H}\sum_{j=1}^{V}\left|f(i,j) - g(i,j)\right|^2 \tag{3.8}$$

Based on the parameters computed in (3.3) – (3.8), a new evaluation function, E is proposed as shown in (3.9):

$$\mathrm{E} = 1 - \exp\left(-\frac{\rho_g}{100}\right) + \frac{N_g + \phi_g}{H \times V} + \frac{\beta_g}{8} \tag{3.9}$$

where E is a linear combination of the normalized parameter values of the different metrics described in (3.3) – (3.8). By parameter normalization, each metric E is made to have values between 0 and 1. Therefore, based on this linear combination, the proposed evaluation function is characterized by a defined scale bounded by a minimum value of zero and maximum value of four. A minimum value of zero represents an entirely black enhanced image, while a maximum value of four represents an entirely white enhanced image.

3.4.3 Metaheuristic algorithms

Metaheuristic algorithms (MA) are generally employed to offer an efficient technique to find solutions for highly non-linear optimization tasks [170]. It is defined as an iterative generation process that directs a subordinate heuristic by automatically fusing various ideas for exploring the search space [108, 171]. Also, learning strategies are applied to arrange information to find efficient optimal solutions. There are different types of metaheuristic algorithm that have been shown in the literature to be effective, such as the Particle swarm optimization algorithm (PSO), Genetic algorithm (GA) algorithm, and the Cuckoo search optimization (CSO) algorithm. The following sub-sections further describe the different metaheuristic algorithms found to be effective, Also, the sections describe how the metaheuristic algorithms have been adopted in this thesis.

3.4.3.1 Particle swarm optimization

The particle swarm optimization (PSO) algorithm optimizes a task by iteratively trying to improve a candidate solution in terms of a given measure of quality. A certain task is solved by considering a population of candidate solutions termed particles, where these particles move around the search space in order to converge to an optimal solution [173].

Following the PSO process, a particle is denoted as $P_i = (P_{i1}, P_{12}, P_{i3}, ...P_{iD})$, where the number of elements, D, in a particle corresponds to the dimensional space of the particle. In this thesis, the elements of a particle typically correspond to the parameters in the transformation function (see (3.1)) whose optimal values should be found. Since four parameters were considered (i.e., a, b, c, k), the dimension was set as D = 4.

The PSO algorithm keeps a record of each particle's previous best position and its velocity along each dimension as $V_i = (V_{i1}, V_{12}, V_{i3}, ...V_{iD})$. The algorithm then iterates to find the particle with the best fitness value *(g)* in the entire population and its corresponding position [173]. The velocity of each particle is then updated using (3.10), and each particle's new position is calculated by using (3.11). At each step (iteration) of the

algorithm, the global best (GBEST) solution over the entire particle population and the local best (LBEST) solution of each particle are noted. The algorithm keeps track of the GBEST value at each iteration and converges to the final GBEST either when no better solution is found or when the pre-set number of iterations is exceeded.

$$V_{t+1} = V_t + c_1 * R(N) * (L_{best} - P_t) + c_2 * R(N) * (G_{best} - P_t) \tag{3.10}$$

$$P_{t+1} = P_t + V_{t+1} \tag{3.11}$$

where V_{t+1} is the updated particle velocity, V_t is the present particle velocity, P_t is the current particle (solution), $R(N)$ is a random number between (0,1), N is the length of the set of random numbers, c_1 and c_2 are learning factors.

3.4.3.2 Genetic algorithm

The genetic algorithm (GA) is based on the process of natural selection where better genotypes are typically passed from one generation to another. GA assumes that an appropriate solution to a task consists of an individual that produces the best fitness value (the evaluation function) over consecutive generations [144].

In using GA, the parameters of the transformation function (see (3.1)) are mapped as the genes of each chromosome structured using a string of binary values. The fitness of each chromosome is evaluated by using the evaluation function (fitness function), in (3.9). As the algorithm iterates (generation), each chromosome has the probability of producing better offspring, i.e. better solutions than in previous generations. Usually, the pool of chromosomes is initiated through a random generation process. In each iteration, a next generation is produced from previous chromosomes. The cycle then continues until a satisfied termination criterion is achieved. This can be set by using the number of generations, the amount of variation in individuals after several generations, or if a predefined fitness value is attained [144].

3.4.3.3 Cuckoo search optimization

The CSO algorithm was developed by [172], and it is among the most recent metaheuristic algorithms used for global optimization, with a process similar to the brood parasitic behaviour of certain cuckoo species. There are some optimization algorithms in the literature. However, the CSO has been considered due to its simplicity, fast convergence and for its useful capability [148, 173, 174]. These attributes are essential requirements to enhance the contrast of images effectively. The CSO algorithm utilizing Levy flight was applied in this work to search for new solutions based on the model given in (3.12):

$$a_i^{(c+1)} = a_i^{(c)} + \alpha \otimes Levy(\lambda) \tag{3.12}$$

where $a_i^{(c+1)}$ depicts the latest solutions for a cuckoo, *i*, using flight function, with λ representing the Levy walk parameter, α representing the step size associated with the scale of the problem of interest, and $\otimes$ product representing entry-wise multiplication [148]. The levy flight produces a random walk while the arbitrary step length is derived from a Levy distribution as $\text{Levy} \sim u = t^{-\lambda}, (1 < \lambda \leq 3)$.

Generally, the CSO algorithm is described as follows: every egg present in a nest denotes a solution. The algorithm uses the latest and supposedly better solutions, i.e., cuckoo in order to swap the less efficient solutions in the nests. Table 3.1 describes the algorithmic process in the CSO [173].

The PSO, GA and CSO algorithms were tested and compared to choose the most suitable optimization algorithm for use in the proposed image enhancement procedure. The results obtained are presented in Section 4.2.2. Following the empirical tests conducted, the CSO algorithm sufficed as the best performer and was then considered in the proposed image enhancement method.

3.4.4 Summary of the process involved in the proposed face image enhancement technique

A summary of the steps involved in the process of the face image enhancement technique is described as follows. The face image is acquired, resized and converted to greyscale as $f(i, j)$. The values of the lower and upper constraints for each parameter in the transformation function are defined. Then, the CSO algorithm is initiated thus: let the number of nests be n, and the dimension of each particle be D, which corresponds to the number of variables to be optimized in the algorithm. In this case, $D = 4$, depicting the four different parameters to be optimized in (3.1).

The probability of discovering an alien egg or solution in a nest is given as P_a, while the number of iterations of the CSO algorithms is given as S. The random and initial solutions (nests) for each parameter are further generated, for every CSO iteration. Levy flights are used to find a new solution for each nest, and each solution is evaluated by using (3.7). The best value among all the nests is then obtained as $\mathrm{E}_{\max}$. If $\mathrm{E}_{S+1}^{\max} > \mathrm{E}_{S}^{\max}$, the new global test is updated, otherwise, a fraction of the worst nests based on P_a, is emptied. Each new solution is updated by using (3.12), and the best nest is reserved. The optimized values of each parameter are reserved after the CSO iterations have been completed, and then they are used to obtain the final enhanced facial image. Table 3.1 describes the steps involved.

Table.3.1. Steps involved in the proposed face image enhancement technique

1. The facial image is acquired, resized and converted to greyscale as $f(i, j)$ see (3.1).
2. The values of the lower and upper constraints for each parameter in the transformation function (see (3.1)) are defined.
3. The CSO algorithm is initiated as follows:

3.1. Let the number of nests be n, and the dimension of each particle be D, which corresponds to the number of variables to be optimized in the proposed algorithm. In this case, D = 4, represent the four different parameters to be optimized in (3.1). The probability of discovering an alien egg or solution in a nest is given as P_a, while the number of iterations of the CSO algorithm is given as S.

3.2. The random and initial solutions (nests) for each parameter are generated

3.3. For every CSO iteration, until S, do the following:

3.4. Use Levy flights to obtain a new solution for each nest

3.5. Evaluate each solution (nest) by using (3.9)

3.6. The best value among all the nests is obtained as E_{max}

3.7. If $E_{S+1}^{max} > E_S^{max}$

3.8. Update the new global best

3.9. End if

3.10. Empty a fraction of the worst nests based on P_a

3.11. Update each new solution using (3.12)

3.12. Keep the best nests

3.13. Return to step 3.4 until S has been completed

4. Obtain the optimized values of each parameter after the CSO iterations have been completed
5. Use the optimized values in (3.1) to obtain a final enhanced image, $g(i, j)$

3.5 PERFORMANCE EVALUATION AND DATA SAMPLES

In this thesis, the focus is on enhancing face images in unconstrained environments. Hence, face images from three different standard benchmark face datasets were used, i.e., the Alexis Martinez (AR) face dataset, Yale face dataset (YF) and the Olivetti research laboratory face dataset (ORL). Face images that had been affected by different lighting conditions, different facial expressions, and different pose variations were selected from each face dataset. The lighting conditions present are right-light on, left-light on and both lights on. Also, for the different face expressions, face images with smile, anger, and scream have been selected. The description of the protocol used in selecting the representative images used in the evaluation of the different image enhancement methods were as follows: Essentially, six different images were selected based on our protocol. First, since our research focused on facial images and their constraints, we used six different types of representative facial condition in unconstrained environments, namely smile, anger, scream, right light illumination, left light illumination and both sides illumination. Secondly, all images in each dataset were categorized under these six different facial conditions. Thirdly, one representative image from each category was then randomly selected, thus accounting for the six different images presented in the result section per dataset.

The effectiveness of an image enhancement method can be assessed qualitatively by visualizing the enhanced output image. Furthermore, a performance evaluation should be in place that can also quantitatively assess the performance of the proposed algorithm. Hence, the following metrics are used: number of edges, number of pixels in the foreground, entropic measure, PSNR, absolute mean brightness error (AMBE) which are defined in the following sub-sections.

3.5.1 Number of edges

The number of edges produced by an image enhancement technique must be greater in the enhanced image than in the initial input. The number of edges N_g can be found as shown in (3.3).

3.5.2 Number of pixels in the foreground

An effective image enhancement technique must be able to reveal more pixels that belong to the foreground object in the enhanced image than in the original image. Hence, a higher value of the desired number of pixels in the foreground shows the effectiveness of an image enhancement technique. The number of pixels in the foreground ϕ_g can be found as shown in (3.5).

3.5.3 Entropic measure

An entropic measure is regarded as the process of quantifying the details of information in the image. A greater entropic measure value depicts a more detailed enhanced image. The entropic value of an image is independent of a different image because the comparison is done on the same image before and after the processing [175]. The entropic measure of the enhanced image β_g can be found as shown in (3.6).

3.5.4 Peak-signal-to-noise ratio

The image enhancement technique must not only improve the images but also control the level at which artifacts are introduced into the enhanced image, i.e., the level of noise should not be increased during the enhancement process. The PSNR ρ_g is used to evaluate the increase in quality between the original and the enhanced image [176]. The PSNR value can be found as shown in (3.7).

3.5.5 Absolute mean brightness error

The AMBE, ξ, is generally used to measure the rate at which the mean brightness has been preserved, and this can be represented mathematically as shown in (3.11). It shows the change in mean brightness value between the original and the enhanced images. Furthermore, the mean brightness of the original and enhanced images can be calculated as shown in (3.12) and (3.13), respectively. Therefore, a lower AMBE value is desired, while a zero AMBE value is considered the ideal result.

$$\xi = |\delta(f(i,j)) - \delta(g(i,j))| \tag{3.13}$$

$$\delta(f(i,j)) = \frac{1}{HV}\sum_{i}\sum_{j}\delta(f(i,j)) \tag{3.14}$$

$$\delta(g(i,j)) = \frac{1}{HV}\sum_{i}\sum_{j}\delta(g(i,j)) \tag{3.15}$$

where $\delta(f(i,j))$ depicts the mean brightness of the original image and $\delta(g(i,j))$ represents the mean brightness of the enhanced image. The graphical flowchart of the proposed image enhancement method is presented in Figure 3.3.

3.6 IMPROVING FACE RECOGNITION SYSTEMS BY USING A NEW IMAGE ENHANCEMENT TECHNIQUE, HYBRID FEATURES, AND THE CONVOLUTIONAL NEURAL NETWORK

In this section, the approach to designing the proposed FRS is described, with the aim of making the recognition performance in unconstrained environments better. The proposed FRS consists of the pre-processing, extraction of features; and feature classification stages. At the pre-processing stage of the FRS, the proposed image enhancement method is introduced and also, at the feature extraction stage, a selection of hybrid features consisting of the pyramid histogram of gradients, edge histogram descriptors, and local binary patterns is presented. Lastly, at the classification stage, the convolutional neural network is used. Experiments to confirm the effectiveness of the system are carried out on three different standard face datasets, i.e. AR, Yale, and LFW. Figure 3.4 shows the framework

of the proposed FRS. In previous sections, the image enhancement method and the different feature extraction methods have been presented and discussed. Hence, the approach to selecting the hybrid features and the convolutional neural network used are presented.

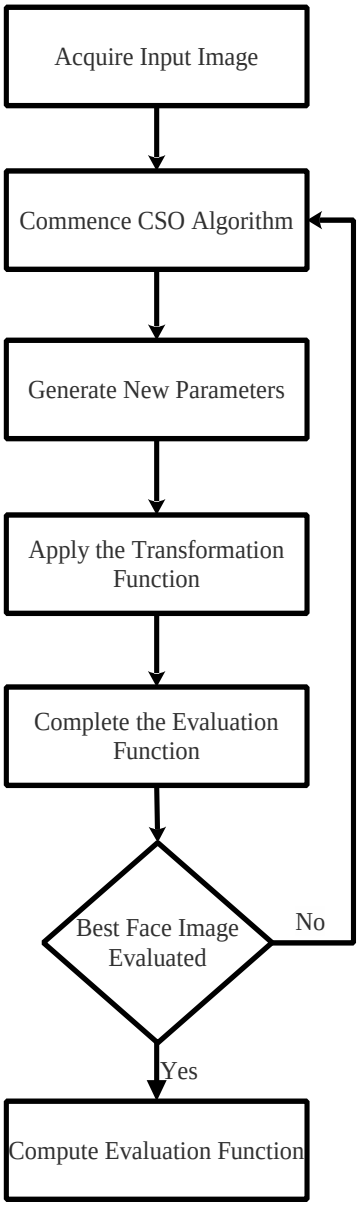


Figure 3.3. Graphical flowchart representation of the proposed image enhancement method.

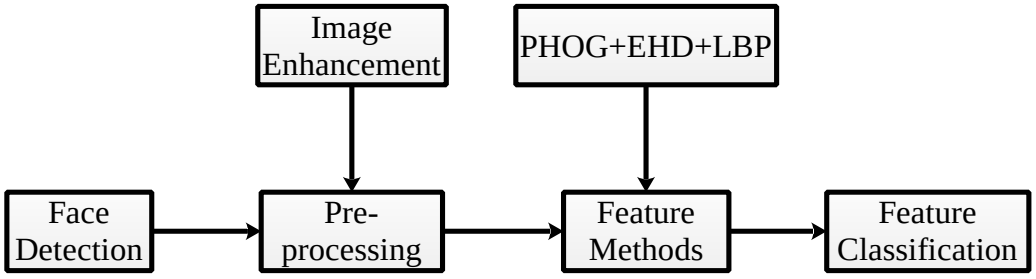


Figure 3.4. The framework of proposed face recognition system

3.6.1 Selected hybrid features

The aim of the feature extraction stage is to extract features from the enhanced facial images in unconstrained environments. This would be effective for classification and, improve recognition performance when using the CNN. To achieve this, three individual feature methods were considered, which were then fused together to become a set of hybrid features. The extensive experiment was carried out with different hybrid methods to confirm the selected and effective method. These feature methods included the PHOG; EHD and LBP. Because of the different individual properties of these methods, they can capture local appearance information; they are invariant to grey changes and can extract important shape and edge information from the enhanced facial image. The selected methods each produce a different number of features, for example: PHOG = 630; EHD = 80; LBP = 755; PHOG+EHD = 710; PHOG+LBP = 1386; and PHOG+LBP+EHD = 1467.

However, not all the features produced are relevant; hence, the information gain ranking method is used to select the most important features produced from the different feature methods. Therefore, Q used in the CNN classifier per selected features (see section 3.3.2) is stated as follows: Q = 26 for PHOG; Q = 70 for LBP; Q = 120; PHOG+EHD= 125; PHOG+LBP=150; and PHOG+LBP+EHD=170.

3.6.2 Feature classification technique

To further confirm the effectiveness of the proposed pre-processing method and selection of the hybrid features, the CNN was utilized as the classification technique in this thesis

because of its recent and effective performance in image processing and computer vision tasks. In addition, it was considered because of its ability to be inspired by biological processes where the connecting patterns amongst neurons are similar to the arrangement of the human visual cortex [177-179]. This was also confirmed by the results of the comparative analysis carried out on the different FRS in unconstrained environments. The description of the way the CNN classification methods operate can be seen in literature such as [138, 180].

In this thesis, the state-of-the-art CNN model architecture in [167] was adopted. The architecture consisted of six layers including the input and the output layers. Other layers included the Convolutional, Rectified Linear Unit (RELU), Pooling, and the Fully-Connected layers. An array of numbers representing the input features was considered in the input layer while another set of arrays of numbers was produced as output. The Input layer holds the [1 x Q] raw feature values of the face images used in this work. The number of features, Q, typically changes based on the type of features used. The Convolutional layer consists of 12 filters, which results in a volume of [1 x Q x 12]. The RELU layer utilizes a max (0, x) thresholding at zero, thus leaving the size of the volume unchanged as [1 x Q x 12]. Down-sampling operations along spatial dimensions are performed in the Pooling layer resulting in a volume size of [1 x Q/2 x 12]. The Fully-Connected layer consists of a [1 x 1 x N] volume size, where N denotes the number of class scores associated with the different image datasets considered in this work. In this way, the CNN architecture typically transforms the input features in a layer-by-layer process to the final class scores.

3.7 CHAPTER SUMMARY

In this chapter, the different approaches to the research objectives of this thesis are presented. In Section 3.1, the different face recognition techniques, i.e., the different feature and classification methods used in this thesis are described. In Section 3.2, the different methods and approach used to design the proposed image enhancement methods are described. These include the transformation function, evaluation function, and the

metaheuristic algorithms. Finally, in Section 3.3., the different approaches to design the proposed FRS are presented. These include the image enhancement method, the selection of hybrid features and the convolutional neural network.

CHAPTER 4 RESULTS

4.1 COMPARATIVE ANALYSIS OF FACE RECOGNITION TECHNIQUES IN UNCONSTRAINED ENVIRONMENTS

In this result section, the results generated from the comparative analysis of the different FRS in unconstrained environments are presented. The comparative analysis was carried out on two standard face datasets, where the different constraints within the face datasets were also considered. These face datasets have been shown to be efficient in confirming the effectiveness of various face recognition techniques in literature. These face recognition techniques involved the different feature extraction and classification methods as described in Section 3.1. The feature extraction techniques used are LBP, EHD, GF and PHOG; while the classification techniques used include CNN, SVM, NB, DT and RF. The average recognition rates were used to evaluate the performance of the different FRS.

4.1.1 Recognition performance on the Yale face dataset

Table 4.1 presents the results obtained by using the different face recognition techniques on the Yale face dataset regardless of the facial constraints. When observed using the LBP features were used, an average recognition rate of 63.7% was achieved with the DT classifier, which presented the lowest recognition performance on this dataset. A slight increase of a 65.8% average recognition rate was achieved with the NB classifier method, while, the highest recognition rate of 71.3% was achieved with the CNN classifier method. Moving on to the EHD features, the lowest average recognition rate of 52.7% was achieved with the NB classifier method, while the highest recognition rate of 56.1% was achieved with the CNN. Similarly, it was noticed that the PHOG feature achieved the

highest average recognition rate across all classifiers. With the SVM classifier, a recognition rate of 76.3% was achieved, 65.4% with the RF, 68.9% with the DT, 76.3% with the SVM, and finally 79.4% with the CNN classification method.

Table 4.1. Recognition performance on Yale.

Feature Methods	**Classification Methods**				
	CNN	SVM	NB	DT	RF
LBP	71.3	67.8	65.8	63.7	66.9
EHD	56.1	54.8	52.7	53.8	55.7
GF	55.6	52.9	51.7	52.8	54.9
PHOG	79.4	76.3	74.8	68.9	65.4

4.1.2 Recognition performance based on illumination, on the Yale face dataset

Table 4.2 presents the results achieved with the different face recognition techniques on the Yale face dataset taking into consideration the illumination constraints. Here, it was observed that with the GF feature, the lowest performance was achieved across the classification methods. With the SVM, an average recognition rate of 51.8% was achieved, while with the RF and CNN, recognition rates of 53.5% and 52.9% respectively were achieved. The highest recognition rate across all classification methods was achieved by using the PHOG features. With the SVM, an average recognition rate of 73.8% was achieved, while there was an increase with the CNN, with an average recognition rate of 75.8%.

Table 4.2. Recognition performance on Yale, based on illumination.

Feature Methods	**Classification Methods**				
	CNN	SVM	NB	DT	RF
LBP	69.7	66.5	62.8	64.7	69.8
EHD	54.8	53.6	53.7	51.6	54.5
GF	52.9	51.8	52.8	50.2	53.5
PHOG	75.8	73.8	71.6	70.3	69.2

4.1.3 Recognition performance based on expression on the Yale face dataset

The results achieved by the different FR systems on the Yale face dataset when the issue of expression was considered are presented in Table 4.3. With the LBP features, an average recognition rate of 63.4% was achieved with the NB classifier, while with the DT, the recognition rate was 61.9%. With the RF, the best recognition performance with this feature was 72.8%. When the EHD features were used across all classifiers, the performance was the lowest, with NB achieving the lowest recognition rate of 49.6%. in contrast, using the PHOG features across all the classifiers produced the best recognition rate. With the RF classifier, a recognition rate of 74.1% was achieved, while the highest was 76.7% with the CNN.

Table 4.3. Recognition performance on Yale, based on Expression.

Feature Methods	Classification Methods				
	CNN	SVM	NB	DT	RF
LBP	70.2	68.1	63.4	61.9	72.8
EHD	53.8	56.7	49.6	50.4	53.2
GF	53.4	52.2	50.7	52.7	51.9
PHOG	76.7	73.9	70.6	69.8	74.1

4.1.4 Recognition performance on the AR face dataset

Table 4.4 presents the results achieved with the different face recognition techniques on the AR face dataset, regardless of the constraints. There was an increase in performance as compared to the Yale face dataset except for the LBP method. With the EHD method, an average recognition rate of 58.9% was achieved when the NB classification method was used and with the SVM classification method, a recognition rate of 60.7% was achieved, while with the CNN method, the average recognition rate was 62.6%. The different classification methods performed best with the PHOG features. With the DT classification method, an average recognition rate of 71.2% was achieved, while the best performance was achieved with the CNN method, with a recognition rate of 83.1%.

Table 4.4. Recognition performance on AR

Feature Methods	Classification Methods				
	CNN	SVM	NB	DT	RF
LBP	45.5	43.4	42.6	43.7	44.1
EHD	62.6	60.7	58.9	58.6	59.2
GF	54.8	52.7	50.8	52.9	53.7
PHOG	83.1	79.8	73.6	71.2	81.7

4.1.5 Recognition performance based on lighting, on the AR face dataset

The results from the analysis of the different face recognition techniques on the AR face dataset, taking into consideration the lighting constraints are presented in Table 4.5. With the LBP, the lowest recognition performance was achieved across the classification methods. With the NB classification method, an average recognition rate of 45.1% was achieved, while with the DT and RF classification methods, the average recognition rates were 44.8% and 46.8% respectively. With the GF features, an average recognition rate of 52.7% was achieved with the SVM classification method, while an average recognition rate of 54.8% was realised with the CNN classification method. The overall best performance considering these conditions was 83.1%, achieved when the PHOG features and the CNN classification method were used.

Table 4.5. Recognition performance on AR, based on lighting

Feature Methods	Classification Methods				
	CNN	SVM	NB	DT	RF
LBP	47.6	45.1	43.2	44.8	46.8
EHD	63.4	62.9	54.7	56.8	62.9
GF	56.2	54.9	52.1	51.6	55.7
PHOG	86.2	83.9	74.2	77.2	81.7

4.1.6 Recognition performance based on expression, on the AR face dataset

Table 4.6 presents the results achieved with the different face recognition techniques on the AR face dataset, taking into consideration the expression constraints. Similar to the previous results, considering previous constraints on this dataset, LBP features produced

the lowest average recognition performance with the different classification methods. With the EHD features, the SVM classification method achieved an average recognition rate of 63.8%, while the RF and CNN classification methods reached average recognition rates of 62,7% and 65.7% respectively.

Table 4.6. Recognition performance on AR, based on expression.

Feature Methods	**Classification Methods**				
	CNN	SVM	NB	DT	RF
LBP	47.9	46.4	41.2	45.2	47.1
EHD	65.7	63.8	55.7	59.4	62.7
GF	59.2	58.1	53.7	56.2	60.4
PHOG	88.2	86.4	62.1	68.3	87.1

4.1.7 Recognition performance based on upper face occlusion, on the AR face dataset

Table 4.7, shows results from the comparative analysis of the different face recognition techniques with the upper face occlusion on the AR face dataset. The NB classification method, with the GF features; achieved an average recognition rate of 45.9%, while the DT and SVM classification methods, achieved average recognition rates of 50.4% and 57.3% respectively. It is observed that the PHOG features with the different classification methods performed better than other features. With the RF classification method, an average recognition rate of 67.4% was achieved, while SVM and CNN, achieved average recognition rates of 69.2% and 72.8% respectively.

Table 4.7. Recognition performance on AR, based on upper face occlusion

Feature Methods	**Classification Methods**				
	CNN	SVM	NB	DT	RF
LBP	58.6	55.9	47.4	50.3	56.1
EHD	66.2	63.8	52.9	54.9	62.9
GF	58.2	57.3	45.9	50.4	56.2
PHOG	72.8	69.2	53.1	55.2	67.4

4.1.8 Recognition performance based on lower face occlusion on the AR face dataset

Table 4.8, shows results from the comparative analysis of the different face recognition techniques on the AR face dataset, considering lower face occlusion. The RF classification method, with the EHD features, achieved an average recognition rate of 64.1%, while with the SVM, the average recognition rate was 63.7%. The NB classification method, with the PHOG features, achieved an average recognition rate of 52.6%, while average recognition rates of 66.9% and 69.5% were achieved with the RF and CNN respectively.

Table 4.8. Recognition performance on AR, based on lower face occlusion

Feature Methods	**Classification Methods**				
	CNN	SVM	NB	DT	RF
LBP	55.4	54.8	41.9	48.8	53.9
EHD	65.8	63.7	51.3	53.9	64.1
GF	52.3	50.1	42.1	45.7	51.9
PHOG	69.5	67.8	52.6	57.8	66.9

4.2 A NEW EVALUATION FUNCTION FOR FACE IMAGE ENHANCEMENT IN UNCONSTRAINED ENVIRONMENTS USING METAHEURISTIC ALGORITHMS.

In this section, the results of the various experiments based on the approaches to implementing the proposed image enhancement method are presented and discussed. The results have been generated using the new evaluation function in (3.9) as a component of the proposed image enhancement technique. First, an experiment to determine the choice of the CSO parameter was carried out to enable us to identify the ideal parameter to be used. Then, simulations of the different metaheuristic algorithms were carried out to confirm their performance based on the fitness value and time of convergence.

Furthermore, to verify the effectiveness of the proposed evaluation function, experiments were carried out to compare it with other evaluation functions using the CSO algorithm. Finally, to confirm the efficacy of the image enhancement method, quantitative and

qualitative comparisons based on standard performance metrics of the proposed enhancement technique with other state-of-the-art methods was carried out on different standard benchmark datasets, such as the AR, Yale and Olivetti research laboratory face datasets.

4.2.1 Choice of CSO parameters

To select the ideal value of the parameter effectively, Pa, for the CSO algorithm, an experiment with different values ranging between Pa = 0.1 to 1.0 was conducted. Based on the analysis result shown in Figure 4.1 an ideal Pa of 0.2 was chosen because it produced the earliest convergence at the 30th iteration. Also, a fitness value of 1.405 was achieved, which outperformed other Pa values.

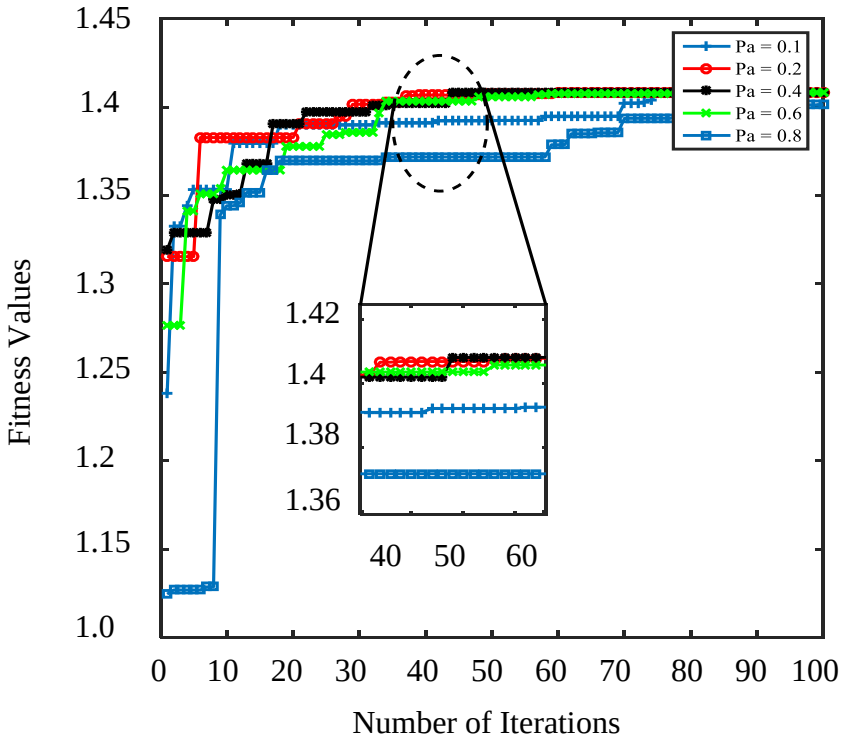


Figure 4.1. Choice of CSO parameter

4.2.2 Evaluation of different metaheuristic algorithms

An evaluation of the different metaheuristic algorithms was carried out to confirm the selection of the metaheuristic algorithm used in this thesis. The selected algorithms included CSO, PSO, and GA. To achieve this, the number of iterations for each metaheuristic algorithm was set at 100, which is plotted against the fitness function value for each metaheuristic algorithm. Also, to further confirm the selection, the different metaheuristic algorithms were compared with the proposed evaluation function based on all the performance evaluation metrics, such as the number of pixels in the foreground, number of edges, PSNR, entropic measure, AMBE, and fitness value. To prevent bias in the experiment, three different images from the AR dataset were selected, on which the performance of the different algorithms were evaluated as shown in Figures 4.2 - 4.4.

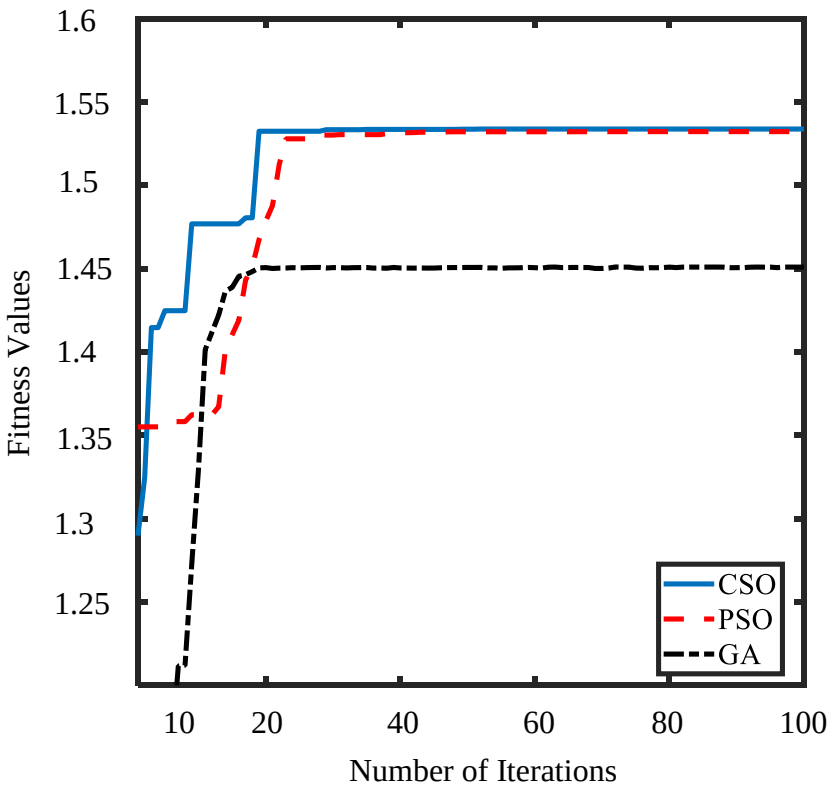


Figure 4.2. Performance and convergence analysis of the different metaheuristic algorithms on Image 1

Figure 4.2., with the selected image 1, shows that the cuckoo search optimization algorithm outperformed other metaheuristic algorithms. The fitness value achieved by the CSO was

1.534; which was closely followed by the PSO with a fitness value of 1.532, and, lastly the GA with a fitness value of 1.4509. Also, considering the convergence analysis, for image 1, the earliest convergence occurred with the CSO at the 20th iteration. This was followed closely by the GA, with a convergence at the 25th iteration and finally, the PSO with a convergence at the 30th iteration.

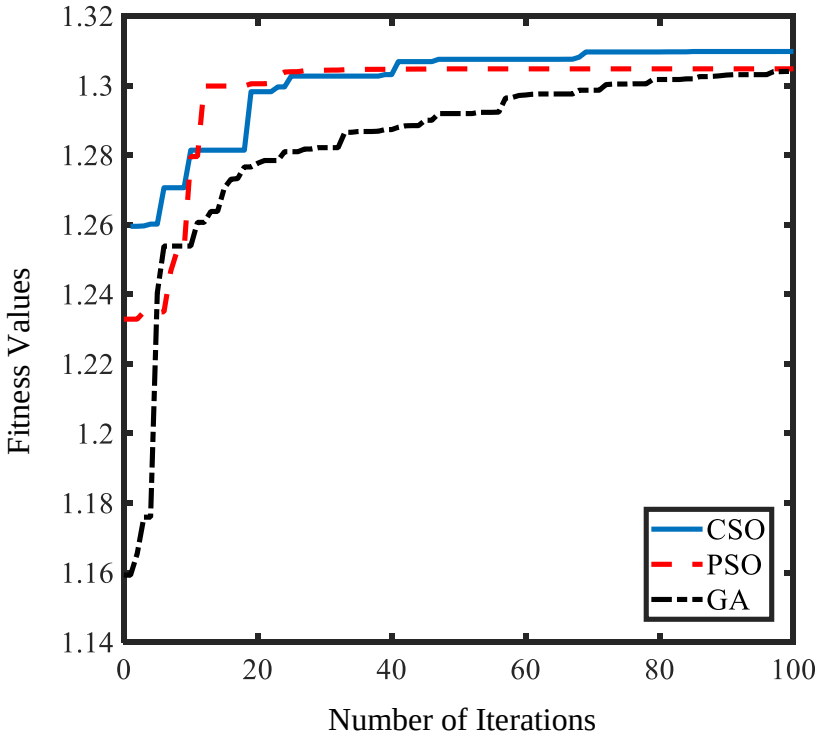


Figure 4.3. Performance and convergence analysis of the different metaheuristic algorithms on Image 2

Figure 4.3, with the selected image 2, shows that the cuckoo search optimization algorithm outperformed other metaheuristic algorithms. The fitness value achieved by the CSO was 1.31; which was closely followed by the PSO with a fitness value of 1.304, and, lastly the GA with a fitness value of 1.302. Also, considering the convergence analysis, for image 1, the earliest convergence occurred with the CSO at the 63rd iteration. This was closely followed by the GA, with a convergence at the 85th iteration and finally, the PSO with a convergence at the 95th iteration.

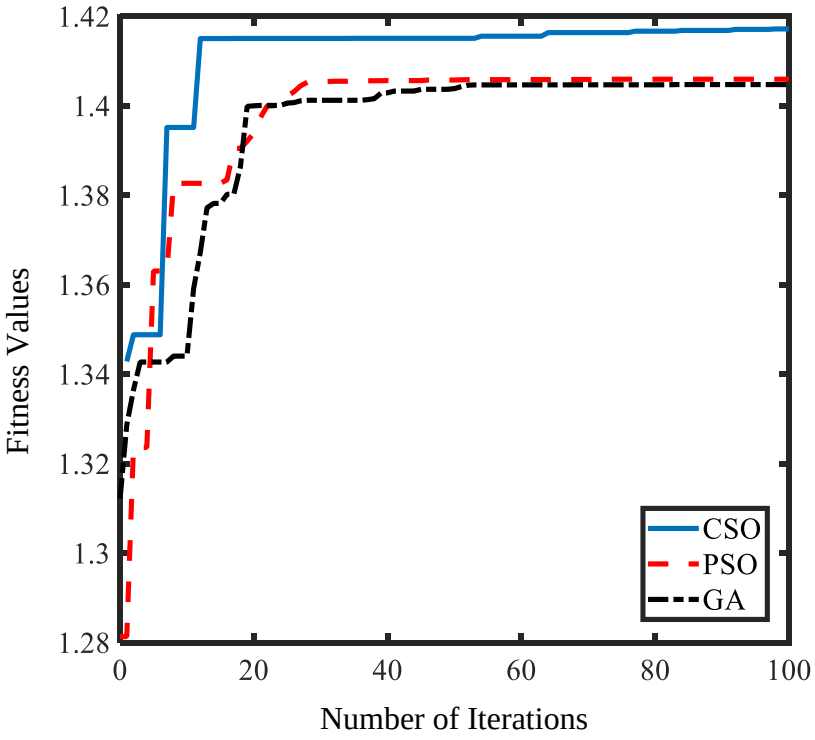


Figure 4.4. Performance and convergence analysis of the different metaheuristic algorithms on Image 3

In Figure 4.4. with the selected image 3, shows that the cuckoo search optimization algorithm outperformed other metaheuristic algorithms. The fitness value achieved by the CSO was 1.417; which was closely followed by the PSO with a fitness value of 1.406, and, lastly once again, the GA with a fitness value of 1.404. Also, as regards the convergence analysis, for image 3, the earliest convergence occurred with the CSO at the 95th iteration. This was closely followed by the PSO, with a convergence at the 98th iteration, and, finally with the GA with a convergence at the 99th iteration.

Table 4.9. Comparison of the different metaheuristic algorithms with the proposed evaluation function based on all the performance evaluation metrics

Metrics	CSO + Proposed EF			PSO + Proposed EF			GA +Proposed EF		
	Img1	Img2	Img3	Img1	Img2	Img3	Img1	Img2	Img3
ϕ_g	6026	2014	4567	4267	2022	4318	5806	2019	4291
N_g	3408	2422	2475	2475	2412	2437	3148	2410	2421
ρ_g	11.9	13.2	12.4	12.4	12.5	11.4	10.2	12.5	11.5
β_g	7.7	7.7	7.7	7.7	7.7	7.7	7.2	7.7	7.5
ξ	0.0	0.1	0.0	0.0	0.1	0.0	0.2	0.1	0.0
E	1.534	1.310	1.417	1.532	1.304	1.406	1.451	1.302	1.404

Legend: ϕ_g - Number of pixels in the foreground; N_g - Number of edge pixels; ρ_g - PSNR; β_g - Entropic measure; ξ - Absolute mean brightness error; E - Fitness value.

The various performance evaluation metrics have been defined earlier in Section 3.2.4 to enable us to observe the performance of the different metaheuristic algorithms with the proposed evaluation function.

In Table 4.9, the CSO with the proposed evaluation function shows the following exciting and useful results: for image 1, a foreground value of 6026 pixels was achieved for the CSO. This outperformed PSO and GA which achieved values of 4267 and 5806 respectively. For image 2, a total of 2433 edge values were achieved with the CSO, which outperformed PSO and GA with the proposed evaluation function achieving 2412 and 2410 respectively. Likewise, for image 3, the highest number of pixels in the foreground was 4567 when the CSO was used with the proposed algorithm. This surpassed the values of 4318 and 4291 generated by PSO and GA respectively. Furthermore, the values generated for the number of edges on the different images by the CSO algorithm with the proposed evaluation function surpassed those of all the other techniques.

As regards the fitness value, for image 1, a value of 1.534 was achieved, while values of 1.532 and 1.451 were produced with the PSO and GA respectively. For, image 2, a fitness value of 1.310 was achieved with the CSO algorithm, while the values with PSO and GA were 1.304 and 1.302 were respectively. For image 3, a fitness value of 1.417 was achieved with the CSO algorithm, while the fitness values achieved with the PSO and GA algorithms were 1.406 and 1.404 respectively. This implies that the CSO algorithm with the proposed evaluation function outperforms other metaheuristic algorithms. This could improve the original images, thereby revealing more information from the image.

To confirm the selection of the CSO metaheuristic algorithm further, a qualitative analysis of the enhanced images produced by the different metaheuristic algorithms with the proposed evaluation function is shown in Figures 4.5 – 4.7.

(a) (b) (c) (d)

Figure 4.5. Qualitative comparison of the different metaheuristic algorithm with the proposed evaluation function on Images 1 (a) original, (b) CSO + proposed evaluation method, (c) PSO + proposed evaluation method and (d) GA + proposed evaluation method.

(a) (b) (c) (d)

Figure 4.6. Qualitative comparison of the different metaheuristic algorithm with the proposed evaluation function on Images 2 (a) original, (b) CSO + proposed evaluation method, (c) PSO + proposed evaluation method and (d) GA + proposed evaluation method.

(a) (b) (c) (d)

Figure 4.7. Qualitative comparison of the different metaheuristic algorithm with the proposed evaluation function on Images 3 (a) original, (b) CSO + proposed evaluation method, (c) PSO + proposed evaluation method and (d) GA + proposed evaluation method.

4.2.3 Comparison of different evaluation functions

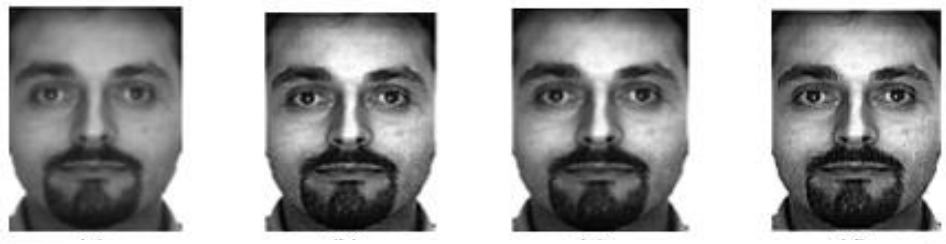
In this section, results showing the performance of the proposed evaluation function with the cuckoo search optimization algorithm is presented. The comparison was carried out by using the CSO algorithm, the different evaluation functions used in Munteanu (MCSO) [168] and Ye (YeCSO) [181] on three different selected face images that varied amongst individuals as shown in Table 4.10. It is seen that using the proposed evaluation function with outperformed the other evaluation functions with the CSO algorithm.

Table 4.10. Quantitative Comparison for different evaluation functions

Metrics	MCSO			YeCSO			Proposed +CSO		
	Img1	Img2	Img3	Img1	Img2	Img3	Img1	Img2	Img3
ϕ_g	1867	1082	2301	762	927	1935	2104	1116	4827
N_g	537	361	630	1314	1673	1719	2306	2057	3368
ρ_g	14.9	13.3	17.0	17.7	15.7	17.8	14.5	13.8	12.9
β_g	7.8	7.7	7.7	7.8	7.6	7.8	7.8	7.7	7.8
ξ	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.1	0.0
E	1.235	1.160	1.279	1.242	1.227	1.323	1.333	1.252	1.510

As shown in Table 4.10, the different evaluation functions were compared based on all the performance evaluation metrics when the CSO algorithm was used. The comparison confirmed the effectiveness of the proposed evaluation and yielded the following good results: across the different images 1, 2 and 3, the proposed evaluation function produced the highest foreground value of 2104, 1116 and 4827 respectively as compared to the other evaluation functions. Likewise, values generated for the number of edges by the proposed algorithms surpassed those of other evaluation function techniques with high values of 2304, 2057 and 3368 for images 1, 2 and 3 respectively.

Furthermore, the fitness values generated by the different algorithms were analyzed, and they showed that the proposed algorithm achieved the highest fitness value for all the images. For example, the fitness value for image 1 with the proposed evaluation function achieved a value of 1.333, while for MCSO, and YeCSO the values were 1.235 and 1.242 respectively. To further confirm the effectiveness of the proposed evaluation function, the different evaluation functions were compared and enhanced images they had produced were qualitatively analyzed as shown in Figures 4.8 – 4.10. Therefore, it is evident that the proposed evaluation function outperformed other algorithms.



(a) (b) (c) (d)

Figure 4.8. Qualitative comparison of the different evaluation functions with the CSO algorithm on Images 1: (a) original, (b) MCSO, (c) YeCSO, (d) Proposed + CSO.

(a) (b) (c) (d)

Figure 4.9. Qualitative comparison of the different evaluation functions with the CSO algorithm on Images 2: (a) original, (b) MCSO, (c) YeCSO, (d) Proposed + CSO.

(a) (b) (c) (d)

Figure 4.10. Qualitative comparison of the different evaluation functions with the CSO algorithm on Images 3: (a) original, (b) MCSO, (c) YeCSO, (d) Proposed + CSO.

4.2.4 Comparison of different image enhancement methods

The experiment carried out in this section was done to confirm the performance of the proposed algorithm by comparing it with other state-of-the-art image enhancement methods. These methods included LCS, HE, image intensity adjustment (IIA), Munteanu with particle swarm optimization (MP), Munteanu with genetic algorithm (MG), Brightness preserving dynamic fuzzy histogram equalization (BPD), Contrast-limited adaptive histogram equalization (CHE), and Low-light image enhancement (LLIE) as shown in the result sections that follow. This experiment was carried out on the different standard benchmark face datasets by selecting six different face images with various real-world facial conditions. The chosen face images selected from each face dataset were labelled images (1 - 6) from each face dataset, and they represented multiple faces with different real-life conditions. The measure of performance used are based on the metrics described in Section 3.5.

4.2.4.1 Performance based on the number of pixels in the foreground value

In Tables 4.11 – 4.13, the results from comparing the different enhancement methods based on the number of pixels in the foreground value on the different standard benchmark face dataset are presented. It is important to note that a higher number of pixels in the foreground shows the effectiveness of an image enhancement technique [176].

Table 4.11. Number of pixels in the foreground value obtained with different enhancement methods on the AR face dataset.

DB		LCS	HE	IIA	MP	MG	BPD	CHE	LLIE	Prop
AR	Img1	451	1006	451	2565	2249	391	1560	977	3294
	Img2	399	1220	399	2407	2388	334	1248	655	2565
	Img3	937	2174	937	3582	3669	891	2660	1734	5424
	Img4	989	1572	989	3491	3475	740	2283	895	5956
	Img5	636	677	636	2902	2912	462	2134	961	6865
	Img6	383	887	383	2317	2324	190	651	371	3947

On the AR face dataset, for image 1, the proposed image enhancement method yielded a value of 3294, MG followed by a value of 2249. For this image, the BPD enhancement method yielded the lowest value of 391. For image 2, values yielded by the proposed method and the MP method came closer to each other with a value of 2565 for the proposed method and 2407 for the MP method. For image 5, there was a significant difference in performance, with the proposed method yielding a value of 6865, and followed by MG with a value of 2912, this was closely yielded by a value of 2902 produced by the MP enhancement method.

Table 4.12. Number of pixels in the foreground value achieved by using different enhancement methods on the Yale face dataset.

DB		LCS	HE	IIA	MP	MG	BPD	CHE	LLIE	Prop
YF	Img1	1308	9563	1308	5693	5955	950	1249	1439	13117
	Img2	1515	13388	1515	8914	9129	1359	1966	1175	17694
	Img3	2285	6298	2285	8547	7003	2098	2010	3006	12425
	Img4	1589	14331	1589	7806	8105	1347	2004	1034	18071
	Img5	1712	4731	1712	5936	5324	1508	1356	834	9751
	Img6	1887	5555	1887	6824	6080	1686	1515	1004	10483

On the Yale face dataset, for image 1, an outstanding performance was produced by the proposed method with a value of 13117. This was followed by the performance of the HE enhancement method, which yielded a value of 9563. For image 2, the HE enhancement method yielded a value of 13388; the LCS method produced a value of 1515; MG produced a value of 9129, and the proposed method performed best by producing a value of 17694. The proposed method performed the best overall on image 4 with a value of 18071, while the overall lowest performance was that of the LLIE method on image 5.

Table 4.13. Number of pixels in the foreground value when different enhancement methods on the ORL face dataset were used.

DB		LCS	HE	IIA	MP	MG	BPD	CHE	LLIE	Prop
ORL	Img1	482	489	482	1795	1909	333	1066	623	3873
	Img2	1938	2191	1938	5277	5279	1318	4209	2873	7618
	Img3	1784	2167	1784	3926	3928	1246	3218	2385	6499
	Img4	1438	1556	1438	3324	2877	1014	2658	1552	6248
	Img5	1330	1016	1330	4312	4353	1025	2513	1958	5946
	Img6	915	738	915	1792	1912	559	1240	1152	4705

As shown in Table 4.13, on the ORL face dataset, the proposed method also outperformed all other standard enhancement methods. For image 1, a value of 3873 was achieved with the proposed enhancement method, which was followed by the MG method with a value of 1909. On image 2, the overall best performance was achieved with the proposed method with a value of 7618, while the BPD method scored the lowest performance rate at 333.

4.2.4.2 Performance based on the number of edges

Tables 4.14 – 4.16, display the results from comparing the different enhancement methods based on the number of edges on the various standard benchmark face datasets. However, it is important to note that a higher number of edges is required for an effective image enhancement technique.

Table 4.14. Number of edges obtained by using different enhancement methods on the AR face dataset

DB		LCS	HE	IIA	MP	MG	BPD	CHE	LLIE	Prop
AR	Img1	624	1240	624	2395	2175	684	1981	1362	2889
	Img2	522	1375	522	2089	2071	514	1882	1010	2223
	Img3	696	1315	696	1976	2016	683	1866	1186	2756
	Img4	781	1017	781	1993	1974	665	1627	992	3044
	Img5	726	672	726	2072	2072	570	1915	1073	3776
	Img6	700	1496	700	1898	1901	733	1657	671	3247

On the AR face dataset, the proposed method outperformed other standard enhancement methods. For image 1, the proposed method achieved a value of 2889, which was followed by the MP method with a value of 2395. On image 5, the proposed method performed the best with a value of 3776 outperforming all other methods. BPD was the method with the lowest performance with a value of 514.

Table 4.15. Number of edge values obtained by using different enhancement methods on the Yale face dataset.

DB		LCS	HE	IIA	MP	MG	BPD	CHE	LLIE	Prop
YF	Img1	1842	4580	1842	5028	5213	1604	2946	2546	9247
	Img2	1699	6158	1699	5727	5804	1514	4064	2631	9106
	Img3	1756	2829	1756	4825	4013	1677	2151	2328	6517
	Img4	1521	7184	1521	5464	5509	2514	4234	2604	9013
	Img5	1349	2103	1349	3780	3409	1217	1845	2209	5461
	Img6	1409	2377	1409	3982	3514	1307	2153	2095	5490

On the YF dataset, the proposed method also outperformed all other enhancement methods. For image 3, a value of 1756 was achieved with the LCS enhancement method, while the proposed method achieved a value of 6517. The overall best performance was

achieved by the proposed method with a value of 9013 on image 4; while the BPD achieved the lowest performance with a value of 1217.

Table 4.16. Number of edge values obtained by using different enhancement methods on the ORL face dataset.

DB		LCS	HE	IIA	MP	MG	BPD	CHE	LLIE	Prop
ORL	Img1	458	546	458	1124	1195	354	911	569	1898
	Img2	603	713	603	1633	1634	395	1293	890	2532
	Img3	714	812	714	1325	1326	517	1122	877	2203
	Img4	668	751	668	1334	1197	519	1346	753	2313
	Img5	566	465	566	1436	1450	470	955	725	2021
	Img6	602	524	602	944	987	424	945	775	1865

As shown in the Table 4.16, on the ORL face dataset, the number of edge values produced by the different enhancements was lower than on the previous face datasets. However, like on the previous face datasets, the proposed enhancement method produced the highest number of values across all images.

4.2.4.3 Performance based on the PSNR value

Tables 4.17 – 4.19 display the results achieved when the different enhancement methods based on the number of edges on the different standard benchmark face datasets were compared. In this case, a lower PSNR value was required for an effective image enhancement technique.

Table 4.17. PSNR values achieved by using different enhancement methods on the AR face datasets.

DB		LCS	HE	IIA	MP	MG	BPD	CHE	LLIE	Prop
AR	Img1	23.14	17.92	23.13	14.43	15.23	27.16	17.02	13.99	13.26
	Img2	21.18	15.38	21.84	12.40	12.39	34.80	16.35	14.28	12.63
	Img3	25.49	19.09	25.49	15.75	15.32	32.37	18.72	13.83	12.44
	Img4	24.01	14.13	24.01	12.85	12.43	40.56	16.87	17.73	8.89
	Img5	24.36	15.56	24.36	14,53	14.41	40.53	17.54	17.65	9.15
	Img6	24.26	9.88	24.26	14.19	14.17	29.95	17.13	21.35	8.24

Table 4.18. PSNR values achieved by using different enhancement methods on the YF face datasets.

DB		LCS	HE	IIA	MP	MG	BPD	CHE	LLIE	Prop
YF	Img1	27.07	12.47	27.07	14.10	13.27	34.11	23.45	26.01	8.07
	Img2	36.54	12.58	36.54	13.71	13.75	35.48	20.53	15.10	8.19
	Img3	37.27	17.08	37.27	13.20	6.61	37.80	21.75	23.83	9.76
	Img4	42.58	13.73	42.58	14.65	14.04	25.33	19.81	14.25	8.84
	Img5	39.03	18.11	39.03	2.296	3.32	32.59	20.95	14.96	9.49
	Img6	35.34	16.07	35.34	13.39	5.06	37.97	21.70	17.01	9.02

Table 4.19. PSNR values achieved by using different enhancement methods on the ORL face datasets.

DB		LCS	HE	IIA	MP	MG	BPD	CHE	LLIE	Prop
ORL	Img1	25.32	20.85	25.32	17.55	16.44	31.98	18.15	16.09	11.63
	Img2	26.78	22.48	26.78	14.86	14.83	41.33	16.41	14.60	10.25
	Img3	25.59	22.10	25.59	16.31	16.21	37.62	17.69	18.66	10.18
	Img4	23.78	19.59	23.78	15.44	16.94	33.86	17.82	18.59	10.52
	Img5	25.62	28.24	25.62	15.37	15.38	36.75	17.31	16.07	12.15
	Img6	22.70	21.72	22.70	18.52	18.13	32.63	17.74	15.35	11.82

4.2.4.4 Performance based on the entropic measure value

Tables 4.20 – 4.22, display the results achieved when the different enhancement methods based on the number of edges on the different standard benchmark face datasets were compared. In this case, a lower PSNR value was required for an effective image enhancement technique.

Table 4.20. Entropic measure values achieved by using different enhancement methods on the AR face dataset.

DB		LCS	HE	IIA	MP	MG	BPD	CHE	LLIE	Prop
AR	Img1	7.446	5.957	7.446	7.738	7.742	7.232	7.815	7.439	7.731
	Img2	7.193	5.950	7.193	7.657	7.658	6.857	7.689	7.180	7.839
	Img3	7.413	5.931	7.413	7.665	7.661	7.066	7.790	7.249	7.780
	Img4	7.635	5.972	7.635	7.782	7.786	7.231	7.760	7.017	7.859
	Img5	7.561	5.827	7.561	7.575	7.572	7.187	7.708	7.051	7.826
	Img6	7.069	5.884	7.069	7.465	7.463	6.648	7.592	6.395	7.351

Table 4.21. Entropic measure values achieved by using different enhancement methods on the Yale face dataset.

DB		LCS	HE	IIA	MP	MG	BPD	CHE	LLIE	Prop
YF	Img1	4.294	3.103	4.294	4.697	4.644	4.196	4.425	4.114	4.549
	Img2	4.710	3.421	4.710	4.884	4.888	4.532	4.811	4.188	4.692
	Img3	4.697	3.416	4.697	4.290	4.072	4.583	4.845	3.734	4.911
	Img4	5.213	3.848	5.213	5.260	5.251	4.981	5.340	4.724	5.083
	Img5	3.801	2.761	3.801	3.416	3.431	3.692	3.993	3.822	3.920
	Img6	3.867	2.746	3.867	3.641	3.383	3.793	3.995	3.709	3.890

Table 4.22. Entropic measures values achieved by using different enhancement methods on the ORL face dataset.

DB		LCS	HE	IIA	MP	MG	BPD	CHE	LLIE	Prop
ORL	Img1	7.291	5.977	7.291	7.747	7.719	7.160	7.783	7.331	7.456
	Img2	7.429	5.981	7.429	7.724	7.724	7.329	7.874	7.430	7.242
	Img3	7.342	5.972	7.342	7.713	7.705	7.240	7.859	7.532	7.268
	Img4	7.368	5.954	7.368	7.578	7.638	7.248	7.818	7.646	7.706
	Img5	7.379	5.964	7.379	7.404	7.397	7.262	7.888	7.618	7.930
	Img6	7.432	5.982	7.432	7.631	7.610	7.290	7.869	7.704	7.199

Table 4.23. AMBE values achieved by using different enhancement methods on AR face datasets.

DB		LCS	HE	IIA	MP	MG	BPD	CHE	LLIE	Prop
AR	Img1	0.066	0.055	0.066	0.058	0.049	0.023	0.065	0.183	0.060
	Img2	0.078	0.061	0.078	0.133	0.137	0.012	0.098	0.182	0.112
	Img3	0.049	0.037	0.049	0.032	0.017	0.014	0.033	0.186	0.037
	Img4	0.053	0.172	0.053	0.165	0.187	0.005	0.055	0.112	0.139
	Img5	0.048	0.153	0.048	0.117	0.121	0.001	0.040	0.103	0.131
	Img6	0.051	0.269	0.051	0.114	0.114	0.003	0.086	0.074	0.114

4.2.4.5 Performance based on the AMBE value

Tables 4.23 – 4.25, display the results achieved when comparing the different enhancement methods based on the number of edges on the different standard benchmark face datasets were compared. Here, a lower AMBE value was required for an effective image enhancement technique.

Table 4.24. AMBE values achieved by using different enhancement methods on Yale face datasets.

DB		LCS	HE	IIA	MP	MG	BPD	CHE	LLIE	Prop
YF	Img1	0.026	0.146	0.026	0.103	0.122	0.008	0.003	0.002	0.115
	Img2	0.009	0.134	0.009	0.135	0.128	0.006	0.015	0.099	0.112
	Img3	0.009	0.060	0.009	0.081	0.334	0.001	0.034	0.002	0.083
	Img4	0.005	0.112	0.005	0.113	0.132	0.023	0.024	0.116	0.111
	Img5	0.007	0.042	0.007	0.629	0.549	0.005	0.044	0.101	0.055
	Img6	0.010	0.071	0.010	0.072	0.446	0.028	0.032	0.075	0.063

Table 4.25. AMBE values achieved by using different enhancement methods on ORL face datasets.

DB		LCS	HE	IIA	MP	MG	BPD	CHE	LLIE	Prop
ORL	Img1	0.016	0.055	0.016	0.065	0.080	0.003	0.014	0.138	0.066
	Img2	0.022	0.050	0.022	0.133	0.134	0.003	0.001	0.168	0.118
	Img3	0.004	0.008	0.004	0.093	0.096	0.006	0.037	0.083	0.083
	Img4	0.013	0.040	0.013	0.086	0.066	0.001	0.036	0.095	0.028
	Img5	0.006	0.012	0.006	0.081	0.076	0.001	0.036	0.131	0.079
	Img6	0.019	0.061	0.019	0.046	0.049	0.016	0.051	0.148	0.044

4.2.4.6 Performance based on the fitness value

Tables 4.26 – 4.28 display the results achieved when the different enhancement methods based on the number of edges on the different standard benchmark face dataset were

compared. A higher fitness value was required for an effective image enhancement technique.

Table 4.26. Fitness value achieved by using different enhancement methods on the AR face datasets.

DB		LCS	HE	IIA	MP	MG	BPD	CHE	LLIE	Prop
AR	Img1	1.192	1.022	1.192	1.352	1.333	1.196	1.312	1.179	1.403
	Img2	1.142	1.017	1.142	1.301	1.299	1.194	1.267	1.115	1.315
	Img3	1.234	1.091	1.234	1.385	1.387	1.236	1.373	1.183	1.482
	Img4	1.257	1.009	1.257	1.370	1.365	1.308	1.326	1.135	1.497
	Img5	1.230	0.941	1.230	1.333	1.333	1.284	1.329	1.146	1.553
	Img6	1.154	0.949	1.154	1.278	1.278	1.136	1.223	1.044	1.361

Table 4.27. Fitness value achieved by using different enhancement methods on the YF face datasets.

DB		LCS	HE	IIA	MP	MG	BPD	CHE	LLIE	Prop
YF	Img1	0.814	0.687	0.814	0.857	0.848	0.846	0.816	0.795	0.934
	Img2	0.936	0.797	0.936	0.927	0.932	0.902	0.864	0.713	1.010
	Img3	0.950	0.701	0.950	0.832	0.715	0.936	0.855	0.747	0.963
	Img4	1.038	0.886	1.038	0.965	0.963	0.896	0.927	0.770	1.068
	Img5	0.838	0.599	0.838	0.575	0.574	0.775	0.729	0.656	0.701
	Img6	0.823	0.594	0.823	0.719	0.596	0.829	0.742	0.660	0.736

Tables 4.17 – 4.25, display the PSNR value, entropic measure value, and AMBE value of each enhancement method son each of the face datasets. The PSNR value achieved with the proposed image enhancement technique gave the lowest value across all face images within the different face datasets. This means that a lower PSNR value produces a more enhanced image.

Table 4.28. Fitness value achieved by using different enhancement methods on the ORL face datasets.

DB		LCS	HE	IIA	MP	MG	BPD	CHE	LLIE	Prop
ORL	Img1	1.226	1.036	1.226	1.413	1.418	1.235	1.331	1.181	1.602
	Img2	1.410	1.231	1.410	1.774	1.774	1.421	1.670	1.430	1.988
	Img3	1.386	1.234	1.386	1.624	1.623	1.390	1.566	1.428	1.850
	Img4	1.337	1.146	1.337	1.542	1.506	1.342	1.529	1.349	1.844
	Img5	1.332	1.352	1.332	1.626	1.630	1.360	1.482	1.361	1.779
	Img6	1.279	1.066	1.279	1.388	1.399	1.285	1.358	1.292	1.649

The entropic measure value with the proposed technique produces a higher value in most of the face images than other image enhancement techniques across the different face datasets. This shows that more information to make a face image unique has been added to the enhanced image. Also, the AMBE value which represents the amount of mean brightness that has been preserved has been compared amongst all image enhancement techniques. The proposed image enhancement technique compared with other methods shows that they all performed well in preserving the absolute mean brightness error value.

Tables 4.26 – 4.28 display the fitness function value obtained from the different image enhancement methods on the different face datasets. The fitness value is an important metric that determines the effectiveness of the various techniques, and a higher fitness value is desired. The proposed algorithm produces the highest fitness value across most face images on the different face datasets when compared to other image enhancement techniques. This is followed by the MP enhancement method across all images but one, where MG had a higher fitness value. As regards image 3 on the ORL face dataset, methods such as CHE and BPD also performed satisfactorily. Lastly, the HE enhancement method yielded the lowest fitness value across all images. To confirm the performance of the proposed algorithm further, qualitative comparison of the enhanced images produced by the different image enhancement methods is presented as shown in Figures 4.11 – 4.13 with each representing different real-life facial conditions.

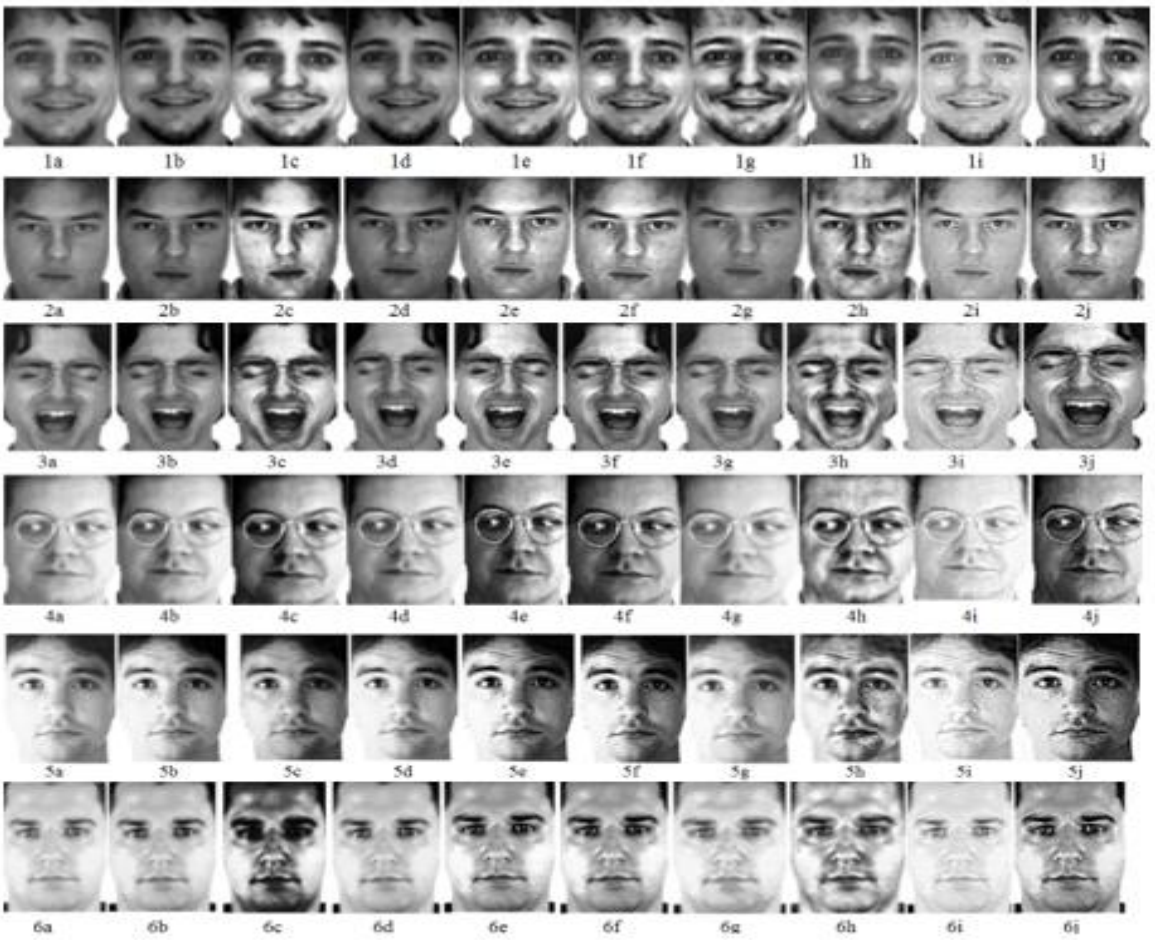


Figure 4.11. Qualitative comparison of the different image enhancement algorithms on the AR face dataset where Figures 1 – 6 represent images of different subjects respectively; and a – j denotes the methods labelled as (a) original, (b) LCS (c) HE (d) IIA (e) MP (f) MG (g) BPD (h) CHE (i) LLIE (j) Proposed**.**

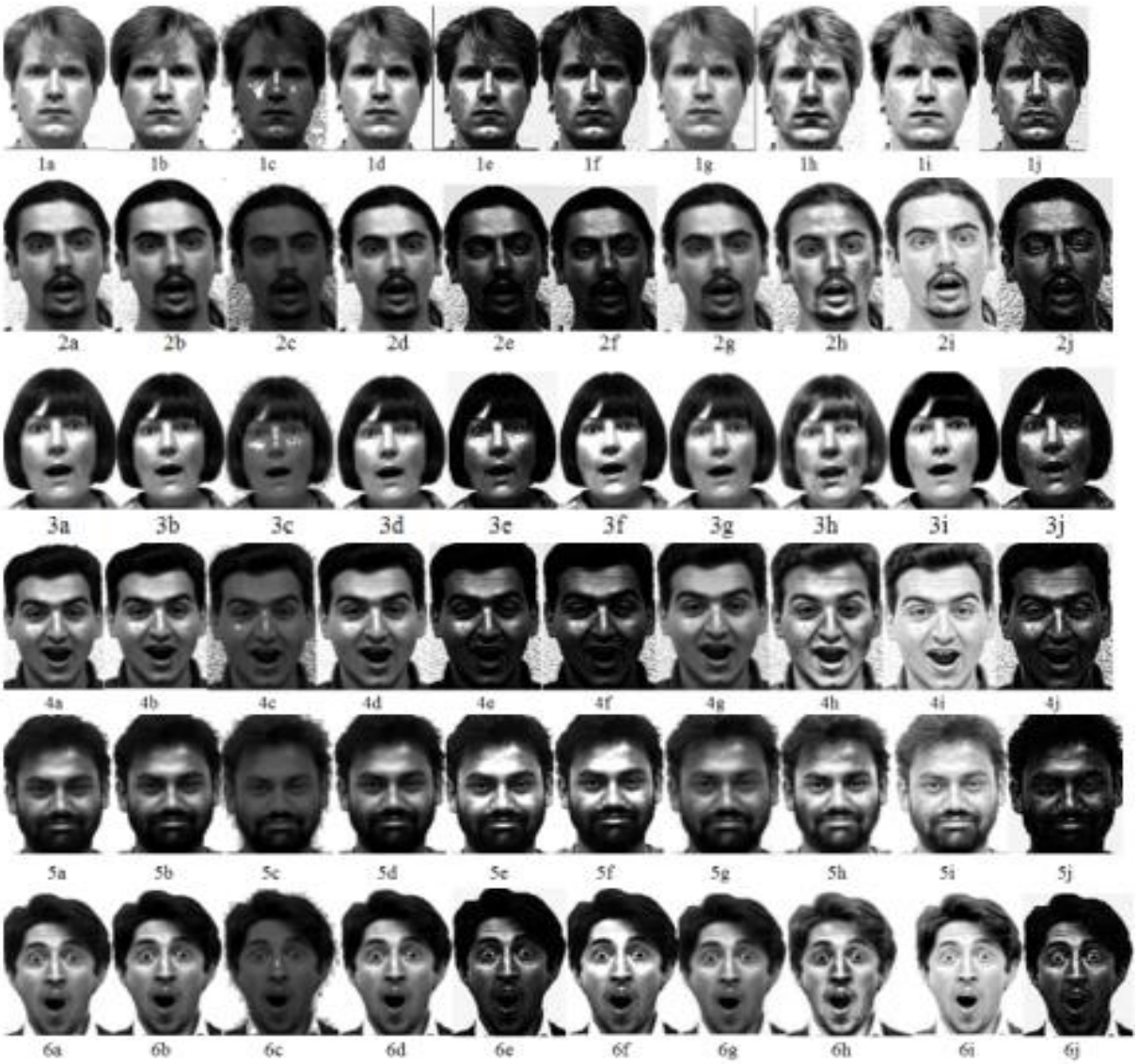


Figure 4.12. Qualitative comparison of the different image enhancement algorithms on the Yale face dataset where Figures 1 – 6 represent images of different subjects respectively; and a – j denotes the methods labelled as (a) original, (b) LCS (c) HE (d) IIA (e) MP (f) MG (g) BPD (h) CHE (i) LLIE (j) Proposed.

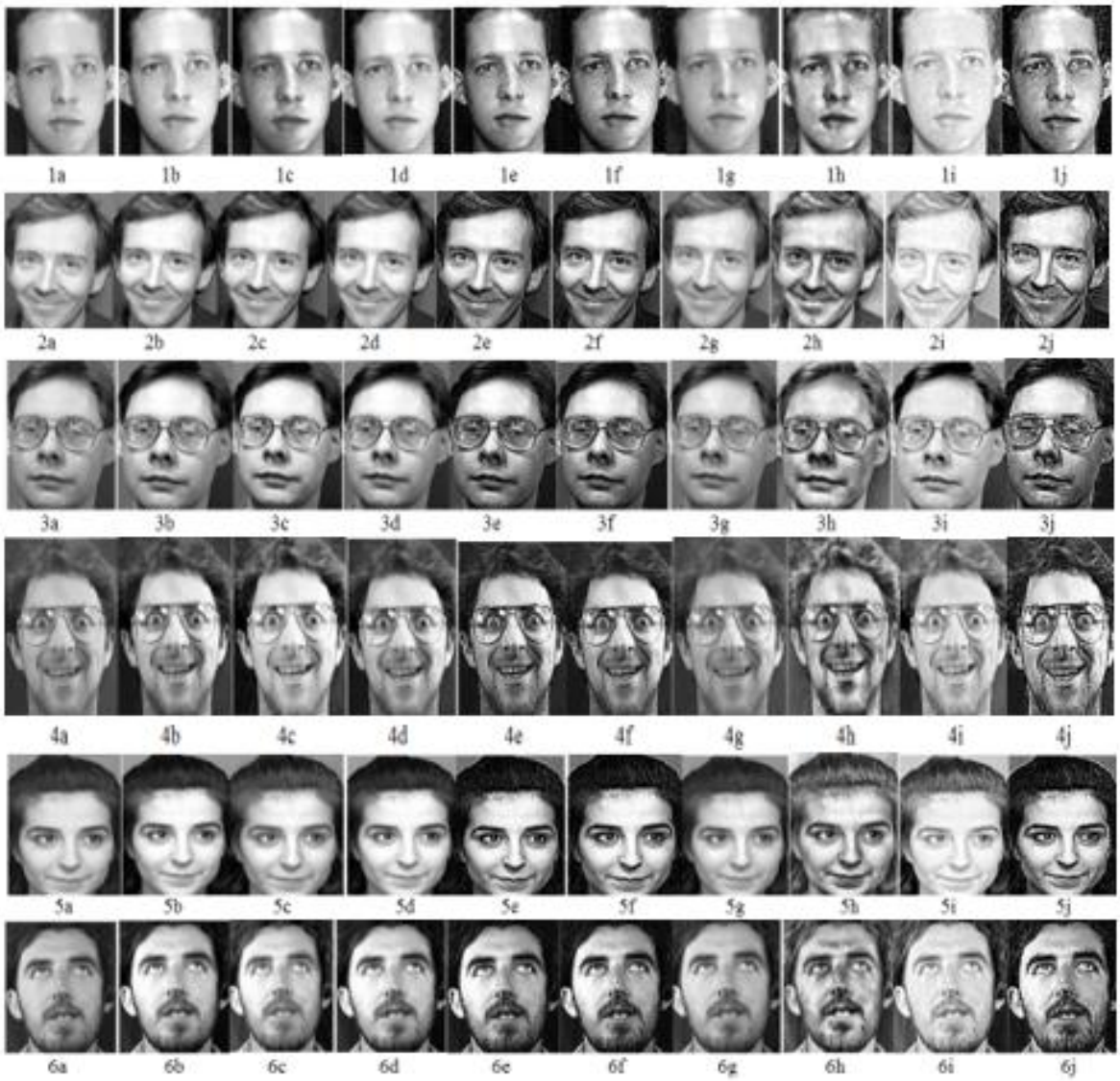


Figure 4.13. Qualitative comparison of the different image enhancement algorithms on the ORL face dataset where Figures 1 – 6 represent images of different subjects respectively; and a – j denotes the methods labelled as (a) original, (b) LCS (c) HE (d) IIA (e) MP (f) MG (g) BPD (h) CHE (i) LLIE (j) Proposed.

Figures 4.11 – 4.13 display the effectiveness of the proposed image enhancement technique on different standard face datasets. The enhanced images produced by the proposed algorithm show much positive difference when compared to the original and other enhanced images in quality. Therefore, compared to those of other enhancement

techniques, images affected by the different facial conditions that depict the real-life scenarios have been effectively enhanced.

Furthermore, the proposed image enhancement technique produced good enhanced images of facial images with different pose and expressions. Generally, unlike the images produced by other image enhancement techniques, the proposed image enhancement technique has produced more explicit facial images. This is due to the different parameters that have been taken into consideration when designing the image enhancement technique. Parameters such as the number of edges and the number of the pixel value in the foreground will undoubtedly add more features to the image, thus, producing a more enhanced image. Finally, to confirm the effectiveness of an image enhancement method, the performance should be based on both qualitative and quantitative measures.

4.3 IMPROVING FACE RECOGNITION SYSTEMS USING A NEW IMAGE ENHANCEMENT TECHNIQUE, HYBRID FEATURES, AND THE CONVOLUTIONAL NEURAL NETWORK

In this section, the results obtained from the approach in implementing the FRS as discussed in Section 3.3 are discussed. Firstly, the results of an experiment to determine the selection of the proposed hybrid features are presented.

Results from the experiments carried out on the different enhanced face datasets are presented, to confirm the effect of the enhancement method on face images using the state-of-the-art CNN architecture model and the selected hybrid features. Furthermore, experiments are carried out based on the different types of constraint present in the face datasets. Also presented are results from experiments carried out on the LFW face dataset regardless of the constraints to further confirm the effect the enhancement method on face images by using an 18-Layer Residual Network (ResNet) state-of-the-art CNN architecture with the selected hybrid features.

Finally, a qualitative comparison of the different image enhancement algorithms on the LFW face dataset is presented. It is important to note that all the enhanced face datasets produced by the different image enhancement methods have been experimented on by using the same selected hybrid features and the CNN architectures.

4.3.1 Choice of proposed hybrid features

This section presents the results from experiments done to confirm the effectiveness of the selection of new hybrid features on the different image enhancement methods by using the CNN classifier. In achieving this, all the data samples in the individual face dataset, regardless of the facial constraints are used. The performance of the proposed enhancement method and the different selected features on the AR and Yale face dataset are shown in Tables 4.29 and 4.30 respectively.

Table 4.29. Performance of the proposed hybrid feature on the AR dataset based on different image enhancement methods.

	FEATURE METHODS					
Method	**Phog**	**EHD**	**LBP**	**Phog+EHD**	**Phog+LBP**	**Phog+LBP+EHD**
UnEh	83.1	62.6	45.5	85.6	87.4	89.4
LCS	88.7	65.1	50.1	89.7	90.5	91.8
HE[182]	87.5	65.4	50.5	90.5	91.7	92.4
AHE[183]	88.4	67.6	52.2	92.1	93.3	94.8
IIA	87.9	66.5	45.8	89.1	90.8	92.6
MP[168]	89.4	68.3	57.5	93.3	94.1	94.9
MG[168]	89.7	67.5	56.5	92.3	93.8	94.5
Prop.	92.3	70.6	60.3	95.4	97.6	98.46

UnEh – Unenhanced; LCS – Linear contrast stretching; HE – Histogram equalization; AHE – Adaptive histogram equalization; IIA – Image intensity adjustment; MP – Munteau's method with PSO; MG – Munteau's method with genetic algorithm; Prop – Proposed method.

Various individual and hybrid features were selected in this work. The selected features were used in this thesis after extensive experiments with other methods which performed poorly on the enhanced face datasets. The features selected included PHOG, LBP, and

EHD. In Table 4.29, when the individual features were used, the Phog method performed best on all types of enhanced methods. With the unenhanced image, an average recognition rate of 83.1% was achieved, while a recognition rate of 92.3% was achieved on the proposed enhancement method. The LBP feature extraction method was the least successful with an average recognition rate of 45.5% on the unenhanced image; 50.5 % with the HE method, 57.5 % with the MP method and 60.3 % with the proposed enhancement method. When the individual methods, i.e., Phog and EHD, were combined, there was a slight increase in the recognition performance as compared to that of the feature extraction methods used individually.

On the unenhanced images, an average recognition rate of 85.6% was achieved; while with the LCS, HE MG and MP methods, average recognition rates of 89.7%, 90.5%, 92.3%, and 93.3% were achieved respectively. The recognition performance of this selection of new hybrid features with the proposed enhanced method was 95.4%. Also, the combination of the Phog and LBP feature extraction methods performed even better across all enhanced and the unenhanced image dataset. This led to the proposed selection of hybrid features where all the methods were combined to extract essential features from the enhanced images effectively.

The proposed method, which was a combination of the Phog, EHD, and LBP, interestingly performed better as regards the recognition rate. On the unenhanced image, an average recognition rate of 89.4% was achieved, while recognition rate of 91.8%, 94.8%, 92.6%, and 94.9% were achieved with the LCS, AHE, IIA and MG enhancement methods. Finally, the new hybrid features performed best with the proposed enhancement method with an average recognition rate of 98.46%.

In Table 4.30 shown, a similar experiment was carried out on the Yale face dataset. Unlike on the AR face dataset, the LBP outperformed the EHD feature extraction method across all types of enhanced dataset type.

Table 4.30. Performance of the proposed hybrid feature on the Yale dataset based on different image enhancement methods.

	FEATURE METHODS					
Method	**Phog**	**EHD**	**LBP**	**Phog+EHD**	**Phog+LBP**	**Phog+LBP+EHD**
UnEh	79.4	56.1	71.3	83.5	80.1	86.9
LCS	81.3	59.8	80.6	86.1	82.4	87.5
HE	82.8	60.1	76.4	87.7	83.6	86.4
AHE	83.2	60.8	80.6	88.3	84.2	87.5
IIA	81.5	59.5	80	86.5	81.3	88.2
MP	85.7	63.4	76.9	90.3	85.9	90.8
MG	85.2	63.5	75.8	90.8	86.4	92.3
Prop.	87.5	65.5	80.4	92.7	88.6	94.6

For instance, on the LCS enhanced dataset, average recognition rate of 80.6% was achieved with the LBP method as compared to 59.8% when using the EHD technique. Also, on the MG enhanced dataset, an average recognition rate of 75.8% was achieved with the LBP method, while an average recognition rate of 63.5% was achieved when using the EHD technique. This was due to the change in the dataset, where the Yale face dataset had more background in the images. Also, the combination of Phog and EHD surpassed the rates of Phog and LBP used individually across all types of image. Therefore there is some inconsistency in the recognition rates of these methods. However, the feature method performed best on all enhanced and unenhanced datasets.

On the unenhanced image, an average recognition rate of 86.9% was achieved, 87.5% was achieved with the AHE enhancement method, 92.3% with MG enhancement method and finally, 94.6% with the proposed enhancement method. The consistency in performance of the selected hybrid features proved its ability to extract features from enhanced images effectively to increase recognition performance when using the CNN classification method.

4.3.2 Recognition based on constraints

Here, the performance of the proposed approach to different constraints present in the face dataset when using the new selected features and the selected CNN classification method is confirmed. To achieve this, the different constraints present in each dataset were grouped,

i.e., lighting condition, expression, and occlusion from the AR face dataset. Lighting conditions and expressions were also grouped from the Yale face dataset. Figures 4.14 – 4.16 and Figures 4.17 – 4.18 further detail the results from the AR and Yale face datasets based on different constraints.

4.3.2.1 Recognition performance based on the AR face dataset

Figure 4.14 shows the performance of the different enhancement methods and the selected hybrid features on the CNN classifier, taking into consideration the issue of lighting conditions. Compared to the unenhanced dataset, performance improved when the different enhanced datasets were used. On the unenhanced dataset, an average recognition rate of 89.8% was achieved while on the LCS, HE, MP and MG average recognition rates of 91.8%, 92.4%, 94.1%, and 93.3% were achieved respectively. On the proposed enhanced method, the highest recognition rate was achieved with 95.5%.

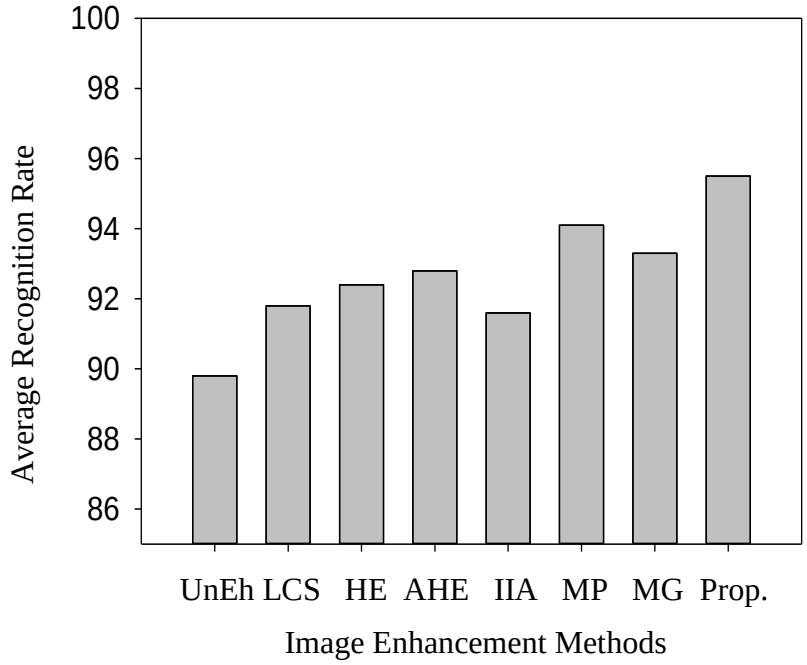


Figure 4.14. Average recognition performance based on lighting.

Similarly, for the issue of expression, as shown in Figure 4.15, the recognition rate was the lowest with the unenhanced dataset with an accuracy of 88.4%, while there was an increase in performance with the different enhancement methods. For instance, average recognition

rates of 91.8%, 92.3%, 93.5%, 90.8% were achieved on the LCS, HE, AHE and MP enhanced datasets respectively. Furthermore, performance on the enhanced dataset when using the proposed approach produced the highest recognition rate with an accuracy of 93.2%.

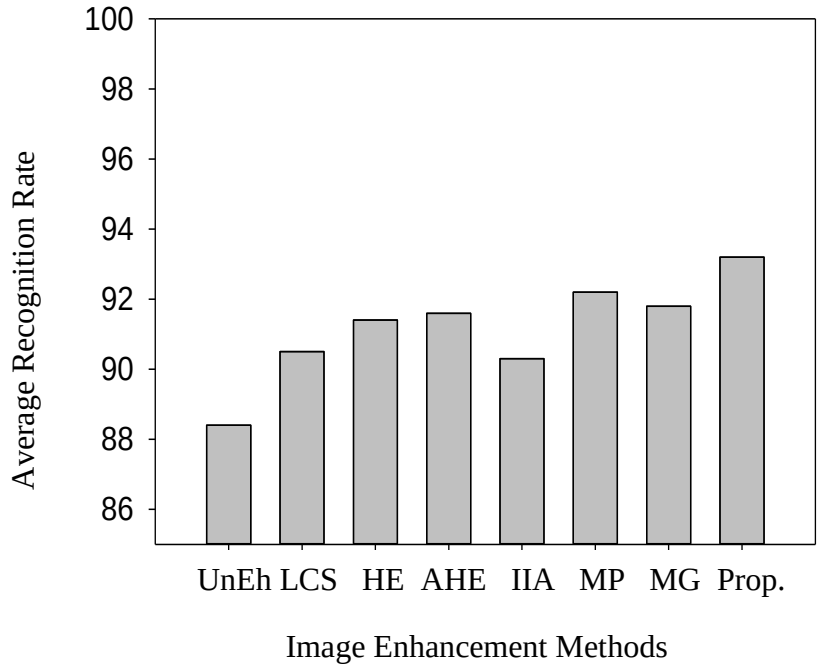


Figure 4.15. Average recognition performance based on expression.

As regard the issue of occlusion, results were based on occlusion of both upper and lower face, i.e., either sunglasses or scarf, as shown in Figure 4.16. Results on the upper face occlusion show that there was a reduction in the performance of the FRS. However, the highest recognition accuracy was achieved on the enhanced face dataset when using the proposed approach. A performance recognition accuracy of 85.0% was achieved as compared to 74.3% on the unenhanced face dataset.

Performance accuracy on the other enhanced methods equally outperformed the unenhanced face dataset, however not reaching the level of performance of the enhanced face dataset using the proposed approach. Likewise, with lower face occlusion, the recognition performance decreased even more. An average recognition rate of 70.1% was achieved on the unenhanced face dataset. The highest recognition rate achieved was on the

enhanced face dataset with the proposed approach with 80.66%, outperforming other enhanced face datasets. This clearly shows the effectiveness of the proposed approach; however, the issue of lower face occlusion reduces the performance of the proposed FR approach.

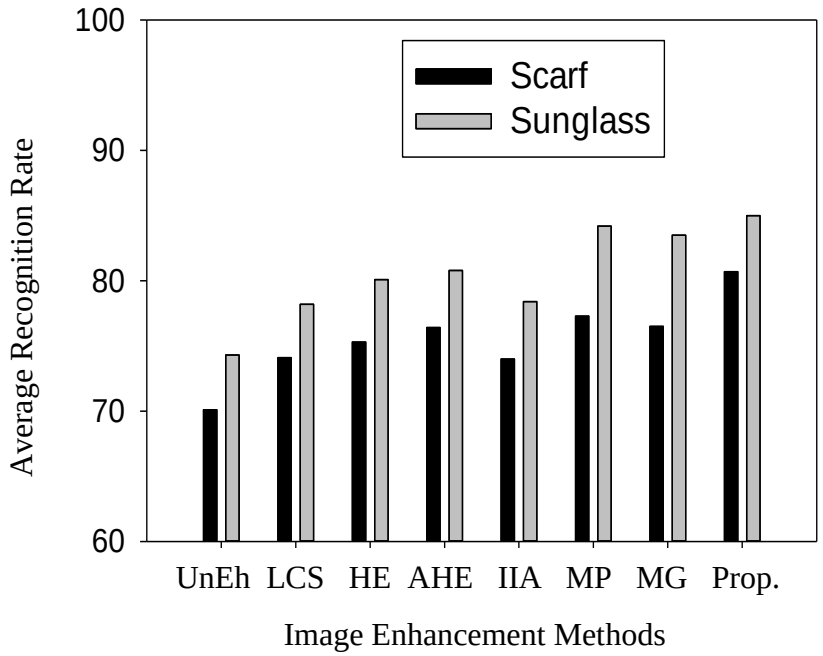


Figure 4.16. Average recognition performance based on occlusion.

4.3.2.2 Recognition performance based on Yale face dataset

Figure 4.17 shows results from the FRS approach on the different enhanced images based on lighting conditions in the Yale face dataset. An average recognition rate of 82.7% was achieved on the unenhanced face dataset, while 91.6% was achieved on the enhanced face dataset when using the proposed approach. This indicates that there is an increase in recognition accuracy when using the proposed approach. Also, other enhanced methods achieved better results as compared to the unenhanced face dataset, for example, average recognition rates of 86.4%, 87.6%, 88.4%, 86% were achieved on the LCS, HE, AHE and IIA enhanced datasets respectively. None performed better than the proposed approach.

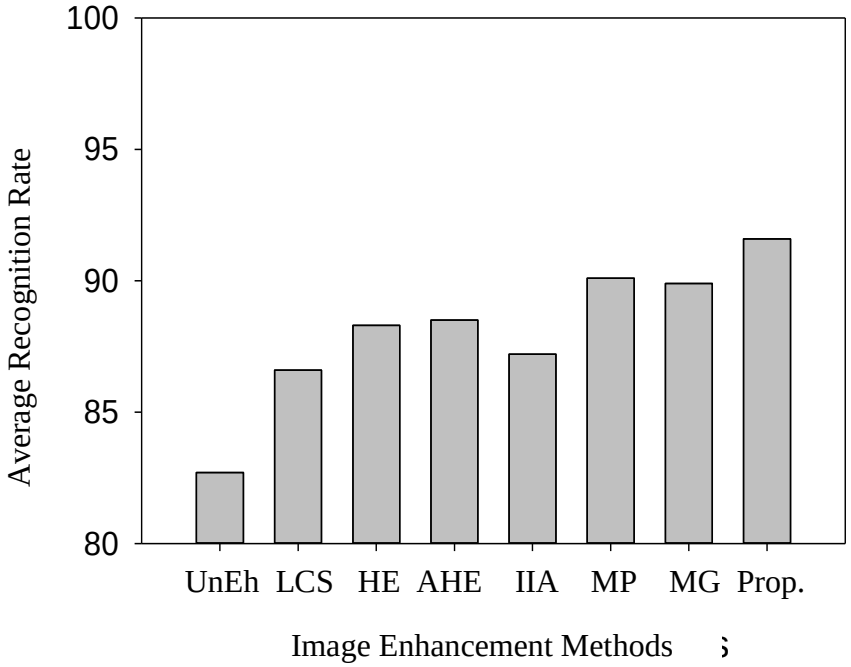


Figure 4.17. Average recognition performance based on lighting.

As regards the issue of expression on Yale face dataset, Figure 4.18, shows the experimental results. On the unenhanced face dataset, there was a slight increase as compared to the issue of lighting with an average recognition rate of 84.4%. Also, an average recognition rates of 86.8%, 88.3%, 88.5%, and 90.1% were achieved respectively on the face datasets of the LCS, HE, AHE and MP enhancement methods. Finally, an average recognition rate of 93.8% was achieved on the face datasets when using the proposed method.

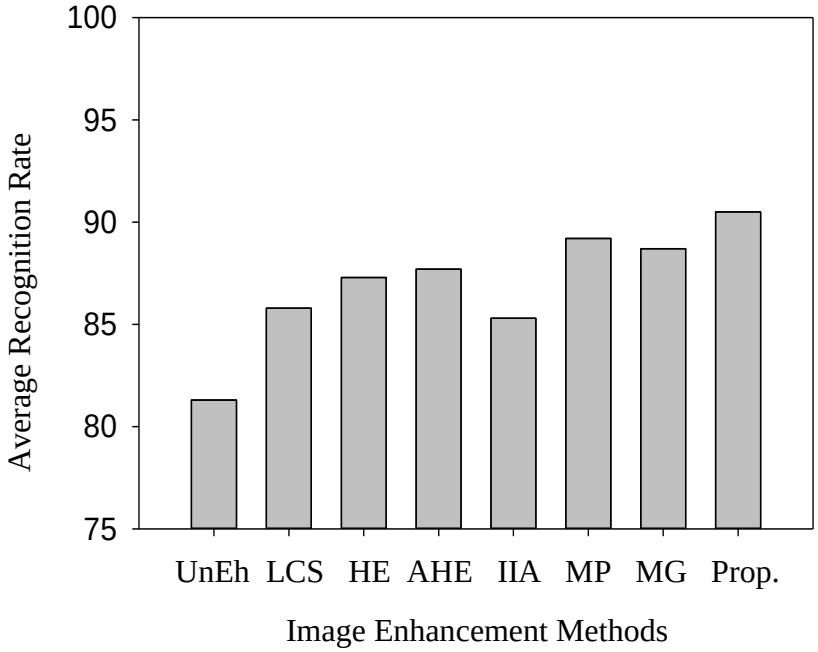


Figure 4.18. Average recognition performance based on expression.

4.3.3 Recognition performance on the LFW dataset

In this result section, the performance of the FRS approach on the LFW face dataset when using the new enhancement method; proposed selected hybrid features, and an 18 -Layer Residual Network (ResNet) state-of-the-art CNN architecture described in [184] was confirmed. ResNet CNN architecture was considered because of its high-level performance in the recent ImageNet Large Scale Visual Recognition Competition (ILSVRC).

The LFW face dataset was considered for this experiment due to the various facial conditions present in the dataset. Like in Section 4.2, all the facial conditions of selected subjects in the face dataset were used. Subjects were selected with four or more facial images, which varied in terms of the different facial constraints such as occlusion, expression, pose and lighting conditions. In this work, a total of 2110 LFW facial images have been used for the results reported in Table 4.31. These consists of subjects that have five different images with different facial conditions. Hence, 422 subjects have been selected with five different images per subject. Figure 4.19. displays the enhanced face

images of four different subjects when using the proposed and different enhancement methods.

Table 4.31. Recognition performance of the 18-Layer ResNet CNN architecture on the LFW dataset using different features and image enhancement methods.

	FEATURE METHODS					
Method	**Phog**	**EHD**	**LBP**	**Phog+EHD**	**Phog+LBP**	**Phog+LBP+EHD**
UnEh	82.4	65.2	75.6	85.7	82.3	88.9
LCS	84.8	68.8	84.2	88.1	84.6	89.2
HE	85.4	71.3	79.8	89.5	85.8	88.5
AHE	86.2	72.5	84.4	90.3	86.2	89.1
IIA	84.5	68.5	83.7	89.1	83.4	90.5
MP	88.7	75.2	80.2	92.3	87.7	92.5
MG	88.2	75.4	79.6	92.8	88.4	94.3
Prop.	89.8	77.5	84.5	94.5	90.2	96.7

As shown in Table 4.31, the results of the experiments carried out on the LFW face dataset to confirm the effectiveness of the proposed approach when using an 18-layer ResNet state-of-the-art CNN architecture.

It is evident that the proposed enhancement method with the hybrid selection of features presented outperforms other approaches. With the LCS method, an average recognition rate of 65.2% was achieved with the EHD feature, while the proposed enhancement method achieved an average recognition rate of 77.5%. With the MP method, an average recognition rate of 92.3% was achieved with the Phog+EHD features, while with the proposed enhancement method, there was a slight increase in performance with an average recognition rate of 92.5%.

With the proposed enhancement method, with the Phog+EHD features, an average recognition rate of 94.5% was achieved; with the Phog+LBP features, the average recognition rate was 90.2%. While for the hybrid feature presented, i.e., Phog+LBP+EHD an average recognition rate of 96.7% was achieved. This experiment further confirmed the effectiveness of the new enhancement method, and the right selection of proposed features to be extracted from the enhanced face dataset.

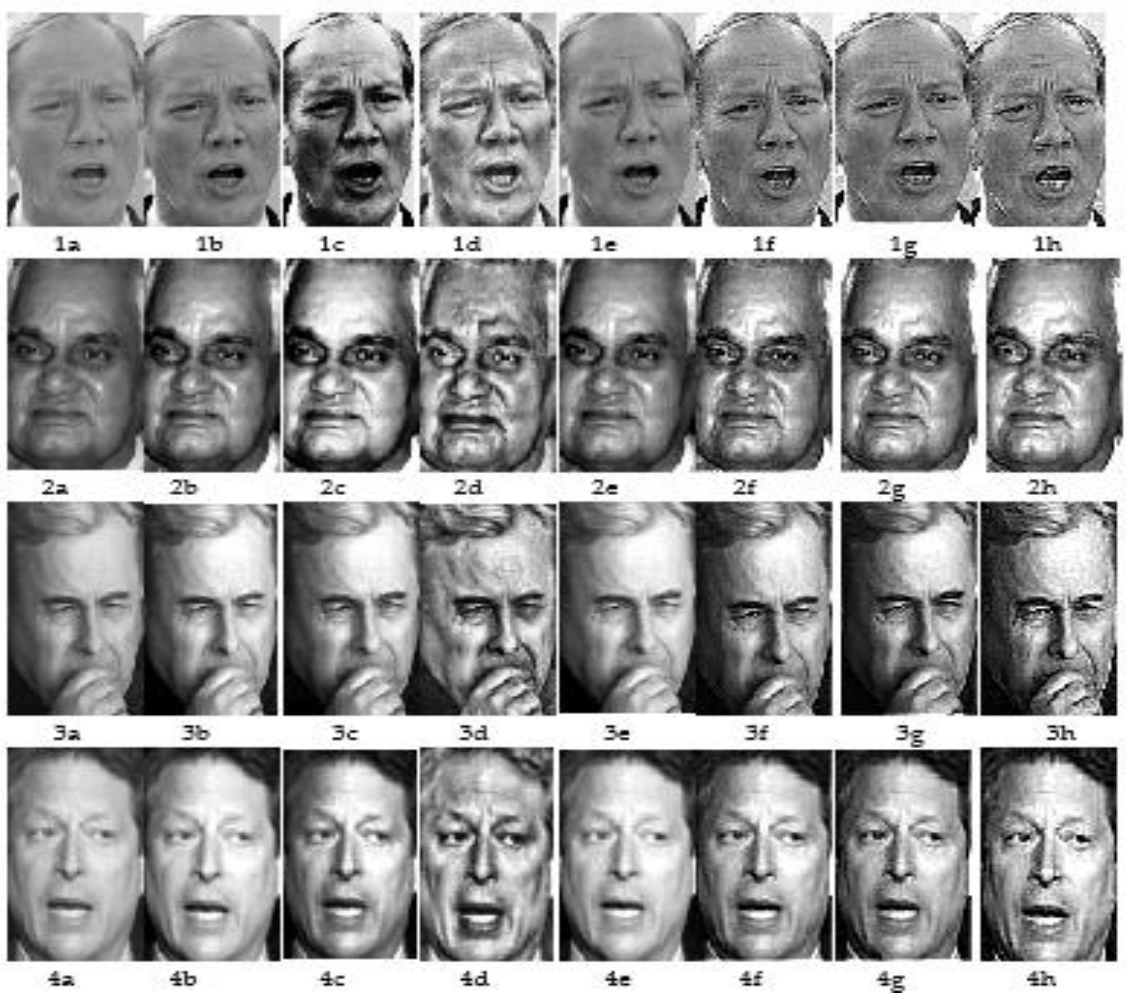


Figure 4.19. Qualitative comparison of the different image enhancement algorithms on the LFW face dataset where Figures 1 – 4 represent images of different subjects respectively; and a – h denote the methods labelled as (a) original, (b) LCS (c) HE (d) AHE (e) IIA (f) MP (g) MG (h) Proposed.

4.4 CHAPTER SUMMARY

In this chapter, the results of the different experiments carried out in this thesis are presented. First, in Section 4.1, the results from the comparative analysis of the different face recognition techniques are presented. This involved the combination of different feature and classification methods. Also, the comparative analysis was carried out on

different standard benchmark datasets, while also considering various facial conditions that depict unconstrained environments.

In Section 4.2, the results obtained from experiments in designing the proposed image enhancement techniques are discussed. This involved the choice of CSO parameters, the confirmation of the performance of different metaheuristic algorithms, the effectiveness of the new evaluation function, and also the performance of the proposed image enhancement technique.

Finally, in Section 4.3., the results obtained from experiments carried out towards the design of the proposed FRS are presented. These includes results to determine the selection of features, the confirmation of the effect of the proposed image enhancement method using the state-of-the-art CNN architecture model and the selected hybrid feature methods.

Lastly, results from an experiment carried out on the LFW face dataset are presented using 18-Layer Residual Network (ResNet) state-of-the-art CNN architecture with the selected hybrid features. The various experiments carried out have shown the effectiveness of the proposed method.

CHAPTER 5 CONCLUSION AND FUTURE WORK

In conclusion, it is necessary to highlight the objectives of this research. The purpose of this research is to achieve the following primary objectives:

- To carry out a comparative analysis of various face recognition techniques to enable us to confirm the performance of these methods in selected unconstrained scenarios.
- To propose a novel image enhancement technique as a pre-processing approach for FRS.
- To propose the right features to be extracted from the enhanced facial image such that their recognition performance is enhanced.
- To design a FRS where an effective image enhancement technique is put in place to increase recognition performance.

In this thesis, the first objective was to carry out a comparative analysis of different FRS in unconstrained environments. This study used a combination of different features and classification methods for this purpose. This was done to confirm the performance of the different face recognition techniques in unconstrained environments. Experiments were carried out in various unconstrained environments, using different standard and benchmark face datasets.

The experiments entailed using both the whole face dataset, regardless of the constraints and considering different unconstrained conditions. There are many face recognition techniques that involve the combination of different features and classification methods. However, it is necessary to confirm the performance of these methods especially in

unconstrained environments. In this regard, the recognition performance rate was used to compare these methods because it describes how well a FBRS accurately recognizes a subject. The results obtained from the experiments showed that using different feature methods with the convolutional neural network classification yielded better results than other classification methods.

The second objective of this thesis was to develop an effective metaheuristic-based image enhancement technique as a pre-processing approach for facial images in unconstrained environments. The proposed face image enhancement method consisted of a transformation function, evaluation function and the metaheuristic algorithm. The transformation function adapted consisted of four different parameters, i.e. *a, b, c* and *k*. Parameter *a* introduced a brightening bias in the output image based on the last term *m(i,j)* in (3.1). It also enabled further control over the amount of smoothening effect required in the output image. Parameter *b* ensured that a zero-standard deviation value in the local neighbourhood pixels did not have a huge whitening effect on the final output image. Therefore, by its introduction, the denominator component in (3.1) typically remained nonzero. Parameter *c* allowed for only a fraction of the mean *m(i,j)* to be subtracted from the original pixels of the input image. That controlled the degree of darkening introduced in the output image. Parameter *k* was introduced to create a fair balance between pixels existing in the mid-range boundaries of the grey scale. Essentially these pixels were controlled from being made either too dark or too white. The following values of the individual parameters were found to be highly effective based on an extensive empirical parameter-tuning exercise conducted in this thesis: $2 \le a \le 2.5$; $0.3 \le b \le 0.5$; $0 \le c \le 3$ and $3 \le k \le 4$.

In addition, this thesis proposes a new evaluation function for improving the performance of metaheuristic-based image enhancement approaches. The new evaluation function used in combination with metaheuristic-based optimization algorithms makes it possible to automatically select the best-enhanced face image based on a linear combination of different key quantitative measures. Also, this was made possible through the inclusion of an intelligent and defined scale mechanism in the fitness function for evaluating the

enhanced facial images such that extreme values represent either dark or too bright images. Experiments have shown that the proposed evaluation function algorithms outperform state-of-the-art methods that include other standard and metaheuristic evaluation functions.

To further confirm the effectiveness of the proposed approach, an extensive quantitative and qualitative comparison with both metaheuristic and state-of-the-art image enhancement techniques has been carried out. The proposed image enhancement technique has been evaluated by using different facial images from different standard benchmark face datasets that represent real-life scenarios. The different computed performance metrics showed the proposed algorithm's superiority over other metaheuristic and state-of-the-art enhancement techniques.

The final objective of this thesis was to come up with a FRS model in unconstrained environments that entailed a new enhancement method at the pre-processing stage and a selection of new hybrid features capable of effectively extracting features from the enhanced image. A major task in this respect was to identify the features that were effective from within an enhanced image. As a result, extensive experiments were conducted using different feature methods based on the convolutional neural network classification method. The newly introduced hybrid feature methods that consisted of the Phog, EHD and LBP features proved to be effective in extensive experiments. This thesis proved that using an effective image enhancement technique in the pre-processing stage of a FRS increased its performance as compared to using unenhanced facial images particularly using the convolutional neural network as the classification method. The results from experiments carried out showed that there was a significant increase in the recognition performance of the proposed enhancement method and the newly selected hybrid features as compared to other approaches with two state-of-the-art CNN classification methods. The selection of the presented new hybrid features from the enhanced face images were also shown to have an impact by enhancing recognition performance.

The thesis can be extended in different directions. In this thesis, the recognition rate used to evaluate performance has been considered. The study can be extended to consider other metrics such as time to build the system. Most especially, the proposed FRS and other metaheuristic-based approaches take longer time as compared to traditional techniques. This is because metaheuristic-based approaches require the need to search very large space for candidate solutions. Hence, the study can be extended to reduce the time taken using metaheuristic-based FRS approach. The issue of lower face occlusion has been observed to significantly reduce the performance of FRS models as compared to other facial conditions. The study can be extended to investigate approaches to increase the recognition rate in such conditions. The research can also be extended to study the effect of enhancement on other facial conditions in unconstrained environments such as plastic surgery and aging.

Deep neural networks are effective techniques for computer vision tasks and most recently for FR tasks. Notwithstanding, the attempts by previous researchers to solve facial constraints during FR, there still is the possibility of improving individual layers of the CNN that can make the network perform optimally. The pooling layer of the CNN architecture is a very important layer because it can reduce the size of the feature map that portrays the filtered image in the convolution layer. Also, it enhances the performance of the CNN by transforming invariance such as expansion, rotation and translation. Hence, it is one component of the CNN architecture that can be used to address the major issues of FR to achieve optimal performance. The pooling layer consists of both the max and average pooling, therefore, the performance of the CNN can largely depend on the type of pooling method used; which is also a function of the input image. Even though the max and average pooling methods perform well on certain datasets, it is unsure which pooling method will outperform the other with different FR problems. Also, other approaches can be introduced into the pooling layer so as to select features that can perform effectively.

Access to the facial dataset is somewhat difficult as a lot of process is involved, thus making research in the area slow. The face datasets should be developed such that a collection of faces can include most of the facial constraints. By doing this, face

recognition techniques can be evaluated with just a single face dataset that depicts a typical real-life scenario.

REFERENCES

[1] M. Merone, P. Soda, M. Sansone, and C. Sansone,"ECG Databases for Biometric Systems: A Systematic Review", *Expert Systems with Applications*, vol. 67, pp. 189-202, January 2017.

[2] M. L. Ali, J. V. Monaco, C. C. Tappert, and M. Qiu, "Keystroke Biometric Systems for User Authentication", *Journal of Signal Processing Systems*, vol. 86, no. 2-3, pp. 175-190, March 2017.

[3] F. Jan, "Segmentation and Localization Schemes for Non Ideal Iris Biometric Systems", *Signal Processing*, vol. 133, pp. 192-212, April 2017.

[4] M. Lafkih, P. Lacharme, C. Rosenberger, M. Mikram, S. Ghouzali, M. El Haziti, W. Abdul and D. Aboutajdine, "Application of New Alteration Attack on Biometric Authentication Systems", in *Proc. 1st International Conference on Anti Cybercrime IEEE-ICACC*, Riyadh, Saudi Arabia, November 2015, pp. 1-5.

[5] M. Gomez-Barrero, J. Galbally, and J. Fierrez, "Efficient Software Attack to Multimodal Biometric Systems and Its Application to Face and Iris Fusion", *Pattern Recognition Letters*, vol. 36, pp. 243-253, January 2014.

[6] M. Nappi, V. Piuri, T. Tan, and D. Zhang, "Introduction to the Special Section on Biometric Systems and Applications", *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 44, no. 11, pp. 1457-1460, November 2014.

[7] K. Nandakumar and A. K. Jain, "Biometric Template Protection: Bridging the Performance Gap Between Theory and Practice", *IEEE Signal Processing Magazine*, vol. 32, no. 5, pp. 88-100, September 2015.

[8] K. Takahashi and T. Murakami, "A Measure of Information Gained Through Biometric Systems", *Image and Vision Computing*, vol. 32, no. 12, pp. 1194-1203, December 2014.

[9] E. Zureik and K. Hindle, "Governance, Security and Technology: The Case of Biometrics", *Studies in Political Economy*, vol. 73, no. 1, pp. 113-137, March 2004.

[10] Y. Jain and M. Juneja, "A Comparative Analysis of Iris and Palm Print Based Unimodal and Multimodal Biometric Systems", *Innovations in Computer Science and Engineering*: *Springer*, pp. 297-306, June 2017.

[11] M. Elhoseny, E. Essa, A. Elkhateb, A. E. Hassanien, and A. Hamad, "Cascade Multimodal Biometric System Using Fingerprint and Iris Patterns", in *Proc. International Conference on Advanced Intelligent Systems and Informatics*, Cairo, Egypt, August 2017, pp. 590-599.

[12] Z. Akhtar, C. Micheloni, and G. L. Foresti, "Biometric Antispoofing on Mobile Devices", *Mobile Biometrics*, vol. 3, no. 15, pp. 375-406, September 2017.

[13] M. Oloyede, A. Adedoyin, and K. Adewole, "Fingerprint Biometric Authentication for Enhancing Staff Attendance System", *International Journal of Applied Information Systems*, vol. 5, no. 3, pp. 19-24, February 2013.

[14] G. S. Walia, T. Singh, K. Singh, and N. Verma, "Robust Multimodal Biometric System Based on Optimal Score Level Fusion Model", *Expert Systems with Applications*, vol. 116, pp. 364-376, September 2018.

[15] R. Al-Hmouz, W. Pedrycz, K. Daqrouq, and A. Morfeq, "Development of Multimodal Biometric Systems with Three Way and Fuzzy Set Based Decision Mechanisms", *International Journal of Fuzzy Systems*, vol. 20, no. 1, pp. 128-140, January 2018.

[16] L. B. Neto, F. Grijalva, V. Regina, L. Cesar, D. Florencio, and M. Cecilia, "A Kinect Based Wearable Face Recognition System to Aid Visually Impaired Users", *IEEE Transactions on Human Machine Systems*, vol. 47, no. 1, pp. 52-64, February 2017.

[17] W. Xu, Y. Shen, N. Bergmann, and W. Hu, "Sensor Assisted Multi View Face Recognition System on Smart Glass", *IEEE Transactions on Mobile Computing*, vol. 17, no. 1, pp. 197-210, January 2018.

[18] I. Masi, F. Chang, J. Choi, J. Kim and J. Leksut, "Learning Pose Aware Models for Pose Invariant Face Recognition in the Wild", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.41, no.2, pp. 379-393, January 2018.

[19] R. He, X. Wu, Z. Sun, and T. Tan, "Wasserstein CNN: Learning Invariant Features for NIR-VIS Face Recognition", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.41, no.4, pp.251-275, June 2018.

[20] P. Karczmarek, A. Kiersztyn, W. Pedrycz, and M. Dolecki, “An Application of Chain Code Based Local Descriptor and Its Extension to Face Recognition”, *Pattern Recognition*, vol. 65, pp. 26-34, May 2017.

[21] Y. Gao, J. Ma, and A. L. Yuille, “Semi Supervised Sparse Representation Based Classification for Face Recognition with Insufficient Labelled Samples”, *IEEE Transactions on Image Processing*, vol. 26, no. 5, pp. 2545-2560, May 2017.

[22] C. Ding and D. Tao, “Trunk Branch Ensemble Convolutional Neural Networks for Video Based Face Recognition”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 4, pp. 1002-1014, April 2018.

[23] B. F. Klare, B. Klein, E. Taborsky, A. Blanton and J. Cheney, “Pushing the Frontiers of Unconstrained Face Detection and Recognition: Iarpa Janus Benchmark A”, in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, Boston, MA, USA, June 2015, pp. 1931-1939.

[24] M. O. Oloyede and G. P. Hancke, “Unimodal and Multimodal Biometric Sensing Systems: A Review”, *IEEE Access*, vol. 4, pp. 7532-7555, September 2016.

[25] V. A. Kumar, V. D. Kumar, S. Malathi, K. Vengatesan, and M. Ramakrishnan, “Facial Recognition System for Suspect Identification Using A Surveillance Camera”, *Pattern Recognition and Image Analysis*, vol. 28, no. 3, pp. 410-420, July 2018.

[26] X. Jin and X. Tan, “Face Alignment in the Wild: A Survey”, *Computer Vision and Image Understanding*, vol. 162, pp. 1-22, September 2017.

[27] B. Shi, X. Bai, W. Liu, and J. Wang, “Face Alignment with Deep Regression”, *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 1, pp. 183-194, November 2018.

[28] S. Tulyakov, L. A. Jeni, J. F. Cohn, and N. Sebe, “Consistent 3D Face Alignment”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 9, pp. 2250-2264, September 2018.

[29] S. Karamizadeh, S. M. Abdullah, M. Zamani, J. Shayan, and P. Nooralishahi, “Face Recognition Via Taxonomy of Illumination Normalization”, *Multimedia Forensics and Security*, vol. 115, pp. 139-160, July 2017.

[30] M. Gharbi, J. Chen, J. T. Barron, S. W. Hasinoff, and F. Durand, “Deep Bilateral Learning for Real Time Image Enhancement”, *ACM Transactions on Graphics (TOG)*, vol. 36, no. 4, pp. 1-12, July 2017.

[31] O. Oktay, E. Ferrante, K. Kamnitsas, and M. Henrich, "Anatomically Constrained Neural Networks (ACNNs): Application to Cardiac Image Enhancement and Segmentation", *IEEE Transactions on Medical Imaging*, vol. 37, no. 2, pp. 384-395, February 2018.

[32] C. Ding and D. Tao, "Pose Invariant Face Recognition with Homography Based Normalization", *Pattern Recognition*, vol. 66, pp. 144-152, June 2017.

[33] X. Ning, W. Li, B. Tang, and H. He, "BULDP: Biomimetic Uncorrelated Locality Discriminant Projection for Feature Extraction in Face Recognition", *IEEE Transaction on Image Processing*, vol. 27, no. 5, pp. 2575-2586, May 2018.

[34] Z. Chen, W. Huang, and Z. Lv, "Towards A Face Recognition Method Based on Uncorrelated Discriminant Sparse Preserving Projection", *Multimedia Tools and Applications*, vol. 76, no. 17, pp. 17669-17683, September 2017.

[35] K. Cao, Y. Rong, C. Li, X. Tang, and C. C. Loy, "Pose Robust Face Recognition Via Deep Residual Equivariant Mapping", in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, Salt Lake City, USA, March 2018, pp. 5187-5196.

[36] M. Belahcene, A. Chouchane, and H. Ouamane, "3D Face Recognition in Presence of Expressions by Fusion Regions of Interest", in *Proc. 22nd IEEE Conference on Signal Processing and Communications Applications (SIU)*, Trabzon, 23-25, September 2014, pp. 2269-2274.

[37] J. X. Mi and T. Liu, "Multi Step Linear Representation Based Classification for Face Recognition", *IET Computer Vision*, vol. 10, no. 8, pp. 836-841, May 2016.

[38] K. W. Bowyer, K. Chang, and P. Flynn, "A Survey of Approaches and Challenges in 3D and Multimodal 3D + 2D Face Recognition", *Computer Vision and Image Understanding*, vol. 101, no. 1, pp. 1-15, January 2006.

[39] L. Chen and N. Hassanpour, "Survey: How Good Are the Current Advances in Image Set Based Face Identification? Experiments on Three Popular Benchmarks with A Naïve Approach", *Computer Vision and Image Understanding*, vol. 160, pp. 1-23, July 2017.

[40] M. Haghighat, M. Abdel-Mottaleb, and W. Alhalabi, "Fully Automatic Face Normalization and Single Sample Face Recognition in Unconstrained Environments", *Expert Systems with Applications*, vol. 47, pp. 23-34, April 2016.

[41] A. F. Abate, M. Nappi, D. Riccio, and G. Sabatino, "2D and 3D Face Recognition: A Survey", *Pattern Recognition Letters*, vol. 28, no. 14, pp. 1885-1906, October 2007.

[42] S. A. Angadi and V. C. Kagawade, "A Robust Face Recognition Approach Through Symbolic Modeling of Polar FFT Features", *Pattern Recognition*, vol.71, pp. 235-248, November 2017.

[43] M. Chihaoui, A. Elkefi, W. Bellil, and C. Ben Amar, "A Survey of 2D Face Recognition Techniques", *Computers*, vol. 5, no. 4, pp. 1-28, September 2016.

[44] S. P. Mudunuri and S. Biswas, "Low Resolution Face Recognition Across Variations in Pose and Illumination", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 5, pp. 1034-1040, May 2016.

[45] N. P. Ramaiah, E. P. Ijjina, and C. K. Mohan, "Illumination Invariant Face Recognition Using Convolutional Neural Networks", in *Proc. IEEE Signal Processing, Informatics, Communication and Energy Systems (SPICES)*, Calicut, India , February 2015, pp. 1-4.

[46] M. R. Faraji and X. Qi, "Face Recognition Under Varying Illumination Based on Adaptive Homomorphic Eight Local Directional Patterns", *IET Computer Vision*, vol. 9, no. 3, pp. 390-399, November 2014.

[47] X. Cao, W. Shen, L. Yu, Y. Wang, J. Yang, and Z. Zhang, "Illumination Invariant Extraction for Face Recognition Using Neighbouring Wavelet Coefficients", *Pattern Recognition*, vol. 45, no. 4, pp. 1299-1305, April 2012.

[48] G. Chen, T. D. Bui, and A. Krzyżak, "Filter Based Face Recognition Under Varying Illumination", *IET Biometrics*, vol. 7, no. 6, pp. 628-635, March 2018.

[49] L. Zhuang, T. H. Chan, A. Y. Yang, S. S. Sastry, and Y. Ma, "Sparse Illumination Learning and Transfer for Single Sample Face Recognition with Image Corruption and Misalignment", *International Journal of Computer Vision*, vol. 114, no. 2-3, pp. 272-287, September 2015.

[50] X. Ma, H. Song, and X. Qian, "Robust Framework of Single Frame Face Superresolution Across Head Pose, Facial Expression, and Illumination Variations", *IEEE Transactions on Human Machine Systems*, vol. 45, no. 2, pp. 238-250, April 2015.

[51] I. A. Kakadiaris, G. Toderici, G. Evangelopulos, G. Passalis, and D. Chu, "3D-2D Face Recognition with Pose and Illumination Normalization", *Computer Vision and Image Understanding*, vol. 154, pp. 137-151, January 2017.

[52] K. Wang, Z. Chen, Q. J. Wu, and C. Liu, "Illumination and Pose Variable Face Recognition Via Adaptively Weighted ULBP_MHOG and WSRC", *Signal Processing: Image Communication*, vol. 58, pp. 175-186, October 2017.

[53] J. Yang, L. Luo, J. Qian, Y. Tai, F. Zhang, and Y. Xu, "Nuclear Norm Based Matrix Regression with Applications to Face Recognition with Occlusion and Illumination Changes", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 1, pp. 156-171, January 2017.

[54] J. Chen, and Z. Yi, "Sparse Representation for Face Recognition by Discriminative Low Rank Matrix Recovery", *Journal of Visual Communication and Image Representation*, vol. 25, no. 5, pp. 763-773, August 2014.

[55] H.D. Liu, M. Yang, Y. Gao, and C. Cui, "Local Histogram Specification for Face Recognition Under Varying Lighting Conditions", *Image and Vision Computing*, vol. 32, no. 5, pp. 335-347, May 2014.

[56] Y.F. Yu, D.Q. Dai, C.X. Ren, and K.K. Huang, "Discriminative Multilayer Illumination Robust Feature Extraction for Face Recognition", *Pattern Recognition*, vol. 67, pp. 201-212, July 2017.

[57] O. Alkkiomaki, V. Kyrki, Y. Liu, H. Handroos, and H. Kalviainen, "Multimodal Force Vision Sensor Fusion in 6 DOF Pose Tracking", in *Proc. 14th IEEE Conference on Advanced Robotics (ICAR)*, Munich, Germany, June 2009, pp. 1-8.

[58] I. Masi, S. Rawls, G. Medioni, and P. Natarajan, "Pose Aware Face Recognition in the Wild", in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, USA, September 2016, pp. 4838-4846.

[59] C. Ding and D. Tao, "A Comprehensive Survey on Pose Invariant Face Recognition", *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 7, no. 3, pp. 77-86, April 2016.

[60] J. Yim, H. Jung, B. Yoo, C. Choi, D. Park, and J. Kim, "Rotating Your Face Using Multitask Deep Neural Network", in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, Boston, MA, USA, June 2015, pp. 676-684.

[61] R. Huang, S. Zhang, T. Li, and R. He, "Beyond Face Rotation: Global and Local Perception GAN for Fhotorealistic and Identity Preserving Frontal View Synthesis", in *Proc. IEEE International Conference on Computer Vision and Pattern Recognition*, Venice , Italy, December 2017, pp. 2458-2467.

[62] J. W. Wang, N. T. Le, J. S. Lee, and C. C. Wang, "Illumination Compensation for Face Recognition Using Adaptive Singular Value Decomposition in the Wavelet Domain", *Information Sciences*, pp. 69-93, April 2018.

[63] X. Luan, B. Fang, L. Liu, W. Yang, and J. Qian, "Extracting Sparse Error of Robust PCA for Face Recognition in the Presence of Varying Illumination and Occlusion", *Pattern Recognition*, vol. 47, no. 2, pp. 495-508, February 2014.

[64] D. Jeong, M. Lee, and S.W. Ban, "(2D) 2 PCA-ICA: A New Approach for Face Representation and Recognition", in *Proc. IEEE International Conference on Systems, Man and Cybernetics*, August 2009, pp. 1792-1797.

[65] Z.H. Feng, J. Kittler, M. Awais, P. Huber, and X.J. Wu, "Face Detection, Bounding Box Aggregation and Pose Estimation for Robust Facial Landmark Localisation in the Wild", in *Proc. IEEE International Conference on Computer Vision and Pattern Recognition Workshop,* Carlifonia, United States, September 2017, pp.2106-2111.

[66] F. A. Bhat and M. A. Wani, "Elastic Bunch Graph Matching Based Face Recognition Under Varying Lighting, Pose, and Expression Conditions", *IAES International Journal of Artificial Intelligence (IJAI)*, vol. 3, no. 4, pp. 177-182, August 2016.

[67] L.F. Zhou, Y.W. Du, W.S. Li, J.X. Mi, and X. Luan, "Pose Robust Face Recognition with Huffman LBP Enhanced by Divide and Rule Strategy", *Pattern Recognition*, vol.78, pp. 43-55, June 2018.

[68] G.S. J. Hsu, A. Ambikapathi, S.L. Chung, and H.C. Shie, "Robust Cross Pose Face Recognition Using Landmark Oriented Depth Warping", *Journal of Visual Communication and Image Representation*, vol. 53, pp. 273-280, May 2018.

[69] M. Patacchiola and A. Cangelosi, "Head Pose Estimation in the Wild Using Convolutional Neural Networks and Adaptive Gradient Methods", *Pattern Recognition*, vol. 71, pp. 132-143, November 2017.

[70] S. J. Goyal, A. K. Upadhyay, R. Jadon, and R. Goyal, "Real Life Facial Expression Recognition Systems: A Review", in *Smart Computing, Informatics, Smart Innovation, Systems and Technologies*, February 2018, pp. 311-331.

[71] Y. Peng and H. Yin, "Facial Expression Analysis and Expression Invariant Face Recognition by Manifold Based Synthesis", *Machine Vision and Applications*, vol. 29, no. 2, pp. 263-284, February 2018.

[72] A. Moeini and H. Moeini, "Real World and Rapid Face Recognition Toward Pose and Expression Variations Via Feature Library Matrix", *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 5, pp. 969-984, May 2015.

[73] H.H. Tsai and Y.C. Chang, "Facial Expression Recognition Using A Combination of Multiple Facial Features and Support Vector Machine", *Soft Computing*, vol.22, no. 13, pp. 1-17, July 2018.

[74] A. Savran and B. Sankur, "Non Rigid Registration Based Model Free 3D Facial Expression Recognition", *Computer Vision and Image Understanding*, vol. 162, pp. 146-165, September 2017.

[75] H. Ali, V. Sritharan, M. Hariharan, S.K. Zaaba, and M. Elshaikh, "Feature Extraction Using Radon Transform and Discrete Wavelet Transform for Facial Emotion Recognition", in *Proc. 2nd IEEE International Symposium on Robotics and Manufacturing Automation (ROMA),* September 2016, pp. 1-5.

[76] J. A. Martins, R. Lam, J. Rodrigues, and J.D Buf, "Expression Invariant Face Recognition Using A Biological Disparity Energy Model", *Neurocomputing*, vol. 297, pp. 82-93, July 2018.

[77] I. M. Revina and W. S. Emmanuel, "Face Expression Recognition Using LDN and Dominant Gradient Local Ternary Pattern Descriptors", *Journal of King Saud University Computer and Information Sciences*, April 2018.

[78] Y. Li, Y. Wang, J. Liu, and W. Hao, "Expression Insensitive 3D Face Recognition by the Fusion of Multiple Subject Specific Curves", *Neurocomputing*, vol. 275, pp. 1295-1307, January 2018.

[79] A. T. Lopes, E. D. Aguiar, A. F. De Souza, and T. Oliveira-Santos, "Facial Expression Recognition with Convolutional Neural Networks: Coping with Few Data and the Training Sample Order", *Pattern Recognition*, vol. 61, pp. 610-628, January 2017.

[80] M. Nappi, S. Ricciardi, and M. Tistarelli, "Deceiving Faces: When Plastic Surgery Challenges Face Recognition", *Image and Vision Computing*, vol. 54, pp. 71-82, October 2016.

[81] A. Bansal and N. P. Shetty, "Matching Before and After Surgery Faces", *Procedia Computer Science*, vol. 132, pp. 141-148, June 2018.

[82] R. D. Rakshit, S. C. Nath, and D. R. Kisku, "Face Identification Using Some Novel Local Descriptors Under the Influence of Facial Complexities", *Expert Systems with Applications*, vol. 92, pp. 82-94, February 2018.

[83] A. S. O. Ali, V. Sagayan, A. Malik, and A. Aziz, "Proposed Face Recognition System After Plastic Surgery", *IET Computer Vision*, vol. 10, no. 5, pp. 344-350, February 2016.

[84] G. George, R. Boben, B. Radhakrishnan, and L. P. Suresh, "Face Recognition on Surgically Altered Faces Using Principal Component Analysis", in *Proc. IEEE Conference of the Circuit, Power and Computing Technologies (ICCPCT)*, Kollam, India, March 2017, pp. 1-6.

[85] C. C. Chude-Olisah, G. Sulong, U. A. Chude-Okonkwo, and S. Z. Hashim, "Face Recognition Via Edge Based Gabor Feature Representation for Plastic Surgery Altered Images", *EURASIP Journal on Advances in Signal Processing*, vol. 2014, no. 102, pp. 1-15, December 2014.

[86] R. Raghavendra, S. Venkatesh, K. B. Raja, and C. Busch, "Transgender Face Recognition with Off the Shelf Pretrained CNNs: A Comprehensive Study", in *Proc. IEEE International Workshop on Biometrics and Forensics (IWBF)*, Sassari, Italy, January 2018, pp. 1-5.

[87] N. Kohli, D. Yadav, M. Vatsa, R. Singh, and A. Noore, "Supervised Mixed Norm Autoencoder for Kinship Verification in Unconstrained Videos", *IEEE Transactions on Image Processing*, May 2018.

[88] R. Singh, M. Vatsa, H. S. Bhatt, S. Bharadwaj, A. Noore, and S. S. Nooreyezdan, "Plastic surgery: A New Dimension to Face Recognition", *IEEE Transactions on Information Forensics and Security*, vol. 5, no. 3, pp. 441-448, September 2010.

[89] A. H. Sable, S. N. Talbar, and H. A. Dhirbasi, "Recognition of Plastic Surgery Faces the Surgery Types: An Approach with Entropy Based Scale Invariant Features", *Journal of King Saud University Computer and Information Sciences*, April 2017.

[90] I. Manjani, H. Sumerkan, P. J. Flynn, and K. W. Bowyer, "Template Aging in 3D and 2D Face Recognition", in *Proc. IEEE Conference on Biometrics Theory, Applications and Systems (BTAS)*, New York, USA, September 2016, pp. 1-6.

[91] C.C. Ng, M. H. Yap, Y.T. Cheng, and G.S. Hsu, "Hybrid Ageing Patterns for Face Age Estimation", *Image and Vision Computing*, vol. 69, pp. 92-102, January 2018.

[92] Z. Li, D. Gong, X. Li, and D. Tao, "Aging Face Recognition: A Hierarchical Learning Model Based on Local Patterns Selection", *IEEE Transactions on Image Processing*, vol. 25, no. 5, pp. 2146-2154, May 2016.

[93] K. Ricanek and T. Tesafaye, "Morph: A Longitudinal Image Database of Normal Adult Age Progression", in *Proc. 7th IEEE Conference on Automatic Face and Gesture Recognition*, Southampton, United Kingdom, July 2006, pp. 341-345.

[94] H. Zhou and K.M. Lam, "Age Invariant Face Recognition Based on Identity Inference from Appearance Age", *Pattern Recognition*, vol. 76, pp. 191-202, April 2018.

[95] C. Xu, Q. Liu, and M. Ye, "Age Invariant Face Recognition and Retrieval by Coupled Auto Encoder Networks", *Neurocomputing*, vol. 222, pp. 62-71, January 2017.

[96] D.X. Zhang, P. An, and H.X. Zhang, "Application of Robust Face Recognition in Video Surveillance Systems", *Optoelectronics Letters*, vol. 14, no. 2, pp. 152-155, February 2018.

[97] L. Best-Rowden and A. K. Jain, "Longitudinal Study of Automatic Face Recognition", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 1, pp. 148-162, January 2018.

[98] F. Zhao, J. Feng, J. Zhao, W. Yang, and S. Yan, "Robust Lstm Autoencoders for Face De-occlusion in the Wild", *IEEE Transactions on Image Processing*, vol. 27, no. 2, pp. 778-790, February 2018.

[99] G. Gao, J. Yang, X.Y. Jing, F. Shen, W. Yang, and D. Yue, "Learning Robust and Discriminative Low Rank Representations for Face Recognition with Occlusion", *Pattern Recognition*, vol. 66, pp. 129-143, June 2017.

[100] H. Jia, and A. M. Martinez, "Face Recognition with Occlusion in the Training and Testing Sets", in *Proc. 6th IEEE International Conference on Image and Video Processing*, Ohio, United States, September 2008, pp. 1-6.

[101] S. Zhao, "Pixel Level Occlusion Detection Based on Sparse Representation for Face Recognition", *Optik International Journal for Light and Electron Optics*, vol. 168, pp. 920-930, September 2018.

[102] Y. Fu, X. Wu, Y. Wen, and Y. Xiang, "Efficient Locality Constrained Occlusion Coding for Face Recognition", *Neurocomputing*, vol. 260, pp. 104-111, October 2017.

[103] Y. Long, F. Zhu, L. Shao, and J. Han, "Face Recognition with A Small Occluded Training Set Using Spatial and Statistical Pooling", *Information Sciences*, March 2017.

[104] B. Li, H. Chang, S. Shan, and X. Chen, "Low Resolution Face Recognition Via Coupled Locality Preserving Mappings", *Signal Processing Letters*, vol. 17, no. 1, pp. 20-23, January 2010.

[105] Y. Chu, T. Ahmad, G. Bebis, and L. Zhao, "Low Resolution Face Recognition with Single Sample Per Person", *Signal Processing*, vol. 141, pp. 144-157, December 2017.

[106] P. Zhang, X. Ben, W. Jiang, R. Yan, and Y. Zhang, "Coupled Marginal Discriminant Mappings for Low Resolution Face Recognition", *Optik International Journal for Light and Electron Optics*, vol. 126, no. 23, pp. 4352-4357, December 2015.

[107] J. F. Pereira, R. M. Barreto, G. D. Cavalcanti, and R. Tsang, "A Robust Feature Extraction Algorithm Based on Class Modular Image Pincipal Component Analysis for Face Verification", in *Proc. IEEE Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Prague, Czech Republic, June 2011, pp. 1469-1472.

[108] A. Aktel, B. Yagmahan, T. Özcan, M. M. Yenisey, and E. Sansarcı, "The Comparison of the Metaheuristic Algorithms Performances on Airport Gate Assignment Problem", *Transportation Research Procedia*, vol. 22, pp. 469-478, January 2017.

[109] M. A. Turk and A. P. Pentland, "Face Recognition Using Eigenfaces", in *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Anchorage, Alaska, May 1991, pp. 586-591.

[110] S. Yi, Z. Lai, Z. He, Y.M. Cheung, and Y. Liu, "Joint Sparse Principal Component Analysis", *Pattern Recognition*, vol. 61, pp. 524-536, January 2017.

[111] E. Naz, U. Farooq, and T. Naz, "Analysis of Principal Component Analysis Based and Fisher Discriminant Analysis Based Face Recognition Algorithms", in *Proc. IEEE Conference on Emerging Technologies, Singapore*, March 2006, pp. 121-127.

[112] K. Shrivastava, S. Manda, P. Chavan, T. Patil, and S. Sawant-Patil, "Conceptual Model for Proficient Automated Attendance System Based on Face Recognition and Gender Classification Using Haar-Cascade, LBPH Algorithm Along with LDA Model", *International Journal of Applied Engineering Research*, vol. 13, no. 10, pp. 8075-8080, July 2018.

[113] M. Li and B. Yuan, "2D-LDA: A Statistical Linear Discriminant Analysis for Image Matrix", *Pattern Recognition Letters*, vol. 26, no. 5, pp. 527-532, April 2005.

[114] M. S. Bartlett, J. R. Movellan, and T. J. Sejnowski, "Face Recognition by Independent Component Analysis", *IEEE Transactions on Neural Networks*, vol. 13, no. 6, pp. 1450-1464, July 2002.

[115] K. Delac, M. Grgic, and S. Grgic, "Independent Comparative Study of PCA, ICA, and LDA on the FERET Data Set", *International Journal of Imaging Systems and Technology*, vol. 15, no. 5, pp. 252-260, November 2005.

[116] H. Hu, "ICA Based Neighbourhood Preserving Analysis for Face Recognition", *Computer Vision and Image Understanding*, vol. 112, no. 3, pp. 286-295, December 2008.

[117] D. Huang, C. Shan, M. Ardabilian, Y. Wang, and L. Chen, "Local Binary Patterns and Its Application to Facial Image Analysis: A Survey", *IEEE Transactions on Systems, Man, Cybernetics, Applications and Review*, vol. 41, no. 6, pp. 765-781, November 2011.

[118] Y. Duan, J. Lu, J. Feng, and J. Zhou, "Context Aware Local Binary Feature Learning for Face Recognition", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 5, pp. 1139-1153, May 2018.

[119] T. Ahonen, A. Hadid, and M. Pietikainen, "Face Description with Local Binary Patterns: Application to Face Recognition", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 12, pp. 2037-2041, December 2006.

[120] C. Kotropoulos, I. Pitas, S. Fischer, and B. Duc, "Face Authentication Using Morphological Dynamic Link Architecture", in *Proc. International Conference on Audio and Video Based Biometruc Person Authentication*, Berlin, Germany, January 1997, pp. 169-176.

[121] M. Lades, J.C. Vorbruggen, J. Buhmann, J. Lange, and C. V. Malsburg, "Distortion Invariant Object Recognition in the Dynamic Link Architecture", *IEEE Transactions on Computers*, vol. 42, no. 3, pp. 300-311, March 1993.

[122] A. Tefas, C. Kotropoulos, and I. Pitas, "Variants of Dynamic Link Architecture Based on Mathematical Morphology for Frontal Face Authentication", in *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Anchorage, Alaska, June 1998, pp. 814-819.

[123] D. S. Bolme, “Elastic Bunch Graph Matching”, Masters Dissertation, Department of Computer Science, Colorado State University, Col., USA, September 2003.

[124] B. M. Lahasan, I. Venkat, M. A. Al-Betar, S. L. Lutfi, and P.D. Wilde, “Recognizing Faces Prone to Occlusions and Common Variations Using Optimal Face Subgraphs”, *Applied Mathematics and Computation*, vol. 283, pp. 316-332, June 2016.

[125] L. Wiskott, J.M. Fellous, N. Kruger, and C.V. Malsburg, “Face Recognition by Elastic Bunch Graph Matching”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 5, pp. 775-779, September 1997.

[126] R. Brunelli and T. Poggio, “Face Recognition: Features Versus Templates”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, no. 10, pp. 1042-1052, October 1993.

[127] D. Chawla and M. C. Trivedi, “A Comparative Study on Face Detection Techniques for Security Surveillance”, *Advances in Computer and Computational Sciences*: *Springer*, Singapore, January 2018, pp. 531-541.

[128] I. Kotsia and I. Pitas, “Facial Expression Recognition in Image Sequences Using Geometric Deformation Features and Support Vector Machines”, *IEEE Transactions on Image Processing*, vol. 16, no. 1, pp. 172-187, January 2007.

[129] A. Rehman and T. Saba, “Neural Networks for Document Image Preprocessing: State of the Art”, *Artificial Intelligence Review*, vol. 42, no. 2, pp. 253-273, August 2014.

[130] Q. V. Le, “Building High Level Features Using Large Scale Unsupervised Learning”, in *Proc. IEEE Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Vancouver, British Columbia, Canada, July 2013, pp. 8595-8598.

[131] R. Ranjan, S. Sankaranarayanan, C. D. Castillo, and R. Chellappa, “An All in One Convolutional Neural Network for Face Analysis”, in *Proc. 12th IEEE Conference on Automatic Face and Gesture Recognition (FG 2017)*, Washington, DC, USA, June 17, pp. 17-24.

[132] Y. Zhang, Y. Lu, H. Wu, C. Wen, and C. Ge, “Face Occlusion Detection Using Cascaded Convolutional Neural Network”, in *Proc. Chinese Conference on Biometric Recognition*, March 2016, pp. 720-727.

[133] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real Time Object Detection with Region Proposal Networks", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, no. 6, pp. 1137-1149, June 2017.

[134] F. Juefei-Xu, V. N. Boddeti, and M. Savvides, "Local Binary Convolutional Neural Networks", in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, Hawaii, USA, February 2017, pp.4284-4293.

[135] S. Hijazi, R. Kumar, and C. Rowen, "Using Convolutional Neural Networks for Image Recognition", *Cadence Design Systems Inc*., Ed., San Jose, CA, January 2015.

[136] P. Rasti, T. Uiboupin, S. Escalera, and G. Anbarjafari, "Convolutional Neural Network Super Resolution for Face Recognition in Surveillance Monitoring", in *Proc. International Conference on Articulated Motion and Deformable Objects*, Kiev, Ukraine, March 2016, pp. 175-184.

[137] S.A. Radzi, K. Syazana-Itqan, and N.M. Saad, "A MATLAB Based Convolutional Neural Network Approach for Face Recognition System", *Journal of Bioinformatics, Proteomics and Image Analysis*, vol. 2, no. 1, pp. 71-75, August 2015.

[138] A. Krizhevsky and G. Hinton, "Learning Multiple Layers of Features from Tiny Images", *Technical Report*, University of Toronto, April 15, 2009.

[139] H. Li, Z. Lin, X. Shen, J. Brandt, and G. Hua, "A Convolutional Neural Network Cascade for Face Detection", in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, Boston, MA, USA, August 2015, pp. 5325-5334.

[140] C. Szegedy, W. Liu, Y. Jia, and P. Sermanet, "Going Deeper with Convolutions", in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, Boston, MA, USA, 2015, pp. 1-9.

[141] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large Scale Image Recognition", in *Proc. IEEE Asian Conference on Pattern Recognition*, Kuala Lumpur, Malaysia, 2015, pp.730-734.

[142] K. He, X. Zhang, S. Ren, and J. Sun, "Identity Mappings in Deep Residual Networks", in *Proc. 14th European Conference on Computer Vision*, Netherlands, 2016, Amsterdam, Netherlands, pp. 630-645.

[143] H. Khalajzadeh, M. Mansouri, and M. Teshnehlab, "Face Recognition Using Convolutional Neural Network and Simple Logistic Classifier", in *Proc. International*

Conference on Soft Computing in Industrial Applications: Springer, pp. 197-207, June 2014.

[144] A. Rikhtegar, M. Pooyan, and M. T. Manzuri-Shalmani, "Genetic Algorithm Optimised Structure of Convolutional Neural Network for Face Recognition Applications", *IET Computer Vision*, vol. 10, no. 6, pp. 559-566, February 2016.

[145] I. Ševo and A. Avramović, "Convolutional Neural Network Based Automatic Object Detection on Aerial Images", *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 5, pp. 740-744, May 2016.

[146] L. Breiman, "Random Forests", *Machine Learning*, vol. 45, no. 1, pp. 5-32, October 2001.

[147] Y. See, N. Noor, J. Low, and E. Liew, "Investigation of Face Recognition Using Gabor Filter with Random Forest as Learning Framework", in *Proc. 1EEE Region 10 Conference TENCON*, Penang, Malaysia, 2017, pp. 1153-1158.

[148] J. Gu, L. Jiao, S. Yang, R. Wang and P. Chen, "Random Subspace Based Ensemble Sparse Representation", *Pattern Recognition*, vol. 74, pp. 544-555, February 2018.

[149] B. Yang, J.M. Cao, D.P. Jiang, and J.D. Lv, "Facial Expression Recognition Based on Dual Feature Fusion and Improved Random Forest Classifier", *Multimedia Tools and Applications*, vol. 77, no. 16, pp. 20477-20499, December 2017.

[150] D. Ghimire, J. Lee, Z.N. Li, and S. Jeong, "Recognition of Facial Expressions Based on Salient Geometric Features and Support Vector Machines", *Multimedia Tools and Applications*, vol. 76, no. 6, pp. 7921-7946, March 2017.

[151] Z. Qi, Y. Tian, and Y. Shi, "Robust Twin Support Vector Machine for Pattern Classification", *Pattern Recognition*, vol. 46, no. 1, pp. 305-316, January 2013.

[152] G. Lei, X.H. Li, J.I. Zhou, and X.G. Gong, "Geometric Feature Based Facial Expression Recognition Using Multiclass Support Vector Machines", in *Proc. IEEE International Conference on Granular Computing*, Nanchang, China, 2009, pp. 318-321.

[153] J. Zheng, H. Pan, and J. Cheng, "Rolling Bearing Fault Detection and Diagnosis Based on Composite Multiscale Fuzzy Entropy and Ensemble Support Vector Machines", *Mechanical Systems and Signal Processing*, vol. 85, pp. 746-759, February 2017.

[154] L. Khedher, I. A. Illán, J. M. Górriz, J. Ramírez, A. Brahim, and A. Meyer-Baese, "Independent Component Analysis Support Vector Machine Based Computer-Aided Diagnosis System for Alzheimer's with Visual Support", *International Journal of Neural Systems*, vol. 27, no. 3, pp. 152-174 , May 2017.

[155] F. Noroozi, T. Sapiński, D. Kamińska, and G. Anbarjafari, "Vocal Based Emotion Recognition Using Random Forests and Decision Tree", *International Journal of Speech Technology*, vol. 20, no. 2, pp. 239-246, June 2017.

[156] K. Wu, Z. Zheng, and S. Tang, "BVDT: A Boosted Vector Decision Tree Algorithm for Multi Class Classification Problems", *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 31, no. 5, pp. 198-215, May 2017.

[157] L. Du and H. Hu, "Modified Classification And Regression Tree for Facial Expression Recognition with Using Difference Expression Images", *Electronics Letters*, vol. 53, no. 9, pp. 590-592, March 2017.

[158] Z. Zhou, A. Wagner, H. Mobahi, J. Wright, and Y. Ma, "Face Recognition with Contiguous Occlusion Using Markov Random Fields", in *Proc. 12th IEEE International Conference on Computer Vision*, Kyoto, Japan, 2009, pp. 1050-1057.

[159] L. Jiang, C. Li, S. Wang, and L. Zhang, "Deep Feature Weighting for Naive Bayes and Its Application to Text Classification", *Engineering Applications of Artificial Intelligence*, vol. 52, pp. 26-39, June 2016.

[160] C.Z. Gao, Q. Cheng, P. He, W. Susilo, and J. Li, "Privacy Preserving Naive Bayes Classifiers Secure Against the Substitution then Comparison Attack", *Information Sciences*, vol. 444, pp. 72-88, May 2018.

[161] D. Chen, X. Cao, D. Wipf, F. Wen, and J. Sun, "An Efficient Joint Formulation for Bayesian Face Verification", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 1, pp. 32-46, January 2017.

[162] R. Yan, J. Wen, J. Cao, Y. Xu, and J. Yang, "An Optimized Naive Bayesian Method for Face Recognition", in *Proc. International Conference on Cognitive Systems and Signal Processing*, Beijing, China, 2016, pp. 126-135.

[163] S. Liao, Z. Lei, D. Yi, and S.Z. Li, "A Benchmark Study of Large Scale Unconstrained Face Recognition", in *Proc. IEEE International Joint Conference on Biometrics* (IJCB), Clearwater, Florida, USA, 2014, pp. 1-8.

[164] S. Jia, T. Lansdall-Welfare, and N. Cristianini, "Gender Classification by Deep Learning on Millions of Weakly Labelled Images", in *Proc. IEEE 16th International Conference on Data Mining Workshops (ICDMW)*, Barcelona, Spain, 2016, pp. 462-467.

[165] A. S. O. Ali, V. Sagayan, A. Malik, and A. Aziz, "Proposed Face Recognition System After Plastic Surgery", *IET Computer Vision*, vol. 10, no. 5, pp. 342-348, February 2016.

[166] Y. Guo, L. Zhang, Y. Hu, X. He, and J. Gao, "Ms-celeb-1m: A Dataset and Benchmark for Large Scale Face Recognition", in *Proc. European Conference on Computer Vision*, Amsterdam, Netherlands, 2016, pp. 87-102.

[167] B. Hu, Z. Lu, H. Li, and Q. Chen, "Convolutional Neural Network Architectures for Matching Natural Language Sentences", in *Proc. Advances in Neural Information Processing Systems*, Montreal, Canada, 2014, pp. 2042-2050.

[168] C. Munteanu and A. Rosa, "Grey Scale Image Enhancement as an Automatic Process Driven by Evolution", *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 34, no. 2, pp. 1292-1298, April 2004.

[169] W. Xi, T. Wu, K. Yan, X. Yang, X. Jiang, and N. Kwok, "Restoration of Online Video Ferrography Images for Out of Focus Degradations", *EURASIP Journal on Image and Video Processing*, vol. 2018, no. 31, pp. 1-11, December 2018.

[170] J.P. Pelteret, B. Walter, and P. Steinmann, "Application of Metaheuristic Algorithms to the Identification of Nonlinear Magneto Viscoelastic Constitutive Parameters", *Journal of Magnetism and Magnetic Materials*, vol. 462, pp. 116-131, October 2018.

[171] X.S. Yang, "A New Metaheuristic Bat Inspired Algorithm", in *Proc. Nature Inspired Cooperative Strategies for Optimization (NICSO 2010)*, Berlin, Germany, 2010, pp. 65-74.

[172] X.S. Yang and S. Deb, "Cuckoo Search Via Lévy Flights", in *Proc. IEEE World Congress on Nature and Biologically Inspired Computing (NaBIC)*, Coimbatore, India, 2009, pp. 210-214.

[173] A. K. Bhandari, V. K. Singh, A. Kumar, and G. K. Singh, "Cuckoo Search Algorithm and Wind Driven Optimization Based Study of Satellite Image Segmentation for Multilevel Thresholding Using Kapur's Entropy", *Expert Systems with Applications*, vol. 41, no. 7, pp. 3538-3560, June 2014.

[174] E. Daniel and J. Anitha, "Optimum Wavelet Based Masking for the Contrast Enhancement of Medical Images Using Enhanced Cuckoo Search Algorithm", *Computers in Biology and Medicine*, vol. 71, pp. 149-155, April 2016.

[175] Z. Krbcova and J. Kukal, "Relationship Between Entropy and SNR Changes in Image Enhancement", *EURASIP Journal on Image and Video Processing*, vol. 2017, no. 83, pp. 1-8, December 2017.

[176] S. Suresh and S. Lal, "Modified Differential Evolution Algorithm for Contrast and Brightness Enhancement of Satellite Images", *Applied Soft Computing*, vol. 61, pp. 622-641, December 2017.

[177] N. Kalchbrenner, E. Grefenstette, and P. Blunsom, "A Convolutional Neural Network for Modelling Sentences", in *Proc. IEEE 16th International Conference on Data Mining Workshops (ICDMW)*, Barcelona, Spain, 2016, pp. 400-4410.

[178] C. Wachinger, M. Reuter, and T. Klein, "DeepNAT: Deep Convolutional Neural Network for Segmenting Neuroanatomy", *Neuroimage*, vol. 170, pp. 434-445, April 2018.

[179] U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, H. Adeli, and D. Subha, "Automated EEG Based Screening of Depression Using Deep Convolutional Neural Network", *Computer Methods and Programs in Biomedicine*, vol. 161, pp. 103-113, July 2018.

[180] M. Sabokrou, M. Fayyaz, M. Fathy, Z. Moayed, and R. Klette, "Deep Anomaly: Fully Convolutional Neural Network for Fast Anomaly Detection in Crowded Scenes", *Computer Vision and Image Understanding*, February 2018.

[181] Z. Ye, M. Wang, Z. Hu, and W. Liu, "An Adaptive Image Enhancement Technique by Combining Cuckoo Search and Particle Swarm Optimization Algorithm", *Computational Intelligence and Neuroscience*, vol. 2015, no. 13, pp.1-13, January 2015.

[182] M. Shakeri, M. Dezfoulian, H. Khotanlou, A. Barati, and Y. Masoumi, "Image Contrast Enhancement Using Fuzzy Clustering with Adaptive Cluster Parameter and Sub-Histogram Equalization", *Digital Signal Processing*, vol. 62, pp. 224-237, March 2017.

[183] Y. Chang, C. Jung, P. Ke, H. Song, and J. Hwang, "Automatic Contrast Limited Adaptive Histogram Equalization with Dual Gamma Correction", *IEEE Access*, vol. 6, pp. 11782-11792, January 2018.

[184] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition", in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, Nevada, USA, 2016, pp. 770-778.

[185] Z. Shi. M. mei Zhu, B. Guo, M. Zhao, and C. Zhang, "Nighttime low illumination image enhancement with single image using bright/dark channel prior", *EURASIP Journal on Image and Video Processing,* vol. 2018, pp. 13, 2018.

[186] G. Zhang, L. Niu, Y. Zhang, and Y. Ke, "Number of Faces Affect Performance of Rigid Joint Alignment", in *Proc. IEEE 2nd International Conference on Image, Vision and Computing (ICIVC),* Chengdu, China, 2017, pp. 249-253.

[187] L. Cong, Z. Ting, L. Chongshan, W. Wei, Z. Xiang , and F. Weiyan, "Improved Explicit Shape Regression Face Alignment Algorithm", in *Proc. IEEE Chinese Control and Decision Conference (CCDC),* Shenyang, China, 2018, pp. 1166-1169.

[188] Z. An, W. Deng, J. Hu, Y. Zhong, and Y. Zhao, "APA: Adaptive Pose Alignment for Pose-Invariant Face Recognition", *IEEE Access*, vol. 7, pp. 14653-14670, January 2019.

[189] B.V. Le, S. Lee, T. Le, and Y. Yoon, "Using Weighted Dynamic Range For Histogram Equalization to Improve the Image Contrast", *EURASIP Journal on Image and Video Processing,* vol. 10, pp. 44-68, March 2014.

[190] J.R. Tang and N.A.M. Isa, "Bi-Histogram Equalization Using Modified Histogram Bins", *Applied Soft Computing,* vol. 55, pp. 31-43, June 2017.

[191] Z. Krbcova and J. Kukal, "Relationship Between Entropy and SNR Changes in Image Enhancement", *EURASIP Journal on Image and Video Processing*, vol. 44, pp. 83-97, October 2017.