

Person Authentication based on Multimodal Biometric Systems

THESIS

Submitted in partial fulfillment of the requirement for the degree of

DOCTOR OF PHILOSOPHY

by

Himanshu Purohit
(2015PHXF0502P)

Under the Supervision of

Dr. Pawan Kamalkishor Ajmera



BITS Pilani
Pilani | Dubai | Goa | Hyderabad

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE, PILANI

2023

Dedication

To:

God Almighty,

My Grandparents,

My Parents,

My Sisters and Brothers-in- law,

My Brother and Sister-in- law,

Loving Wife,

Raghav, Devansh, Shritu, Akshita, Heeral, Yuvi, Vihaan, ChandraShekhar, Bhavesh

and best friends,

Along with all hardworking and respected

Teachers,

who instilled in me perseverance and commitment and relentlessly encouraged me to

strive for the excellence



**BIRLA INSTITUTE OF TECHNOLOGY AND
SCIENCE, PILANI-333031, (RAJASTHAN), INDIA**

CERTIFICATE

This is to certify that the thesis entitled “**Person Authentication Based on Multimodal Biometric Systems**” and submitted by **Himanshu Purohit**, ID No. **2015PHXF0502P** for the award of Ph.D. of the institute embodies original work done by him under my supervision.

Dr. Pawan K. Ajmera

Associate Professor,

Electrical and Electronics Engineering Department,

Birla Institute of Technology and Science (BITS), Pilani-Campus,

Rajasthan-333031

14/04/2023

Acknowledgements

I would like to take this opportunity to convey my heartfelt gratitude and indebtedness to my supervisor Dr. Pawan K. Ajmera for his patience, support, motivation, immense knowledge, and kindness during my research tenure. His guidance helped me all the time in writing my thesis and research work. I could not have imagined having a better mentor and advisor for my research work.

Besides my advisor, I would like to thank my Doctoral Advisory Committee members Prof. Surekha Bhanot, Dr. Sandeep Joshi, and Dr. Anantha Krishna Chintanpalli for their encouragement, insightful comments, and hard questions.

I wish to express my humble gratitude to Vice Chancellor, Prof. V. Ramgopal Rao, Ex-VC, Souvik Bhattacharyya, and Director Prof. Sudheer Kumar Barai, for providing me the opportunity to pursue doctoral studies. I express my gratitude to Dr. S. K. Verma, Dean of Administration, Prof. Shamik Chakraborty, Associate Dean AGSRD, Prof. Navneet Gupta, HOD- EEE, and Prof. H. D. Mathur, for their official support and encouragement. I extend my thanks to SRCDD, AGSRD, and Library staff for their support.

I would also like to thank my friend, Mr. Rajeev Gaur, Sir, Dr. Manish Dadhich, Ms. Hemlata Paliwal, Ms. Anima Paliwal, Mr. Amit Purohit, Ms. Vandana Paliwal, Ms. Priyamvada Paliwal, and my friends at BITS & SPSU.

Last but not least I would like to thank my parents, Sh. Sohan Lal ji Paliwal and Mrs. Tulsi Paliwal for their blessings and support, my wife Neelam Purohit for her unconditional love, moral support, and care, and my beloved son Mr. Raghav Purohit for his patience and sacrificing his childhood time for my research work.

April 2023

Himanshu Purohit

Abstract

Biometrics is the science of utilizing natural attributes of the human body for identification and authentication. Applications of biometrics vary from employee attendance, access control, and welfare distribution to traveling, border crossing, forensic science, etc. The global forecast for the biometric system market is expected to grow at a CAGR of 13.4% and reach USD 68.6 billion by 2025. This makes biometric technology an area of research interest.

Public distribution, bank transactions, and property rights are sensitive to imposter access and need to be secure with high-end security and full access control. Biometric authentication can help in controlling malpractices and corruption in such cases. An efficient recognition system with high reliability, feasible implementation, and commercial viability is the need of the hour.

This research work focuses on person authentication based on multimodal biometric systems. Multimodal biometric systems are more efficient and reliable as compared to unimodal biometric systems, where multiple traits are used for authentication. The main backbone of the multimodal system is the fusion of features (extracted from traits) where different features are merged at different levels for better authentication results. To understand the existing fusion methods, optimization of the fusion process, the efficient practical realization of the biometric system, Convolution Neural Network (CNN) based extraction and fusion, and study of user perception about multimodal biometric systems are explored.

The initial part of the work is based on unimodal and multimodal experiments for different feature extraction methods with multiple classifiers. Present multimodal, multilevel-multi-classification authentication gives more weightage to feature-level fusion. The fusion process is efficient and less complex in the case of compatible features. Incompatible features create issues in feature space representation and in classification. Optimization of

the feature selection process and fusion can degrade the dimensionality issue of multimodal biometric fusion. In the present work, Modified Grey Wolf Optimization (MGWO) has been used for optimal feature selection with reduced complexities.

Online and remote user authentication is a popular service to offer for different purposes. Security and reliability are the main factors for the continuous authentication process. In this work, the emphasis is on exploring the implementation feasibility of multimodal systems for continuous authentication in the online environment. Look up Convolution Neural Network (LCNN) based Continuous User Biometric Authentication (CUBA) for online applications, using Salp Swarm Optimization (SSO) has improved the classification capacity of the proposed system.

CNN-based algorithms are known for the efficient implementation of complex feature extraction and classification problems. Still, the size of the computation and execution time (speed) are important factors for CNN-based implementation. The proposed Modified VGG16 (MVGG16) based multimodal authentication process outperformed standard VGG16 methods in terms of magnitude of computations and time efficiency.

Getting insights about user awareness and perceptions of technology adoption, and its correlation with their decision-making is an important area of research. In the final phase, Statistical Equation Modelling- Artificial Neural Network (SEM ANN) approach has for individual's understanding of the biometric system's usability. Data quality is verified via the Reliability and Validity test, before using an exploratory and confirmatory analysis. SEM's limitation of linearity is taken care of using ANN's contribution. This part is implemented in Statistical Package for the Social Sciences (SPSS). The work presented in this thesis can be helpful for the implementation of multimodal biometric fusion-based systems and further study of users' adoption of biometric solutions.

Table of Contents

Dedication	ii
Certificate	iv
Acknowledgement	v
Abstract	vi
Table of Contents	viii
List of Figures	xii
List of Tables	xvii
List of Abbreviations	xix
CHAPTER 1 Introduction	1
1.1 Background	1
1.1.1 Types of Multimodal Biometric System	5
1.1.2 Advantages of Multimodal Biometrics	6
1.1.3 Biometric Fusion	8
1.2 Motivation	10
1.3 Literature Survey/Related Work	12
1.3.1 Introduction	12
1.3.2 Remote Biometric Recognition	15
1.3.3 CNN-based Biometric Recognition	20
1.3.4 Study of Work based on User Perception of Biometric Applications	25
1.4 Research Gaps	30
1.5 Research Objectives	32
1.6 Thesis Outline	33
CHAPTER 2 Database, Fusion and Multimodal Biometric Experiments	34
2.1 Introduction of Biometric Databases	34
2.1.1 Iris Databases	34
2.1.2 Ear Databases	37
2.1.3 Face Databases	38
2.1.4 Fingerprint Databases	40

2.2	Fusion of Features	41
2.3	Performance Metrics of Biometric System	45
2.4	Baseline Methodologies	49
2.4.1	Multilevel Fusion Type 1	49
2.4.1.1	Proposed Methodology	50
2.4.1.2	Results	55
2.4.2	Multimodal Multilevel Fusion of Type 2	58
2.4.2.1	Proposed Methodology	59
2.4.2.2	Results	65
2.4.3	Multimodal Multilevel Fusion of Type 3	67
2.4.3.1	Proposed Methodology	68
2.4.3.2	Results	71
2.5	Summary	73
CHAPTER 3 Optimal Feature Level Fusion for Authentication		74
3.1	Introduction	74
3.1.1	Problem Definition	75
3.2	Proposed Methodology	76
3.2.1	Preprocessing	77
3.2.2	Feature Extraction	78
3.2.3	Gabor Feature Extraction from Palmprint	78
3.2.4	HMSB-based Texture Feature Extraction from Fingerprint	81
3.2.5	Shape and Texture Feature Extraction from Ear	81
3.3	Optimal Feature Level Fusion Gray Wolf Optimization (OGWO)	84
3.3.1	Recognition by Multi-kernel SVM	86
3.4	Results and Discussions	88
3.4.1	Dataset Description	88
3.4.2	Evaluation Metrics	88
3.4.3	Comparative Analysis	89
3.5	Summary	93
CHAPTER 4 Continuous Biometric User Authentication for e- Proctoring		94
4.1	Introduction	94
4.2	Problem Methodology	97
4.2.1	CBUA-OE System	98

4.3	Continuous Biometric User Authentication System for Online Examinations (CBUA-OE)	99
4.3.1	Optimal Feature Extraction using the Modified Wolf Optimization	99
4.3.2	Weight Fusion using Optimal Feature Fusion (OFF) Algorithm	101
4.3.3	Action Classification using LCNN-SSO-based Classifier	104
4.4	Results and Discussions	109
4.4.1	Dataset Description	109
4.4.2	Performance Analysis	110
4.5	Summary	120
CHAPTER 5 Modified VGG16 based Multimodal Biometric Authentication		121
5.1	Research Methodology	122
5.1.1	Overview of Proposed Method Structure	122
5.1.2	Modules Used in CNN-based Multimodal Architecture	124
5.1.3	Proposed Methods	127
5.2	Visual Geometry Group	136
5.2.1	Architecture of VGG	137
5.2.2	Modified VGG16- (MVGG16)	140
5.3	Results and Discussions	144
5.3.1	Dataset	144
5.3.2	Performance Metrics	145
5.3.3	Experimental Parameters	145
5.3.4	Result Analysis	146
5.4	Summary	151
CHAPTER 6 SEM-ANN based Analysis of User’s Awareness and Acceptability of Multimodal Biometrics		152
6.1	Introduction	152
6.2	Conceptual Model	156
6.3	Theoretical Constructs and Development of Hypotheses	157
6.3.1	Hypotheses for the Research	157
6.4	Research Methodology	159
6.4.1	Survey Instruments and Data Collection	159
6.4.2	Measurement Used in Study	160
6.5	Data Analysis and Interpretation	162
6.5.1	Measurement Model	164
6.5.2	Artificial Neural Network	171
6.5.3	Results of ANN Modeling	174

6.6	Theoretical Implications	176
6.7	Limitations and Future Research	178
6.8	Summary	178
CHAPTER 7 Conclusions and Future Scope		180
7.1	Preamble	180
7.2	Motivation	180
7.3	Conclusions	182
7.4	Future Recommendations	183
References		184
List of Publications		201
Brief Biography of Candidate and Supervisor		202
Annexure		204

List of Figures

Figure 1.1	Biometric healthcare market size data (US) from 2014-2025	2
Figure 1.2	Classification of biometric traits	3
Figure 1.3	Fusion levels in multimodal biometric systems	8
Figure 1.4	Finger multimodal biometric schemes using t-norm	21
Figure 2.1	Sample iris images	35
Figure 2.2	CASIA iris v1 images	36
Figure 2.3	CASIA iris v2 images	36
Figure 2.4	MMU v1 iris	36
Figure 2.5	CUHK v1.0 iris images	36
Figure 2.6	BATH iris images	37
Figure 2.7	MMU v2.0 iris images	37
Figure 2.8	IITD ear database images	38
Figure 2.9	ORL face images	39
Figure 2.10	Yale B face images	39
Figure 2.11	CASIA fingerprint image database images	40
Figure 2.12	FVC databases image	41
Figure 2.13	BioSec fingerprint database images	41
Figure 2.14	Different types of the multibiometric combination	42
Figure 2.15	Standard score level fusion	44
Figure 2.16	Equal error rate	47
Figure 2.17	Fusion of iris, face & ear biometric traits	50
Figure 2.18	Iris recognition process from sample image	51
Figure 2.19	Ear recognition process block diagram	52
Figure 2.20	Ear recognition process using canny edge detection	52
Figure 2.21	Face recognition process from sample image	54
Figure 2.22	ROC curve for ear	56
Figure 2.23	ROC curve for face	56
Figure 2.24	LTP histogram values of face images	57

Figure 2.25	ROC curve for iris	57
Figure 2.26	ROC curve for feature level fusion	57
Figure 2.27	ROC curve for multilevel multimodal fusion	57
Figure 2.28	Comparative analysis of different models	57
Figure 2.29	Comparative analysis of different models	58
Figure 2.30	Structural flow of multimodal multilevel fusion of face, ear, and iris	59
Figure 2.31	Face processing	61
Figure 2.32	Ear processing	61
Figure 2.33	Iris processing	62
Figure 2.34	ROC curve for unimodal face	65
Figure 2.35	ROC curve for unimodal ear	65
Figure 2.36	ROC curve for unimodal iris	65
Figure 2.37	ROC curve for feature level fusion	65
Figure 2.38	ROC curve for multilevel fusion	65
Figure 2.39	Comparison of unimodal and proposed work	66
Figure 2.40	SIFT-based multimodal biometrics classification using RF & KNN	67
Figure 2.41	Face pre-processing	69
Figure 2.42	SIFT and LBP features	69
Figure 2.43	DWT-based feature extraction	69
Figure 2.44	Histogram for LBP features	70
Figure 2.45	Iris pre-processing	70
Figure 2.46	SIFT & DCT output	70
Figure 2.47	Ear pre-processing	70
Figure 2.48	MFRT features and LBP histogram	71
Figure 2.49	ROC for face (unimodal)	71
Figure 2.50	ROC for Iris (unimodal)	71
Figure 2.51	ROC for Ear (unimodal)	72
Figure 2.52	ROC for feature level fusion model	72
Figure 2.53	ROC for Multi-level fusion model	72
Figure 2.54	Comparison of multilevel fusion	72
Figure 2.55	Performance parameters of the proposed method	72

Figure 2.56	Performance analysis for various methods based on accuracy	73
Figure 3.1	Proposed oppositional grey wolf optimization-based biometric image recognition	77
Figure 3.2	Graphical representation for evaluation measures for our proposed methodology	88
Figure 3.3	Graphical representation for comparison of proposed and existing sensitivity measures	89
Figure 3.4	Graphical representation for comparison of proposed and existing specificity measures	90
Figure 3.5	Graphical representation for comparison of proposed and existing accuracy measures	90
Figure 3.6	Graphical representation of our proposed research ROC Curve	91
Figure 3.7	ROC curve analysis for existing research (a) GWO+L (b) GWO+LQ (c) GWO+Q (d) OGWO+L (e) OGWO+Q	92
Figure 4.1	Proposed efficient continuous biometric user authentication system for online examination	99
Figure 4.2	Proposed model for fusion of face, fingerprint, and keystroke for efficient continuous biometric user authentication system for online examination	103
Figure 4.3	Long short-term memory (LSTM) architecture	105
Figure 4.4	Sample image from ORL database	109
Figure 4.5	Sample image from YALE database	109
Figure 4.6	Sample image from FASSEG database	109
Figure 4.7	Sample data with the fused format of (a) User 1 face, fingerprint, and keystrokes (b) User 2 face, fingerprint, and keystrokes (c) User 3 face, fingerprint, and keystrokes	110
Figure 4.8	Accuracy of proposed and existing classifiers	113
Figure 4.9	Precision of proposed and existing classifiers	113
Figure 4.10	Recall of planned and obtainable classifiers	113
Figure 4.11	F-measure of planned and obtainable classifiers	114
Figure 4.12	Sensitivity of planned and obtainable classifiers	114
Figure 4.13	Specificity of planned and presented classifiers	114
Figure 4.14	False Positive Rate of planned and presented classifiers	116
Figure 4.15	Comparative analysis of planned and existing techniques for ORL database	117

Figure 4.16	Comparative analysis of planned and existing techniques for YALE database	118
Figure 4.17	Comparative analysis of planned and existing techniques for FASSEG Database	119
Figure 5.1	Unimodal recognition (a) face (b) ear (c) iris	123
Figure 5.2	Proposed multimodal architecture	124
Figure 5.3	Multimodal MVGG16 CNN model's framework (with feature level fusion)	125
Figure 5.4	Image augmentation	128
Figure 5.5	Various deep learning methodologies	129
Figure 5.6	Key applications of deep Learning	130
Figure 5.7	Convolution operation example with kernel size 3x3	131
Figure 5.8	An example of max pooling and average pooling	133
Figure 5.9	Different activation functions used in CNN	134
Figure 5.10	Transfer learning mechanism	136
Figure 5.11	VGG net architecture	137
Figure 5.12	Architecture of VGG19	138
Figure 5.13	Architecture of VGG16	138
Figure 5.14	MVGG16 Architecture	141
Figure 5.15	Sample images of face, ear, and iris traits from SDUMLA-HMT and IIT Delhi database	144
Figure 5.16	Training process accuracy and loss – unimodal face	147
Figure 5.17	Training process accuracy and loss – unimodal ear	147
Figure 5.18	Training process accuracy and loss – unimodal iris	148
Figure 5.19	Training process accuracy and loss – multimodal (face + ear + iris)	148
Figure 5.20	Recognition accuracy analysis for unimodal and multimodal using MVGG16	149
Figure 5.21	EER analysis for unimodal and multimodal using MVGG16	150
Figure 5.22	Recognition time analysis for unimodal and multimodal using MVGG16	150

Figure 6.1	Universal framework of biometric system	154
Figure 6.2	Typical biometric attributes	155
Figure 6.3	Conceptual framework of study	156
Figure 6.4	Methodology of research	162
Figure 6.5	Estimates of CFA model	168
Figure 6.6	ANN model used for process	172
Figure 6.7	RMSE statistics of training and testing	176

List of Tables

Table 1.1	Biometric trait's characteristics identifier	4
Table 1.2	Recognition rates of different method on PolyU database	23
Table 1.3	Challenges for future research	24
Table 2.1	Performance matrix of biometric systems	45
Table 2.2	Confusion matrix	48
Table 2.3	Comparison of unimodal methods with proposed multimodal system	58
Table 2.4	Accuracy obtained for different experiments	66
Table 2.5	Performance comparison with other methods	67
Table 3.1	Procedure for modified region growing	83
Table 3.2	The evaluation measures result of proposed and existing methods	89
Table 4.1	Performance comparison of classifier for different databases	115
Table 4.2	Comparative analysis of proposed and existing techniques for ORL database	117
Table 4.3	Comparative analysis of proposed and existing techniques for YALE database	117
Table 4.4	Comparative analysis of proposed and existing techniques for CBUA-OE database	119
Table 5.1	Hyperparameters of CNN	132
Table 5.2	VGG16 model	139
Table 5.3	Comparison between VGG16 and MVGG16	141
Table 5.4	Proposed MVGG16 network model	143
Table 5.5	MVGG16 training parameters for unimodal and multimodal	146
Table 5.6	Performance analysis of unimodal recognition using MVGG16	149
Table 5.7	Performance analysis of multimodal recognition using MVGG16	149
Table 5.8	Comparative performance analysis of multimodal biometric recognition	151
Table 6.1	Multimodal biometrics and execution measures	158
Table 6.2	Demographic profile	163
Table 6.3	Statistical distribution of constructs	165

Table 6.4	Goodness-of-Fit for MMB adoption	165
Table 6.5	Validity and reliability test of Standards	166
Table 6.6	Fornell-Larcker criterion for discriminant validity	167
Table 6.7	Summary of standard regression weights of constructs	169
Table 6.8	Estimation of the hypotheses and comparison	170
Table 6.9	RMSE value for training and testing data (N-530)	174
Table 6.10	Normalized and sensitivity analysis	177

Abbreviations

AF	Activation Function
AHE	Adaptive Histogram Equalization
ANN	Analysis of Neural Network
BT	Biometric Template
CNN	Convolutional Neural Network
CUBA- OE	Continuous User Biometric Authentication- Online E Proctoring
CFA	Confirmatory Factorial Analysis
DBM	Deep Belief Machine
DBN	Deep Belief Network
EER	Equal Error Rate
EFA	Exploratory Factorial Analysis
FAR	False Acceptance Rate
FMR	False Match Rate
FNMR	False Non-Match Rate
FRR	False Rejection Rate
FTE	Failure to Enroll Rate
GWOA	Grey Wolf Optimization Algorithm
LBP	Local Binary Pattern
LCNN	Looked up Convolutional Neural Network
LDA	Linear Discriminate Analysis
LGBP	Local Gradient Binary Pattern
LSTM	Long Short Term Memory
MBS	Multimodal Biometric System
MBM	Multimodal Biometric Mechanism
MKSVM	Multi Kernel Support Vector Machine
MWO	Modified Wolf Optimization
OGWA LQ	Optimum Grey Wolf Algorithm Lazy Querying
PCA	Principal Component Analysis

SSO	Slap Swarm Optimization
SEM	Statistical Equation Modeling
TPR	True Positive Rate
TNR	True Negative Rate
VGG	Visual Geometry Group
MVGG	Modified VGG
UBS	Unimodal Biometric System

Chapter 1

Introduction

1.1 Background

In today's era, biometrics is a field of great importance for mankind. As services and systems are transforming into digitized version, so access grant is digital. Earlier memory-based methods like passwords and pin number were used for authentication of users and still in many categories are under use. The challenge with these systems is comparatively easier spoofing or high probability of imposter attacks as well as difficulty of keeping track of multiple passwords for different services.

“Bio” is for life and “Metrics” is for measurement of something. Biometric authentication is based on user's physiological and behavioral characteristics which are unique such as Face, Iris, Ear, Finger Print, Retina, Finger Vein, Finger Knuckle, Palm print, Gait, Signature, Voice, key stroke pattern and ECG etc. These features are by default present with the user all the time, no need of remembering, easy to use at the time of a service access.

Natural attributes are unique and more reliable, hard to spoof and static for longer time span, so their use gives better and secure performance in authentication or verification model. Initially biometrics was used for high end security areas, border access, forensic science or for person verification by security/police teams. Down the line it got access into civilian life such as attendance of employees or students, financial transactions, welfare services, social security database like Aadhar and many more. With the

development of computer technology and semiconductor chips, applicability of the biometrics is increasing day by day. In general, its reach will be more universal in coming years.

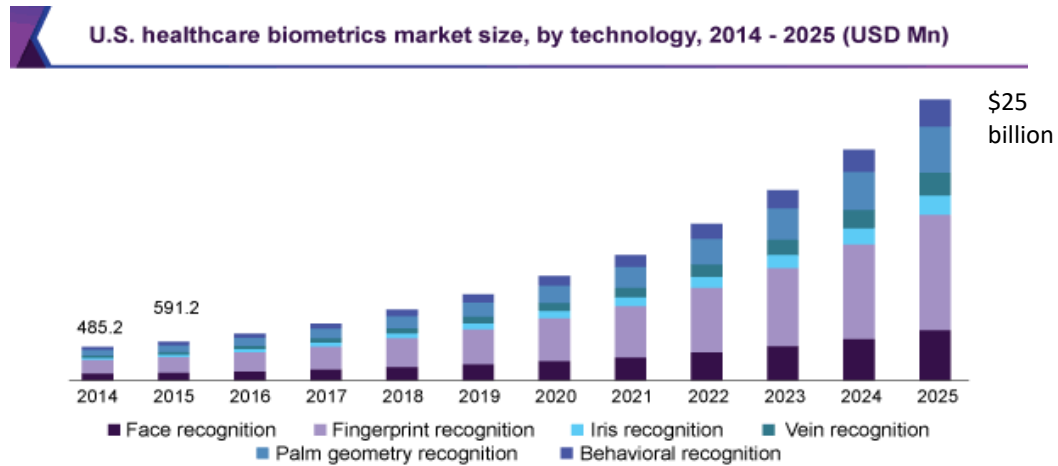


Fig. 1.1 Biometric healthcare market size data (US) from 2014-2025 [2]

Fig. 1.1 highlights the year-on-year increase in the market size of healthcare biometrics in the USA, 2014 onwards. Commercial significance is the main driving factor behind any technology's lifespan. The user's awareness and perception of the use of biometrics in important tasks such as banking has highlighted the bright future of this technology [1].

Biological traits like palm, hand, iris, face, vein, etc. are called physiological biometric traits. Whereas behavioral traits like the way we talk (speech), the way we walk (gait), keystroke patterns, and signature patterns fall under the behavioral category as shown in Fig. 1.2.

Biometric attributes or traits are having different features and with the advent of high-end capturing sensors researchers are able to use finger vein kind of internal traits as well. The selection of biometric traits depends upon the type of application and financial implications. Different traits can be compared on standard parameters like Universality, Permanence, Collectability, Acceptance, Circumvention, Distinctiveness, and Performance.

Universality is about the availability of the traits among all the users like face, iris, etc. Distinctiveness is about how unique it is from other users. Like iris has high uniqueness level than the face. A biometric trait has high “Permanence” if its properties don’t change for many years, the face will be different in old age and childhood, or the fingerprint may change over the years if someone using/doing work with hands. The trait should be stable for longer durations like 10 to 15 years so that fewer updates are required in the process. Collectability is how easy it is to collect data from users like face images can be taken easily whereas iris requires more effort. Acceptance is users’ willingness to give the data, so it is more for face or fingerprint than for iris, DNA, etc. Circumvention is how easy it is to spoof or fake identity. Iris is very tough to replicate or spoof, whereas face is a comparatively easier pathway for imposters. Performance is in a way the level of accuracy achieved during the verification or authentication process.

Different biometric traits can be compared on the basis of these characteristics and selected for an application. In Table 1.1, the level of these attributes is shown as High (H), Medium (M), and Low(L).

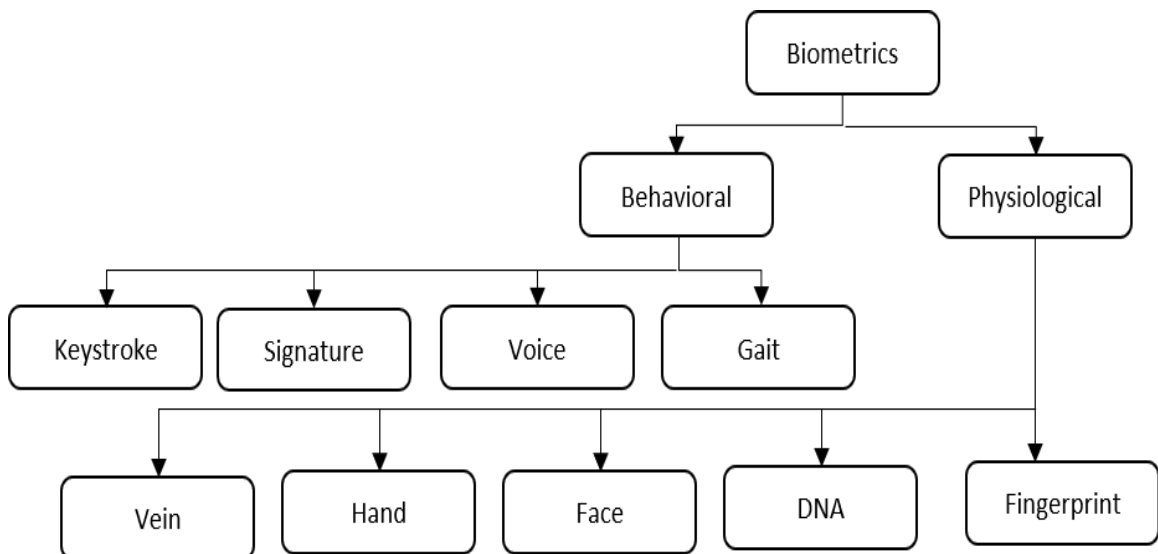


Fig. 1.2 Classification of biometric traits

Table 1.1 Biometric trait's characteristics identifier [12]

Biometric Trait	Universality	Distinctiveness	Permanence	Collectability	Performance	Acceptability	Circumvention
DNA	H	H	H	L	H	L	L
Ear	M	M	H	M	M	H	M
Face	H	L	M	H	L	H	H
Facial Thermogram	H	H	L	H	M	H	L
Fingerprint	M	H	H	M	H	M	M
Gait	M	L	L	H	L	H	M
Hand Geometry	M	M	M	H	M	M	M
Hand vein	M	M	M	M	M	M	L
Iris	H	H	H	M	H	L	L
Keystroke	L	L	L	M	L	M	M
Odor	H	H	H	L	L	M	L
Palm print	M	H	H	M	H	M	M
Retina	H	H	M	L	H	L	L
Signature	L	L	L	H	L	H	H
Voice	M	L	L	M	L	H	H

L- Low, M-Medium, H-High

In earlier days' the majority of the systems were based on faces or fingerprints. Later on, the retina, iris, and palm print came into the picture. System which is using only one type of trait is known as a unimodal biometric system whereas a system with more than one biometric trait is known as multimodal biometric system.

Whenever there is a dilemma in choosing which type of system to prefer, specific advantages associated with the multimodal biometric system make their deployment and development chances as compared to unimodal biometric systems. Also, for improving the recognition rate, a multimodal biometric system is preferred. Even the limitations of unimodal systems can be removed with the help of multimodal systems. Challenges associated with a unimodal system are that possess a lack of secrecy, non-universalities of samples, and even the extent of user's comfort. The multimodal biometric system generally merges two or more biometric traits like face, iris, ear, fingerprints, etc. and

they take inputs from single or multiple sensors for the measurement of two or more different biometric characteristics. For multimodal systems, it doesn't matter whether these systems take input from single or multiple sensors for the measurement of two or more than two biometric characteristics. For example, if a multimodal biometric system is developed using face and fingerprint recognition, then under such circumstances, the individual user verification can be processed by either of these modalities [2]. Even to simplify the customer experience and for the improvement of recognition rate, combining two or more modalities and developing multimodal biometric systems can be a feasible solution.

These multimodal biometric systems basically consist of modules such as (1) Sensor module, (2) Feature extraction module, (3) Matching module, and (4) Decision-making module. For the purpose of fusion, two or more biometric traits are run against two or more different algorithms to make a decision [3]. This technique is applicable for civil ID scenarios at a large scale in which authentication of multiple user identities needs to be processed.

There can be two types of biometric recognition operations: Verification and, Authentication. Biometric authentication is the process when the input image is matched against the entire range of stored templates 1: N, whereas in the verification process, it's a 1:1 match between the input image and claimed identity's store template [4].

1.1.1 Types of Multimodal Biometric Systems

A multimodal biometric system can be classified as below:

- a) Multi-algorithmic biometric systems: here two or more than two different algorithms are preferred for the processing of single biometric sample input obtained from a single sensor.
- b) Multi-instance biometric systems: here one or more than one sensor is used for the capturing of different samples obtained from the same biometric trait. For e.g., any system which captures images of multiple fingers comes under this category.

- c) Multi-sensorial biometric systems: here two or more sensors are used for the capturing purpose of the same instance of a biometric trait. After that, a single or combination of algorithms is preferred for the processing of captured samples. For e.g., if in any system, for capturing a facial image, two different cameras like an IR camera and visible light camera are preferred then such a system belongs to this category.

Before discussing the specific advantages of a multimodal biometric system as compared to a unimodal system, let's concentrate on the errors associated with image acquisition and matching of biometric traits. Failure to acquire (FTA) rate and failure to enroll (FTE) rate are the basic image acquisition errors. False nonmatch rate (FNMR), where the true individual's rejection is measured and false match rate (FMR) where the intruder is granted access, come under the category of matching errors. The specialty of these multimodal biometric systems is that they possess almost zero FTA, FTE, FNMR, and FMR rates. Sometimes few segments of the population are facing problems with one of their biometric traits so in that case, another biometric trait can be helpful to fulfill the criteria needed for verification purposes. For e.g., due to aging, old people suffer from fingerprint issues so in that case alternate modality can be used by the system [5]. Thus, in this manner due to the availability of other options with multimodal biometric systems, an almost zero FTE rate can be achieved. In this manner, data distortion can be significantly reduced by multimodal biometrics. Even the chances of spoofing are also very difficult for multimodal biometric systems and if any one modality got spoofed then the user can use his other modality for authentication purposes.

1.1.2 Advantages of Multimodal Biometrics

From the industry's perspective, the following are the few advantages of considering this multimodal biometric system.

-
- a) **Accuracy:** Technically speaking, the lesser the error, the higher will be the accuracy. Since multimodal biometrics possess almost zero FTA, FTE, FNMR, and FMR rates, so accuracy for such types of systems will be very high.
 - b) **Enhanced Security:** Historically, security is the necessity that gave rise to the invention of biometrics. So, this is one of the most important advantages associated with a multimodal biometric system. Here the system administrator for a multimodal biometric system has the option to choose the number of biometrics to be deployed as per the need of the security zone. It means that for a high-security area, he will be using three biometric traits but for a less secure area, even one or two credentials are sufficient to get the desired result. If any of the identifiers fail for any unknown reason, then also identification is possible with the help of other modalities in a multimodal biometric system.
 - c) **Increased and Reliable Recognition:** In a multimodal biometric system, since multiple biometric traits are deployed so higher recognition rate can be achieved. For e.g., the gaits or the pattern of movements of two people belonging to the same family or even different families can be the same so any unimodal system built on this gait modality is not reliable [6]. Thus, a multimodal system can use the other modality along with gait for proper identification purposes as it is quite impossible for any two persons possessing the same two or more modalities.
 - d) **Vulnerability:** Spoofing is one of the biggest threats which is associated with any authentication system. Both unimodal as well as multimodal systems are vulnerable to spoofing. The phenomenon of spoofing takes place when some unauthorized person invades the system. Here the potential threat is due to the possibility of artificial fingers which are cloned with plastic molds to be accepted by the system for the verification stage with a probability of 68 to 100%. To avoid spoofing, alternate hardware devices like finger vein readers or smart fingerprint readers with liveness detection can be deployed along with the regular systems. Now this liveness means the ability of a system to make a judgment between a

living and a fake sample and this is accomplished by measuring biometric features like temperature, humidity, pulse, blood flow, etc.

- e) User Acceptance: For larger deployment purposes, multimodal systems are preferred as compared to unimodal systems. Even those big databases of the large-scale population are also turning to multimodal biometric systems. So, we can say user acceptance for large biometric data is done by using multimodal biometric systems.

1.1.3 Biometric Fusion

The mechanism of Fusion plays a crucial stage in multimodal biometric system development. More than one decision channels are present due to multiple modalities associated with a multimodal biometric system. So biometric fusion plays a crucial role in combining the classification outcome which is coming from the individual biometric channel. The role of fusion is to combine measurements received from many biometric attributes to increase the strength and decrease the weakness associated with individual measurements. Accuracy, Universality, Robustness, Efficiency, and Applicability can also be addressed with the help of the fusion mechanism [7]. There are various levels of fusion which can be classified as sensor level, feature level, matching score level, and decision level fusion. The block diagram below of Fig1.3 depicts the fusion levels present in a multimodal biometric system.

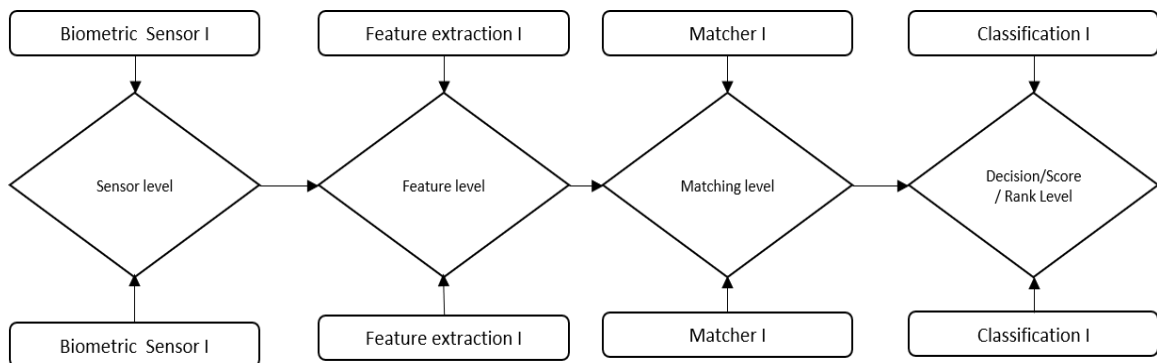


Fig. 1.3 Fusion levels in multimodal biometric systems

Following are the different fusion levels:

-
- a) Sensor level fusion: A merged biometric trait is created by fusing various biometric traits which are coming from different sensors like iris scanner, video camera, fingerprint scanner, etc.
 - b) Feature level fusion: Separate feature vectors are obtained from individual biometric traits after the processing of signals that are coming from different biometric channels. A specific fusion algorithm is used for combining these feature vectors to form a composite feature vector. For the selection of desired useful features, a few reduction techniques need to be applied. Feature level fusion is applied for getting better recognition results as these features are a very rich source of information on biometric traits as compared to other scores. In feature-level fusion, the amount of accuracy is dependent on the compatibility among the different biometric traits used in the multimodal biometric system, the higher the compatibility, the higher will be the accuracy as such.
 - c) Matching (score) level fusion: The matching score is obtained individually as the processing of feature vectors is performed separately. For classification purposes, after getting the individual matching score, a composite matching score is obtained by using many different techniques such as logistic regression, mean fusion, Bayes rule, highest rank, etc. Normalization of scores is performed by using various techniques such as min-max, z-score, piecewise linear, etc.
 - d) Decision level fusion: Pre-classification of each biometric trait has been performed separately. Firstly, individual biometric traits are captured and then the extraction of features is performed. These traits are further categorized as either accept or reject which is based on these extracted features. After that for the final classification purpose, output from different modalities is combined.

1.2 Motivation

Multimodal biometric systems are more efficient and reliable as compared to unimodal biometric systems and this is the major impetus for research and development in this direction. Even the global forecast for the biometric system market is expected to a growth rate from USD 36.6 billion in 2020 to USD 68.6 billion by 2025. It is basically estimated to grow at a Compound Annual Growth Rate (CAGR) of 13.4%. Whereas the Indian biometric market size reached a value of almost USD 1.95 billion in the year 2020. It is further expected to grow at a CAGR of 14.5% between 2021 & 2026 to reach a size of almost USD 4.4 Billion by 2026. The major application area for the biometric system will be surveillance and security, consumer electronic devices, and most importantly automotive applications. As multimodal biometrics is the most advanced industrial topic, so major companies working in this domain and promoting biometric growth are Thales (France), NEC(Japan), Fujitsu (Japan), etc.

Many researchers generally preferred two biometric features for their experimental purpose. The more specific applications like UID card or Aadhar card system prefers face, finger, and iris features for the purpose of authentication of the individual. This motivates the thesis in the direction of a detailed study of different multimodal biometric systems with more focus on feature-level fusion as not much research work is available in this direction of fusion and real-time implementation of feature fusion is a big challenge.

During COVID-19 Pandemic, remote proctoring of the employee or human being has evolved as a big challenge for the information retrieval process. Compared to memory-based system access authentication, they are becoming outdated and less preferred for live applications, especially where data security and customer privacy are crucial. Multi-modal authentication has outperformed the unimodal process with high accuracy and improved security in the user authentication field. Multi-modal biometric verification includes user attributes such as keystrokes, iris, speech, face, etc.

As per the literature, it is verified that the feature level fusion method is better than other fusion methods as it uses original information of input signals in the original form. There are many challenges associated with the feature level fusion so other methods such as score and decision level are more conspicuous. The size of feature sets is significantly larger in feature fusion than in other levels of fusion, so dimensionality, complexity, and increased computation are major concerns. With such challenges, feature-level fusion provides a high level of performance in the authentication process. Hence multimodal biometric fusion-based authentication seems to be a better approach but due to implementation challenges and associated costs and complexities, it is not realized at full scale. Real-time execution of multi-modal biometric fusion-based algorithms may assist in live tracking for the compatible use case, and it can also be deployed in several applications.

Score and decision level-based fusion etc. schemes [8] have been popular but in the era of technological development, spoofing, attacks, and cost-time effective real-time implementation is a growing challenge. Larger-size feature sets that are not correlated at all create another challenge for the ensemble or fusion process. Reported fusion schemes are majorly based on linear approaches like mean, max, sum, averages, etc. There is a great scope to develop optimum fusion schemes for multimodal biometrics particularly in the category of feature-level fusion.

Next-generation platforms-based implementation attracts the use of CNN-based network modeling for the biometric authentication system. Portable devices with limited memory and power offer another challenge for the scientific community. The fast-changing world has a different class of users who belong to different generations, communities, and segments which also highlights the non-universality of any technology. The adaptation of new technology has its own challenges and its viability and general perceptions can be analytically explored before commercially launching it.

1.3 Literature Survey/Related Work

1.3.1 Introduction

Biometric technology is frequently used in multiple areas including the work entrance of an organization to the human authentication/verification among the financial exchange [9]. It is also prominent in pattern recognition and machine learning application design [10]. It is the backbone of the person recognition process, and biometric traits such as the hand vein, palm print, face, fingerprint, and speech are being applied to identify an individual [11]. It offers a very convenient and secure mode of identification and verification solutions. It is used in several applications like computer network login, electronic data security, e-commerce, Internet access, ATM, credit card, physical access control, cellular phone, PDA, medical records management, and distance learning [12]. In this technology, biometric systems rely on particular data about unique biological traits to unique work effectively. Uni-modal biometric systems that use only one biometric trait for recognition often affect issues such as biometric data variation, lack of distinctiveness, low recognition accuracy, and spoof attacks. To discover the problem, multimodal biometric systems are used [13].

Multimodal biometric systems, consolidating multiple traits, address limitations of uni-modal biometric systems in matching accuracy, spoofing difficulty, universality, feasibility, etc. [14]. When compared to single-modal biometric systems, a multimodal biometric system increases recognition accuracy, security, and system reliability. Multimodal biometric technologies are used to improve identification and are carried out by allowing lots of biometric features, which enables us to avoid unauthorized access [15].

Gao et al. [16] have proposed that traditional CCA and 2D-CCA algorithms were unsupervised multiple-feature extraction methods. Hence, inaugurating the supervised information of samples into those methods should be capable to promote classification performance. In that paper, a new method was suggested to carry out the multiple feature

extraction for classification, called two-dimensional supervised canonical correlation analysis (2D-SCCA), in which the supervised information was added to the criterion function. Then, by testing the relationship among the GCCA and 2D-SCCA, another feature extraction method named as multiple-rank supervised canonical correlation analysis (MSCCA) was also improved. Vary from 2D-SCCA, in MSCCA k pairs left transforms, and k pairs right transforms were sought to optimize the correlation. The convergence behavior and computational complexity of the algorithms were tested. Experimental results on real-world databases established the viability of the formulation, they also demonstrated that the classification results of our methods were higher than the others and the computing time was competitive. In that manner, the suggested methods proved to be competitive multiple-feature extraction and classification methods. As such, the two methods might well help to develop image recognition tasks, which were required for many advanced expert and intelligent systems.

The multimodal biometrics established on feature-level fusion was a crucial topic in the personal identification research community. Yang et al. [17] have described a new fingerprint-vein-based biometric method that was suggested for making a finger more universal in biometrics. The fingerprint and finger-vein features were first exploited and extracted by applying a unified Gabor filter framework. Then, a novel supervised local-preserving canonical correlation analysis method (SLPCCAM) was suggested to generate fingerprint-vein feature vectors (FPVFs) in feature-level fusion. An established on FPVFs, the nearest neighborhood classifier was utilized for personal identification finally. Experimental results demonstrated that the suggested approach had a high capability in fingerprint-vein-based personal recognition as well as multimodal feature-level fusion.

Biometric characteristics, like finger knuckles and finger veins, are unique and secure. Veluchamy et al. [18] have suggested a multimodal biometric system by equating the finger knuckle and finger vein images at feature-level fusion using fractional firefly (FFF) optimization. Initially, the features were extracted from the finger knuckle and

finger vein images by applying repeated line tracking methods. Then, a newly improved method of feature-level fusion applying FFF optimization was applied. That method was applied to determine the optimal weight score to fuse the extracted feature sets of finger knuckle and finger vein images. Thus, the recognition was carried out by the fused feature set applying layered k-SVM (k-support vector machine) which was newly improved by equating the layered SVM classifier and k-neural network classifier. The experimental results were estimated and the performance was analyzed with false acceptance ratio, false rejection ratio, and accuracy. The outcome of the suggested FFF optimization system found a higher accuracy of 96%.

Multi-biometric systems were being increasingly deployed in several large-scale biometric applications (e.g., FBI-IAFIS, UIDAI system in India) since they had several advantages such as lower error rates and larger population coverage equated to uni-biometric systems, as reported by Abhishek et al. [19]. However, multi-biometric systems needed storage of multiple biometric templates (e.g., fingerprint, iris, and face) for each user, which results in increased risk to user privacy and system security. One method to protect individual templates was to store only the secure sketch generated from the representing template by applying a biometric cryptosystem. That required storage of multiple sketches. In that paper, they suggested a feature-level fusion framework to simultaneously protect multiple templates of a user as a single secure sketch. Their main contributions include:

- 1) Practical implementation of the suggested feature-level fusion framework using two well-known biometric cryptosystems, namely, fuzzy vault and fuzzy commitment, and
- 2) Detailed analysis of the trade-off among matching accuracy and security in the suggested multi-biometric cryptosystems established on two various databases (one real and one virtual multimodal database), each containing the three most popular biometric modalities, namely, fingerprint, iris, and face. Experimental results demonstrated that both the multi-biometric cryptosystems proposed here

had higher security and matching performance equated to their uni-biometric counterparts.

Equated to uni-biometric systems, multi-biometric systems, which fused multiple biometric features, could develop recognition accuracy and security. However, due to challenging issues such as feature fusion and biometric template security, there was little research on cancellable multi-biometric systems. Wencheng et al. [20] suggested a fingerprint and finger-vein-based cancellable multi-biometric system, which rendered template protection and revocability. The suggested multi-biometric system equated the minutiae-based fingerprint feature set and the image-based finger-vein feature set. They improved a feature-level fusion strategy with three fusion options. Matching performance and security strength applying those different fusion options were thoroughly evaluated and analyzed. Moreover, equated with the original partial discrete Fourier transform (P-DFT), the security of the suggested multi-biometric system was strengthened, thanks to the raised partial discrete Fourier transform (EP-DFT) based non-invertible transformation.

1.3.2 Remote Biometric Recognition

Remote authentication for online application/ deployment is important in the new age of the internet. Continuous tracking of web users' interests, navigational actions, and preferences has gained crucial weightage due to the objectives of organizations and companies as reported by Slanzi et al. [21]. Traditionally that field has been analyzed from the Web Mining perspective, particularly through the Web Usage Mining (WUM) concept, which comprises the application of machine learning techniques over data originating in the Web (Web data) for automatic extraction of behavioral patterns from Web users. WUM makes use of data sources that approximate users' behavior, such as weblogs or click streams among others; however, those sources imply a considerable degree of subjectivity to interpret. For that reason, the application of biometric tools with the possibility of measuring actual responses to the stimuli introduced via websites has become of interest in this field. Instead of doing separate analyses, Information Fusion

(IF) tries to develop the results by developing efficient methods for transforming information from different sources into a single representation, which then could be applied to guide biometric data fusion to complement the traditional WUM studies and obtained better results. That paper introduces a survey of Biometric Information Fusion applied to the WUM field, by first defining WUM and its major applications, later explaining how the Biometric Information Fusion could be applied and finally reviewing several analyses that applied the concept to WUM.

Accuracy and usability were the two most crucial issues for a multibiometric system. Most of the multibiometric systems were established on matching scores or features of multiple biometric traits. However, plenty of identity information was lost in the procedure of extracting scores or features from captured multimodal biometric data, and the loss of information stops the accuracy and usability of the multi-biometric system from reaching a higher level as observed by Miao et al. [22]. It was believed that matching scores could recover some identity information, which had not been applied in previous fusion work. That study suggested a framework of a bin-based classifier method for the fusion of multibiometrics, to handle that problem. The suggested method embeds matching scores into a higher-dimensional space by the bin-based classifier, and rich identity information, which was hidden in matching scores, was recovered in this new space. The recovered information was sufficient to distinguish impostors from genuine users more accurately.

A multi-attribute estimation with a convolution tensor correlation fusion network had been explained by Duan et al. [23]. The system includes feature extraction, correlation excavation among facial attribute features, score fusion, and multi-attribute prediction. Subnetworks (Age-Net, Gender-Net, Race-Net, and Smile-Net) were used to extract corresponding features while Main-Net extracts feature not only from an input image but also from corresponding pooling layers of subnetworks. Dynamic tensor canonical correlation analysis (DTCCA) was proposed to explore the correlation of different targets' features in the F7 layers. Then, for binary classifications of gender, race, and

smile, corresponding robust decisions are achieved by fusing the results of subnetworks with those of TCFN while for age prediction, facial image into one of the age groups, and then ELM regressor performs the final age estimation. Duan et al. [24] explained an ensemble CNN2ELM-based age estimation. This system consists of three-layer namely, feature extraction and fusion, age grouping via an ELM classifier, and age estimation via an ELM regressor. Similarly, in [25], Duan et al. explained a hybrid deep learning CNN-ELM-based age and gender classification. CNN was used to extract the features from the input images while ELM classifies the intermediate results.

Peng et al. [26] proposed wearing glasses with a steady negative solidified Glass Guard. Glass Guard isolates the owner and an impostor with social biometrics from six sorts of touch movements and voice orders that are generally open customary customer joint efforts. Using data accumulated from 32 Google Glass customers, after 3.5 customer events, Glass Guard has a 99% ID rate and 0.5% counterfeit alarm speed, with a typical of 3.5 customer events and all customer events. As demonstrated by the program's five ordinary circumstances, the system area rate is over 93%, and the counterfeit alert rate is below 3% after five customer events.

Keystroke Biometrics is an important feature in the digital era. Ali et al. [27] have presented research on keystroke biometrics (KB), and reviews of keystroke biometric structures can be used as an early phase for newcomers to the field. Furthermore, a couple of references are delivered from time to time, differentiating them subject to different procedures and strategies used in other examinations. It fills in as a wellspring of viewpoints for various examiners to evaluate their work to choose the orientation of future investigation.

The approval of understudies with text-based and picture-based troubles has been investigated by Ullah et al. [28]. Their assessment showed that 70 online individuals from nine countries had completed a five-week online route and were standing up to the peril of mercilessness. Excellent review prompts a development in substance requesting ($p < 0.01$), ii) another doable facts section strategy than text requests ($p < 0.01$). Re-

enactment Abuse diversion is used to gauge the impact of database bundles at different levels.

Ullah et al. [29] aimed to brace understudy's characters with dynamic profile questions. In five weeks, 31 online individuals from five countries checked out the examination. The delayed consequences of their usage and security examination were dissipated. Dynamic profile necessities are more capable than text and picture-based essentials ($p < 0.01$). Fake individual abuse is reflected in messages and telephones. Regardless, the email deception attack was not successful, and understudies had the choice to respond persistently using an outcast test situation, achieving a 93% correct answer in the online test. Phone information and pestering response time are astounding from understudy time ($p < 0.01$).

A continuous authentication (CA) system that consistently screens the customer before marking the PC was proposed by Prakash [30]. The CA system shields intruders from starting the structure. It dormant yields the form without interrupting the customer task measure. Here, the score-level mixing gel is proposed for mass improvement measure and torpid glass fire-based ID measure. The essential goal of their specific method is to join customer biometric incorporates and achieve the best-delayed consequences of steady customer confirmation.

By considering the elements of four unique sensors, Yang et al. [31] make a restrictive model using one-class SVM (OCSVM) and disengagement woodland (iForest) and figure out the precision of each sort utilizing this model. Afterward, they determined each kind of certainty level using the Bayesian hypothesis. They accomplished the exactness of the sensor work range through a creative anticipated framework.

SNUSE, a multi-customer secure system approach to manage re-enlisting on BIA structures was developed by Nunes et al. [32]. They train SNUSE to reflect its inactive limit and evaluate its sufficiency and precision using two biometric methodologies, one-of-a-kind imprint, and iris separating. The biometric thought of the SNUSE key doesn't

impact the precision of the extraction program, and their test results are indicated each second by a supportive module using a standard PC Laboure. Executing their model achieves acknowledgment accuracy of more than 90% Genuine Acceptance Rate (GAR) and fewer than 5% False Acceptance Rate (FAR).

A lead biometric count program-level security for customer affirmation was proposed by Chatterjee et al. [33]. Regardless, they drove risk assessments for various exercises. The keystroke ensures complete security and sets the standard multimode modes close to the beginning of the natural gathering. An appraisal of their proposed affirmation instrument has seemed dependent on fake affirmation rate of false acceptance rate (FAR), fake excusal rate or false rejection rate (FRR), and comparable slip-up rate or equal error rate (EER). The standing multi-particular check achieved an FRR of 1.2%, FAR of 0.89%, and EER of 1.04% with J48 course of action computation.

Wang et al. [34] represented BodyPIN, an unending customer confirmation system that fathoms far-off correspondences using Wi-Fi things. In the wake of examining the system work, the structure can follow the character of the current customer. If the check crashes and burns, induction to system security will be set up. Since the customer describes different ascribes from customer to customer, it is hard to use the latest Rich Wireless User ID direct on the BodyPIN. The system interferes with the customer's activities, which is uncommonly problematic and gravely orchestrated. In this examination, the makers developed a BodyPIN of a bioelectromagnetic field to evaluate the effect of two pi-Wi-Fi signals on the BodyPIN and to get a customer-choosing component.

Kiyani et al. [35] have proposed keystroke components: a predictable customer affirmation system with lead biometric structures. Every limit offers another technique for recognizing the customer, which chooses the credibility of the current customer subject to the authentic thought of every limit. The makers developed a two-stage system with alerts and end limits using the novel study of two-passage bundle learning and the constant uncertainty model (R-RCM). Their specific computation masterminds each activity subject to the board's fire glass capacity, which uses R-RCM hyper-limits to

assess the certified trustworthiness of the customer's genuine lead. The system chooses if the customer can use the structure based on assessing the new trust and the endpoint.

1.3.3 CNN-based Biometric Recognition

Convolution Neural networks based has supported the robust implementation of biometric systems. Marieo et al. [36] proposed a Siamese convolution neural network to learn the signature of the motion patterns from users and achieved an accuracy of 97 %. It has demonstrated that changes in motion patterns may help in detecting the unauthorized use of mobile devices. The proposed class SVM model improved the verification rate at the same sampling frequency and window size. It was also observed that with lower frequency, the accuracy rate was higher irrespective of window size.

A hybrid deep learning CNN-ELM has been proposed by Duan et al. [25] where automatic age and gender classification was obtained via a convolutional neural network and extreme learning machine (ELM). CNN was used to extract the features from images while ELM classifies the intermediate results. Experiments were conducted on MORPH-II and Adience Benchmark datasets to verify the proposed hybrid structure. The process is shown in Fig. 1.5 [25]. The age classification with 0.5 and 0.7 dropout rates were 51.4 % and 52.3% respectively. In the case of gender Classification accuracy was 87.3% and 88.2% with similar dropout rates. In extended work, authors used an ELM regressor for decision-making after the ELM age classifier with an accuracy of 0.6649. Here fusion of Gender-Net, and Race-Net outputs with Age-Net have improved the overall performance [26].

CNN-based feature extraction methods using multimodal fusion for high-security applications were used by P. Shende et al. [37]. Here network was of two convolutional layers, two Relu and, two Maxpooling layers with ten hidden layers in fully connected layers. Researchers performed an experiment for Fingerprint, face, and palm vein with 4500 images of each trait and obtained accuracy above 90%.

Over the year scientist have tried different methods to improve the performance with lower complexity levels. Multilevel feature abstraction is a step forward where multiple streams of modality-specific Convolution neural networks are optimized at multiple levels [38]. Features obtained at different levels of modality-oriented CNN are actually input at different levels of abstract representations. Multi-level abstract representations result in a reduction of network parameters, and it improves the overall accuracy. Multi-abstract fusion produced 99.34%, and 99.91 % accuracy with Bio Cop and BIOMDATA databases respectively.

Multimodal fusion is the future of biometric recognition and researchers have made significant progress in the field. Peng et al. [39] have used finger knuckle, finger shape, finger vein, and fingerprint fusion for biometric authentication with score fusion based on triangular norms. The distance between genuine and imposter score distribution was larger than other state-of-the-art approaches.

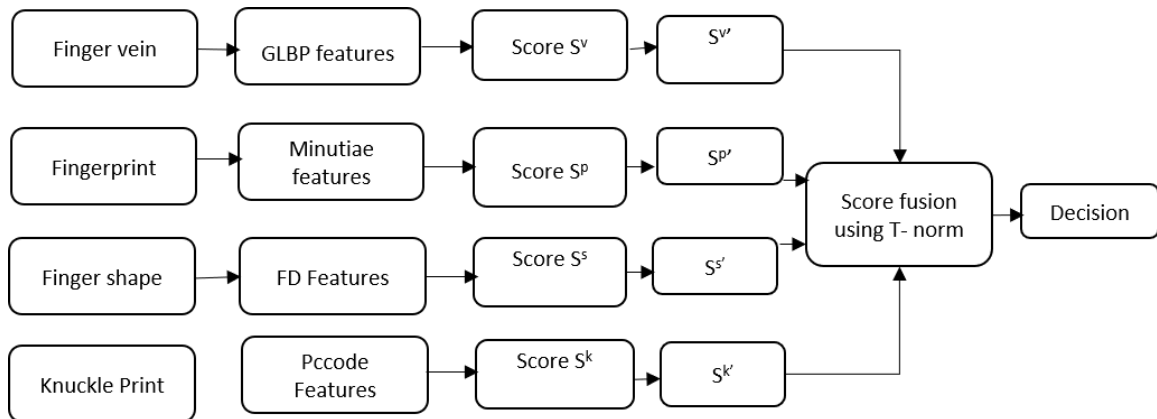


Fig. 1.4 Finger Multimodal biometric schemes using t-norm [39]

Another aspect is to use handcrafted features for a multimodal biometric fusion of features obtained in the unconstrained environment. E. Hansley et al. [40] presented an ear recognition framework that outperforms standard systems with different databases. A Two-stage landmark detector was designed for untrained scenarios and generated results were used to perform geometric image normalization. Authors could realize that

handcrafted features and CNN are complementary and normalization is critical for performance improvement in ear recognition under unconstrained environments.

H Sinha et al. [41] successfully improved the security and protection in ear biometrics via CNN as a feature extractor and support vector machine (SVM) for classification. The joint CNN-SVM framework is used for mapping ear images to random base- n codes. The SHA-3 algorithm is used for generating secure EAR templates. On the web ear database, it demonstrates a 12.52% average equal error rate without any preprocessing. This method was found too strong against hill-climbing attacks, FAR attacks, and linkage attacks.

Further Sinha et al. [42] presented an innovative strategy for a CNN-based neural architecture that was capable of learning sparse representation imitating the receptive neurons in the primary auditory cortex in mammals. It was assessed for audio classification tasks on standard benchmark datasets such as Google speech commands (GSCv1 & GSCv2), UrbanSound8k (US8K). This architecture was called Braided Convolutional Neural Network (BCNN) and it could obtain 97.15 %, 95%, and 91.9% average recognition accuracy on GSCv1 & GSCv2, and US8K respectively.

An efficient and real-time multimodal biometric system was proposed by Al-waisy et al. [43] in which face, left & right iris images were used as input traits. A deep belief network (DBN) was used for face feature extraction and IrisConvnet for Iris image feature extraction. Score and rank fusion schemes were utilized for an accuracy rate above 97%.

Wan et al. [44] fused finger vein and finger shape using a near-infrared camera sensor and a score-level approach for authentication. The process was repeated with VGG16, Resnet 50, and Resnet 101 on a desktop computer environment (Intel core i7-6700 CPU, 3.4 GHz (4 cores) with 32 GB RAM, and NVIDIA GeForce GTX Titan X (3072 CUDA cores). The algorithm was implemented using Caffe Framework and Microsoft Visual Studio. The SDU-DB and PolyU-DB dataset were used for experiments. The EER range was between 1.706% to 4.36% for different networks with different fusion schemes.

Palm print is one of the less used biometric features in comparison to fingerprint and face but has very good collectability and uniqueness. Q. Sun et al. [45] proposed a deep learning model based on CNN-F (fast) architecture and exactly evaluate the convolutional features from different layers in the network for both identification and verification. The experiment results in an EER of 0.25% and an identification rate of 100% on the PolyU database which highlights the effectiveness of CNN features.

O. M Parkhi et al. [46] of Visual Geometry Group presented face recognition with novel goals such as (i) How a very large-scale dataset can be assembled by a combination of automation and human in the loop & trade-off between data purity and time, (ii) to rescue complexities of deep network training and face recognition to present methods and procedures to achieve a comparable state of the art results on the standard LFW & YTF face benchmarks. The authors reported 98.95 % & 97.3% accuracies on LFW and YTF databases respectively.

Table 1.2 Recognition rates of different methods on PolyU database

Methods	Total Samples	Different palms	Train samples	Recognition rate
RLOC	7752	386	1	98.37
Contourlet Transform	7752	386	3	88.91
2D-DOST	900	150	3	97.29
BDCT	2000	100	4	98.93
KPCA+GWR	3860	386	4	99.69
OWE	2000	100	5	98.90
2D Gabor wavelets +PCNN	3860	386	5	97.37
Deep CNN-F	3855	386	3	100

Information fusion is the core part of multimodal biometric systems. Fusion methods shall impact the performance level of the entire model. M Singh et al. [47] have presented state-of-the-art work in a detailed overview of biometric fusion. The authors covered all the progress made to date in the fusion arena as well as continuous authentication, child biometrics, attacks, cryptosystems, designing adaptive and dynamic fusion systems, and multimodal solutions for personal devices like smartphones, health gadgets, vehicle access, and system access.

Finger vein verification using a Siamese CNN which is a kind of lightweight CNN was proposed by Su Tang et al. [48]. Heavy image augmentation is used to cope with the shortage of datasets. A Siamese structure combined with a contrastive loss function for training the CNN was prepared. Later looking at deploying the above CNN on embedded devices a lighter version was also prepared by the authors. It was 1/6th of the original size and ERR 0.08, 0.11, and 0.75 were obtained with SDMULA-HMT FV-USM and MMCBNU_6000 datasets.

Table 1.3 Challenges for future research

Sr. No.	Key challenges
I.	Portability of the Multibiometric Solutions
II.	Designing Adaptive and Dynamic Fusion Systems
III.	Multibiometric Security and Privacy
IV.	Resolving conflicts between information sources
V.	Predicting scalability of Multibiometric systems
VI.	Sensor Configuration in Multibiometric Systems
VII.	Multimodal Solutions for Compact Personal Devices

K. Gunasekaran et al. [49] proposed deep multimodal biometric recognition using deep contour let derivative weighted rank (DCD-WR) fusion with a human face, fingerprint, and iris images. Contour let transform reduced the dimensionality of the feature vector with a great margin. A novel rank-level fusion and deep learning-based template matching method were used to make the final decision.

Deep cascade score level fusion for unconstrained ear recognition was explored by Umit K. et al. [50]. The proposed method represents the first automated fusion learning approach and is also compatible with parallel processing. The authors evaluated Score net using the Unconstrained Ear Recognition Challenge Database which is one of the toughest datasets to work upon.

So far, we could observe that CNN has reduced the feature extraction problem to a great extent. Almost all kind of biometric traits have been used with different structure of CNN and it could generate better results. Having multimodal biometric network based on CNN shall give us ease of implementation for real time solutions and particularly for portable

devices. Here in this research study, we have followed a systematic path starting from basic uni-modal biometrics to optimized feature fusion based multimodal, its real-time application in continuous authentication, CNN-based experiments and implementation, and finally user perception about all these developments.

In the following section, we are presenting work done in the area of user perception of biometric technology and modeling feasibility. Specially Gokul Kumari et al. [1] have done a detailed survey that could highlight user awareness about the use of multimodal biometrics in financial transactions.

1.3.4 Study of Work based on User Perception of Biometric applications

A systematic review of the literature is an effective method for locating, coordinating, and deducing facts relevant to a particular research question rationally, descriptive, dependable, and accurately [51],[52]. Porwik et al. [53] studied Multimodal biometrics is based on a fusion of different physical and behavioral traits like face, ear, iris, gait, keystroke, voice, etc. Anil et al. [54] presented a detailed study on types of fusion methods for MMB. They have highlighted various scoring fusions such as min, max, average, mode, and others. Feature and sensor level come under the first category score, whereas decision & rank level fusion is the second category.

Recent studies [55],[56], [57] advocated a combination of unimodal and multimodal biometrics features that are helpful for accuracy in text and voice output. El-fishway [58] used unimodal score and multimodal score to compare performance for different conditions. They found MMB has the advantage of a more secure and reliable process at the cost of increased complexity.

In the words of Guan et al. [59]; Heracleous and College et al.[60]; & Stylios et al.[61] multimodal biometrics outperformed unimodal in multiple experiments with different feature extraction algorithms. However, authors have used a different combination of face, ear, iris, fingerprint, palm print, and palm knuckle to confine the results. Jagadiswary and Saraswady [62] stated fused fingerprint, retina, and finger vein at feature

level and used RSA key generation to achieve FAR (false acceptance rate) .01, 95.3% GAR (genuine acceptance rate). Modified RSA (Rivest, Shamir and Adleman) [63] is the contribution made by the authors to improve GAR, FAR from 90% & 2.06 %, respectively. Individual unimodal GAR FAR compared with the MBM results. RSA with fingerprint GAR of 80.2% and FAR of 3.25%, RSA with retina GAR of 84.2% and FAR of 2.2%, RSA with finger vein GAR of 87.6% and FAR of 0.52% were obtained. Choudhary et al. [64] reported fast and robust biometrics system design using Ear trait. Later it was combined with other characteristics and compared in precision, recall rate, accuracy, ease of use.

Srinivasu et al. [65] studied the computerized progression of classifying skin disease using deep learning MobileNet V2 and specified that the projected system could assist general practitioners in diagnose skin conditions proficiently and reduce complications and indisposition. In a similar line of research (Panigrahi, Borah, Bhoi, Ijaz, Pramanik, Kumar, et al.) [66] developed a consolidated tree construction (CTC) algorithm to create a sound sample from a high-class imbalanced dataset at the detector's pre-processing phase. The outcomes delineated the accuracy of 99.96%, reflecting the CICIDS2017 dataset and the NSL-KDD data frame using thirty-four topographies.

Scholarly evidence Ajeenkya D. et al. [67]; Liang and Li et al. [68]; Nader et al. [69]; Oloyede et al. [70] a robust and efficient ear-based biometric system using AdaBoost-based ear detection, local features extraction, and stereo matching-based recognition algorithm proposed in this work outperforms other unimodal and multimodal results in similar categories. Furthermore, peripolar geometry and rectification, fast Stereo mapping utilized to compare probe images with ear template images Sinha et al. [71]. Sireesha et al. [72] outlined IITD, UND-F, USTB database used for experimental purpose and comparison. The results were compared with other methods where SVM, plane Adaboost, M.N.N. and Proposed method generated accuracy of 96.20%, 94.40%, 96.40% and 98.50%, respectively.

Stylios et al. [73] highlighted the feature level fusion-based method, where correlation analysis is done in class and intraclass feature sets. For feature fusion, discrete correlation analysis is adopted to get the more pairwise correlation. DCA also reduces between class correlation to create a clear boundary.

Sinha et al. [71] outlined the detailed status of contemporary biometric system design issues, prospects, commercial uses and market demand and supplied part of unimodal vs multimodal biometrics systems. A significant portion of the current market is occupied by a unimodal system which is getting tough competition from multimodal systems robustness. Moreover, users are changing priorities from user-friendly ness to security and fraud protection.

Some previous studies viz. El-fishawy et al. [58]; Teh, Zhang et al. [74], Venkatraman et al.[75]; studied various applications in school attendance systems, public place access, office/institute access, airports, etc. labs, and sensitive regions like border crossing points are static points of use. At the same time, online commercial benefits like purchase, banking transactions, stock trading, shopping, and non-commercial online parts such as necessary access like defense lab systems or high-security areas access are studied in detail [76].

Certain authors Abomhara et al. [77]; Joshi et al. [78] highlighted the increasing impact of artificial intelligence, deep learning impact in biometric system design. Using AI-based algorithm testing and deployment part is becoming smooth. Moreover, smart devices like phone and other tablets are now using multimodal biometric algorithms to enhance user experience.

A brief review on the research dimension of multimodal biometrics was conducted by Dwivedi et al. [79]. It was reported that MBS could be classified in two major areas as synchronous and asynchronous systems. In the first case, two or more biometric systems are under the same authorization systems. In contrast, in the case of asynchronous

category two, the biometrics process is used one after another. Further, subdivisions are serial (cascade), parallel, and hierarchical modes.

Alsadoon, et al. [80] categorized the fusion schemes in two categories like fusion before and after matching: feature level, score level, decision, rank level, and hybrid levels. Soft biometrics is another vital aspect of a user-friendly system where age, gender, height, eye color, skin color, hair color kind of input are added with other traits. It is convenient for the user and enhances perception about ease of service. The template matching process is also vital in an extensive database, so speed up the searching indexing and cloud technologies are used.

Cherrat, et al. [81] used a convolution neural network for advanced biometric system design with cancellable biometrics rules. In their previous work Cherrat et al. [81] and Carrión-ojeda et al. [82], used braided CNN for speech classification, which generated a very high accuracy of 97.5%, 98.8% and 98.06% for US8k, GSCv1, and GSCv2 datasets, respectively. Later in the cited work, they utilized cancellable biometrics for upgrading the security and protection in-ear biometrics. This shows the path to use cancellable biometrics in MBS design to improve safety and privacy. CNN and SVM combination are utilized for feature extraction and classification part.

For multimodal biometrics, gait is an essential feature as user acceptance, and ease of data collection are very high. In 2014 Guan et al. [59] successfully presented a novel approach to obtained gait identification from low-quality videos where the frame rate was 1fps and resolution was 32X22. To reduce error large number of weak classifiers (ensemble of classifiers) were used over average gate image (AGI), and it is noted that performance was highly correlated with the diversity. The experiment was conducted on outdoor and indoor databases like USF, and OU-ISIR-D.

In 2018, a detailed survey on Unimodal and Multimodal biometric was carried out by Oleish [83] and Stylios et al. [73]. They explained all the modules of the biometric system like sensors, feature extractors, pre-processing, matching modules, and decision

modules. Various biometric systems like face, ear, iris, fingerprint, voice, gait, palmprint have been mentioned with details. The fusion methods were classified into three categories such as rule-based method, classification-based, and estimation based fusion methods. Kalman filter and particle filter had been explained for the estimation based fusion category.

CNN-based biometric system design is becoming popular in recent years as traditional feature extractions/ classification methods are less efficient in large datasets [84]. Also, it removes the need to pre-process the inputs, as multiple filters inside the network layers perform the task automatically in the state of the artwork, invalidated the significance of sample size in biometrics identification. The proposed multivariate copula models for correlated biometric acquisitions demonstrated the calculation of minimum numbers of samples required for the authentication process. CNN based system is managing such tasks with more excellent performance [81].

G. Kumari et al. delineated an excellent analytical outlook on customer awareness towards the biometric mechanism of unimodal and multimodal systems in online transactions. It motivates system designers to improve system performance and integration with modern e- platforms such as banking sites, e-commerce platforms and device access [1]. More than 100, 93.4% of users have shown high interest in multimodal biometrics systems based on online transactions in the selected sample size. Furthermore, a multidimensional questionnaire was floated among users from different classes regarding the profession, gender, age etc. As per the PLS-SEM report, 68.87% of customers reflected confidence in the online purchase if platform authentication is MBS equipped.

By analyzing the previous outcomes of biometrics behavioral studies in a recent survey, it has been observed that several studies have subsumed the adoption of technology through a notion of a single theory. A single theory cannot attempt the comprehensive view of research matter, so a model of several theories is conducive and advantageous for panoramic in-depth research [85].

Several studies focus on the organization's internal and external conditions and technological aspects in analyzing drivers for new technology diffusion [86]. The TOE framework [87] considers multidimensional factors when studying technology adoption. It yields a greater explanatory power than other adoption models such as the Technology Acceptance Model [88]. The manifests selected in this study also mapped with the Theory of Planned Behavior matrix [89], the Unified Theory of Acceptance and Use of Technology [76] and Social cognitive theory [90].

Comprehensive usages of these models may help establish whether MMB is appropriate in facilitating the delivery of online financial transactions. The extended model fit describes how suitable MMB is for delivering services, and this may be tested by determining how user-friendly and secured the tasks are. IN the current Covid 19 era, researchers are working on touchless biometrics, remote authentication processes using AI-ML models [91]. These technologies are more user-friendly, robust, and reliable for commercial use and high-security asset access [92]. User Acceptability and Data Privacy Factors extracted from the above theoretical model for this research.

1.4 Research Gaps

- (i) In the multimodal based biometric recognition, several works are done at various fusion levels such as sensor level, matching score level, and decision level, etc. Multimodal fusion operation at feature level normally is difficult to generate desirable results due to the potential incompatibility of feature spaces formed by different modalities. It is observed that the canonical correlation analysis method is insufficient to reveal the complex and nonlinear correlation relationship between two features sets. In the multimodal biometric system, the integration of feature sets is difficult task, especially when the feature sets of different modalities are incompatible; the relationship between the feature space of multiple modalities is unknown and the curse of dimensionality problem due to ensemble approach. The limitation of the feature level fusion is that feature sets from the multimodalities are difficult to be accessible and

are mostly non compatible. In the feature level fusion, the concatenation may yield a very large dimensional feature vector due to the presence of noisy or redundant data, resulting in a decrease in the performance.

- (ii) E-proctoring is a great challenge especially after Covid era. In online education, examination and business meeting etc. live proctoring is important and crucial. In reported work, few methods have been reported but most are facing challenges in real time implementations.
- (iii) Real time implementation of multimodal biometric system for portable devices with limited memory and power is a challenge. Convolutional neural network has significantly reduced the dependence on hand crafted features, but its complexities and large size is also a challenge especially in the era of smart devices of limited power backup.
- (iv) The multimodal biometrics application's commercial success depends upon user acceptance and trust. It's very important to understand the level of awareness and perceptions about integration of biometric process in routine solutions like online transaction, E-Proctoring, remote vigilance an information collection at different interfaces. Less work has been reported in this direction which highlights the requirement of detailed study and analysis of what user thinks about offered solutions.

These are the main drawbacks of various existing multimodal biometric systems, which motivate us to do this research on the multimodal biometric system. We intend to propose a suitable method to work on these research gaps and achieve better performance.

1.5 Research Objectives

This research work focuses on analytical study of various multimodal biometric systems and their fusion technologies, Understanding the scope of improvements in the existing fusion methods, at different level of fusion, and finding out efficient fusion methodology for multimodal biometric system at most suitable.

Covid 19 has also created new challenge of remote proctoring in various dimensions of routine life, e.g., remote plant monitoring, online examinations. The emphasis shall be upon to explore feasibility of implementation of multimodal systems for continuous authentication. CNN based algorithms have also drawn interest for efficient implementation for contemporary and future platforms. Study of User adaptation or multibiometric technology is also our point of interest. The flow of the study and investigation are as follows:

1. A comparative study and analysis of various fusion technologies of multimodal biometrics (e.g., face and finger, finger and iris, iris and face etc.) at different level viz. feature level, rank level, score level and decision level.
2. Design and development of fusion techniques that would improve the performance measure of multimodal biometric system and validation of these techniques through simulation on standard databases. Exploring efficient fusion scheme for multimodal biometric person authentication for continuous application.
3. Analytical study of user awareness and adaption of Multimodal biometric authentication method for online services.

1.6 Thesis Outline

Chapter 2

It highlights the prominent biometric databases available for experimental and research work, base line methodologies of multimodal biometrics system and its performance, challenges and scope of improvement from baseline level.

Chapter 3

It proposes a new optimum feature level fusion method based on Grey-wolf optimization technique for feature level fusion. Optimum feature level fusion has been used for multimodal biometric system design for Finger Print, Iris and Face.

Chapter 4

It elaborates application of multimodal biometrics in continuous user authentication in online proctoring. LCNN based novel method of continuous user biometric authentication in online examination (CUBA-OE) has been shared. The model presents result for fusion of key board stroke pattern, Iris and Face of user during live interaction.

Chapter 5

It highlights the importance of CNN results based multimodal biometrics systems for user authentication process. The CNN based multimodal system architecture for authentication are compared with different methods, with focus on the real time implementation and deployment.

Chapter 6

It presents detailed study of user awareness and perceptions about multimodal biometrics for real time application such as online transactions. A novel SEM-ANN approach has been presented which gives new insights about biometric authentication in daily life.

Chapter 7

It concludes the results obtained from proposed OGWF, CUBA-OE & SEM -ANN methods. It also gives insights of future aspects and further studies related to multimodal biometric system design.

Chapter 2

Database, Fusion and Multimodal Biometric Experiments

2.1 Introduction of Biometric Databases

In the recent decade, there has been a significant increase in interest in biometric recognition systems for person authentication. The availability of biometric databases is one of the important factors in this achievement; these are critical for defining standard criteria that allow for consistent comparison of rival recognition algorithms. The design, acquisition, and gathering of these databases are one of the most time and resource-intensive processes for the research community, particularly when multimodal databases with numerous biometric attributes and acquisition sessions are involved. Biometric experiments depend upon the availability of the databases and their quality. There are several research groups and University labs that are contributing to this field and helping researchers. The digital transformation is making images of traits more accessible and a few clicks away. For the proposed work various database samples have been used. As work is based on Face, Ear, Iris, and Fingerprint, so we are highlighting these databases here.

The following key databases have been referenced for the purpose of the research.

2.1.1 Iris Databases

The IIT Delhi Iris database [93] is primarily based on information provided by students and staff at the institute. During the time in Biometric Research Laboratory from January to July 2007, a JIRIS, JPC1000 digital CMOS camera was employed for this database. This image acquisition application was created to capture and preserve iris images in bitmap

format, which are also accessible upon request. This database contains 224 images in bitmap (*.bmp) format, which were collected from 224 users. This database considers 176 males and 48 females between the ages of 14 and 55. This database, which contains 1120 photos, is arranged into 224 folders, each of which is associated with a unique integer identification/number. These photographs were captured in an indoor setting at a resolution of 320 X 240 pixels.

Iris Database at IIT Delhi (Version 1.0): The research's effectiveness is contingent on the availability of a large-scale iris database that can be easily obtained in a real-time context. IIT Delhi's Biometrics Research Laboratory has been collecting iris data since 2007. The primary concern is to create a large and accurate iris database of Indian users and make it publicly available.

Sample Images

This database is being made public as a whole (none of the photographs have been removed as of yet), with different image quality. The following image collection essentially replicates the example photographs from this database.

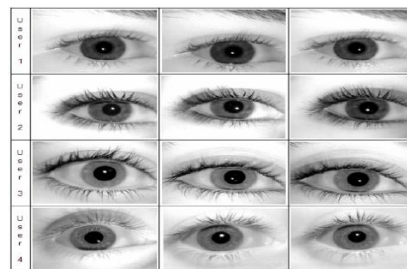


Fig. 2.1 Sample iris images

CASIA-IrisV1

The CASIA-Iris Image Database Version 1.0 (CASIA-IrisV1) [94] was created by the Chinese Academy of Sciences' Institute of Automation and contains 756 photos gathered from 108 eyes using the CASIA close-up camera (CASIA). This device makes use of near-infrared illumination. In two separate sessions, seven pictures were obtained for each eye.

In the first and second sessions, three and four pictures are captured, respectively. Each eye image was saved as a BMP file with a resolution of (320 X 280). The gray-level intensity for the pupil in each image was manually modified using a circular zone of constant gray-level intensity to suppress reflections. This change aids in the detection of iris contours, but it has little or no effect on the remaining system components.

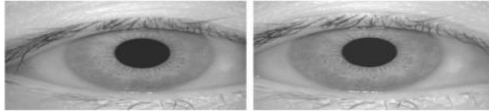


Fig. 2.2 CASIA iris v1 images

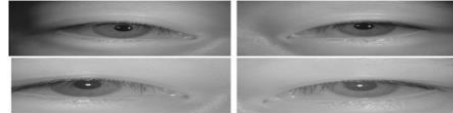


Fig. 2.3 CASIA iris v2 images

CASIA-IrisV2

CASIA's CASIA-Iris Image Database Version 2.0 is referred to as CASIA-IrisV2. It contains two subsets of images taken in an in-door setting using two separate image acquisition devices: CASIA-IrisCamV2 (OKI IRISPASS-h) and CASIA-IrisCamV2 (OKI IRISPASS-h). Each subset has 1200 photos from 60 different classes. Each image was saved in BMP format and had a resolution of 640 X 480. This database's eye images are shown in the figure below.

MMU V1 Iris Database

This database was created by Malaysia Multimedia University (MMU)[95]. The LG IrisAccess®2200 outfitted with NIR illuminators was used to collect 460 photographs from 46 people as shown in the figure. Each person is asked to submit a set of ten photos. Images having a resolution of 320 X 240 pixels were reserved in the BMP format.



Fig. 2.4 MMU v1 iris

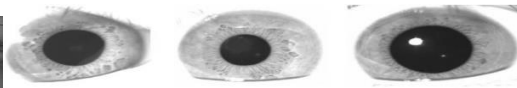


Fig. 2.5 CUHK v1 iris images

CUHK V1 Iris Database

This database was created by the Computer Vision Laboratory [Department of Automation and Computer-Aided Engineering, Chinese University of Hong Kong (CUHK)] [96].

There are 250 eye images in all, including seven images from each of the 36 Asian subjects. Each eye image was saved as a BMP file with varying resolutions. These eye images have a variety of viewing angles. Furthermore, light reflections, eyelashes, and eyelids are prominent sources of noise. This database's eye images are shown in the Fig. 2.5.

BATH Iris Database

The University of Bath in Bath, Somerset, United Kingdom, created it. It has 1000 eye images taken from 50 different eyes, each with 20 pictures. Each image has a resolution of 1280 X 960. At a bit rate of 0.5 bpp, all photos are compressed using JPEG2000 compression. For decompressing pictures, the Kakadu software is recommended. Eyelashes, eyelids, and reflections are common sources of noise. Some sample eye images are shown in the diagram below.

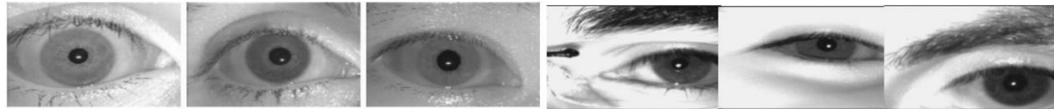


Fig. 2.6 BATH iris images

Fig. 2.7 MMU v2 iris images

The MMU V2.0 was created by Malaysia Multimedia University. The Panasonic BM-ET100US Authenticam was used to capture 995 eye images from 100 people. Participants come from all over the world, including Asia, Africa, the Middle East, and Europe. A total of ten photos are acquired from each individual, five for each eye. In addition, due to cataract illness, five left eye images were removed. Images having a resolution of 320 X 240 pixels were reserved in BMP format. Rotated irises, non-uniform light, blurring, reflections, contact lenses, eyelids, eyelashes, eyeglasses, and hair are among the noisy features in this dataset.

2.1.2 Ear Databases

The IIT Delhi ear image database [98] is a collection of ear images acquired from IIT Delhi students and staff in New Delhi, India. Using a modest imaging setup, this database was obtained on the IIT Delhi campus between October 2006 and June 2007 (still in process).

All of the photographs are taken from a distance (touchless) using a simple imaging setup, and the imaging is done indoors. The database now accessible is made up of 121 individual subjects, each of whom has at least three ear images. All of the subjects in the database are between the ages of 14 and 58. Every user's image in the database of 471 has been consecutively numbered with an integer identification/number. These photos have a resolution of 272 X 204 pixels and are all accessible in jpeg format. This database also includes automatically normalized and cropped ear photos with a size of 50 X 180 pixels, in addition to the original photographs. A larger version of the ear database (automatically cropped and normalized) from 212 users with 754 ear photos has recently been incorporated and made available on demand.

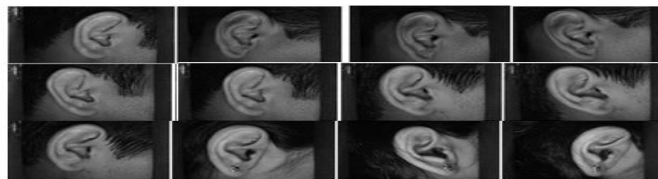


Fig. 2.8 IITD ear database images

2.1.3 Face Databases

A series of tests using different subsets of the face are used to investigate the performance of our unimodal and multimodal systems. FERET, ORL, Yale-B, and Essex [99-102] are the face databases used in this study. Each subsection has a quick overview of the face database. FERET Face Database

The Facial Recognition Technology (FERET) Database [99] ran in 15 sessions from 1993 to 1997. The Defense Advanced Research Products Agency and the Department of Defense's Counterdrug Technology Development Program sponsored the event (DARPA). The final corpus has 14126 face photos from 1564 sets of images with 1199 subjects and 365 duplicate sets of images. Duplicate sets were taken on various days that cover the second image sets of the same people. The photos of the same individual in duplicate sets were taken during a two-year period. The image dimension is 256 X 384 pixels.

FERET Database Naming System shows the naming convention for FERET imagery,

which includes frontal photos and position angles, depending on distinct categories. After cropping the photos to 80 X 64 pixels, a subset of 235 participants were employed in this study. This dataset contains photos with various lighting conditions (right-light, center-light, and left-light), regular and alternate facial emotions (happy, normal, drowsy, sad), a broad range of stances (both frontal and oblique views), and with and without spectacles.

ORL Database [101], AT&T face database, formerly known as the ORL face database, is a standard face database including a set of facial photographs obtained at the lab between April 1992 and April 1994. The ORL database was used in a facial recognition study conducted in partnership with the Cambridge University Engineering Department's Speech, Vision, and Robotics Group. It includes ten different photos for each of the 40 different subjects. They were photographed at various times, with variable lighting, facial emotions such as open and closed eyes, smiling and not smiling, and facial details with and without spectacles, all against a dark homogenous background.

Extended Yale-B Face Database [100]

Kuang-Chih Lee and Jeffrey Ho were the first to report on the expanded database, as compared to the original Yale Face Database B with ten subjects. All photographs in the database are carefully aligned, cropped, and resized to fit within the 168 X 192 image limit. It has 16128 photos of 38 human beings in 9 positions with 64 different lighting settings. In the Extended Yale-B Database each individual provides a total of 20 photos of their face. The image has a resolution of 196 X 196 pixels. All of the photographs were taken with artificial lighting, which included a mix of tungsten and fluorescent overhead lights. It includes photographs of both male and female people of various races.



Fig. 2.9 ORL face images

Fig. 2.10 Yale B face images

2.1.4 Fingerprint Databases

Fingerprint databases are organized collections of fingerprint data that are mostly used for evaluation and operational recognition. Because of its permanence and uniqueness, fingerprint recognition is the most widely used biometric technique in personal identification.

CASIA-Fingerprint V5

Version 5.0 of the CASIA Fingerprint Image Database [103] contains 20,000 fingerprint pictures from 500 people. The fingerprint images of CASIA-Fingerprint V5 were obtained using URU4000 fingerprint sensor in one session. Each subject provided 40 fingerprint photographs of his eight fingers (left and right thumb/second/third/fourth finger), for a total of 5 images per finger. All of the fingerprint images are 8-bit gray-level BMP files with a 328 X 356 resolution.



Fig. 2.11 CASIA fingerprint image database images

FVC Databases

In the years 2000, 2002, 2004, and 2006, [104-107] four worldwide Fingerprint Verification Competitions (FVC) were held. Four databases were gathered for each competition using three different methods, various sensors. There are 110 fingers in each database (150 in total), 8 imprints on each finger, for a total of 880 impressions (FVC2006).



Fig. 2.12 FVC databases images

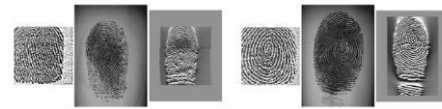


Fig. 2.13 BioSec fingerprint database images

BioSec Fingerprint Database

BioSec was a ‘6th’ European Framework Program Integrated Project (IP) with approximately 20 partners from nine European countries. The baseline corpus consists of 200 subjects with two acquisition sessions per subject, but the BioSec expanded version included four acquisition sessions per subject. Fingerprints were obtained using three separate sensors.

2.2 Fusion of Features

Although a unimodal biometric system performs well in some situations, it has a number of flaws, including non-universality, spoofing susceptibility, disturbances in sensed data, intra-class differences, and interclass similarities. Scientists have recently attempted to combine different modalities for human identification, known as multimodal biometrics. For enrollment, multimodal biometric identity systems can use more than one physical, behavioral, or chemical feature. When compared to systems based on a single biometric modality, multimodal biometric systems combine the information offered by various biometric sources and often deliver greater recognition performance. Sensor level fusion, feature level fusion, score level fusion, and decision level fusion are four different fusion methodologies that have been employed for various multimodal systems to combine multimodal information.

A process known as “fusion” is preferred for joining two or more biometric features. Fusion is the process of merging two or more biometric modalities in biometry. As it has proven to be a promising trend both in tests and real-life authentication applications, multimodal biometrics systems always demand the entire integration of data from diverse modalities such as face, ear, and iris. Fusion can be done both before and after matching for this reason.

Prior to matching, sensor, and feature level fusion can be accomplished, and decision, rank, and score level fusion can be accomplished afterward. In this section, we'll go through the various fusion scenarios used by multimodal biometric systems.

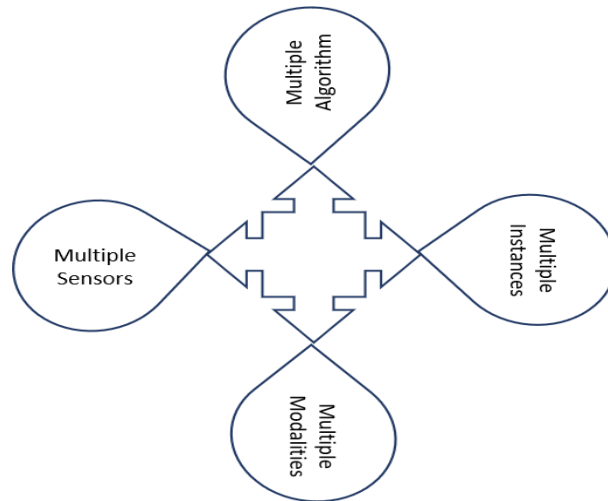


Fig. 2.14 Different types of multibiometric combination

It is important to note that multimodality does not imply the employment of several biometric modalities in the literal sense of the term (for example, combining iris and fingerprint), but rather a broader definition as defined by the numerous fusion scenarios described below.

In multimodal systems, there are five levels of fusion in general as below:

- (i) **Sensor Level-** Multi-sensorial biometric systems collect data from two or more sensors to sample the same instance of a biometric feature. Multiple samples can be processed using a single method or a collection of algorithms. A visible light and infrared camera, along with a specified frequency, could be used in a face recognition application.
- (ii) **Feature Level-** In classification, feature level fusion is useful. Different feature vectors are mixed, either from different sensors or from different feature extraction methods applied to the same raw data.
- (iii) **Decision Level-** Each biometric subsystem completes the operations of feature

extraction, matching, and recognition on its own using this method. Boolean functions are commonly used in decision techniques, with the recognition yielding the majority decision among all current subsystems.

- (iv) **Rank Level-** Partitions of the template are used instead of the complete template. The fusion rank for the categorization is estimated by combining the ranks from template partitions. Combining identifying ranks acquired from several unimodal biometrics is known as rank-level fusion. It combines a ranking that will be utilized to make a final judgement.
- (v) **Score Level-** It is the sum total of the matching scores provided by the various systems. Fixed rules (AND, OR, majority, maximum, minimum, total, product, and mathematical rules) and taught rules are the two types of score level fusion approaches (weighted sum, weighted product, fisher linear discriminate, quadratic discriminate, logistic regression, support vector machine, multilayer perceptron's, and Bayesian classifier).

Sensor and Feature level fusion are considered as “fusion before matching” and remaining three are under “fusion after matching category”.

The following fusion schemes have been considered for this study:

A. Feature-Level Fusion

The fusing of feature vectors obtained from many feature sources is known as feature-level fusion:

- (i) feature vectors based on a single biometric that are gathered from various sensors;
- (ii) feature vectors derived from several entities based on a single biometric, such as iris feature vectors derived from both left and right eyes; and
- (iii) feature vectors derived from a variety of biometric traits.

In general, biometric systems that integrate data early in the processing cycle are thought

to be more effective than systems that primarily do fusion afterward. Furthermore, feature-level fusion produces considerably superior recognition outcomes. However, because of compatibility issues with the feature sets of distinct modalities, fusion at this level is rather difficult to establish.

B. Score Level Fusion

Fusion on the corresponding score level is known as score-level fusion. At this level, multiple matching scores acquired from separate classifiers or from distinct biometrics can be combined for matching. Fusion can be treated at the matching level in two ways: as a classification problem or as an information combination problem. A feature vector is reconstructed utilizing matching scores that are generated based on individual matches in the classification procedure. The feature vectors are then categorized as “Accept” (real user) or “Reject” (impostor). Individual matching scores are fused to yield a single scalar score that is utilized to determine the final decision in the information combination strategy.

Before fusion, the individual matching score must be normalized to a uniform field. Because the acquired scores do not have to be within the same range, this normalizing approach is required.

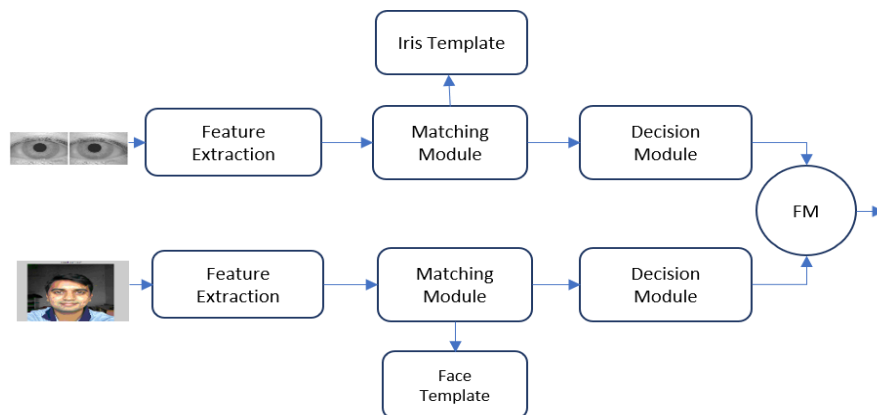


Fig. 2.15 Standard score level fusion

2.3 Performance Metrics of Biometric System

FAR/GAR/FRR/GRR/Precision/Accuracy/Recall Rate, F1 Score, Sensitivity, Confusion Matrix, Equal error rate, the false acceptance rate, or FAR, are the key parameters that are used to evaluate the performance of biometric systems. FAR indicates the likelihood that a biometric security system may wrongly accept an unauthorized user's access attempt. The number of erroneous acceptances divided by the number of identification attempts is typically used to calculate a system's FAR.

Table 2.1 Performance matrix of biometric systems

Sr. No.	Performance matrix
1	$TPR = TP/Actual\ Positive = TP/(TP + FN)$
2	$FNR = FN/Actual\ Negative = FN/(TP + FN)$
3	$TNR = TN/Actual\ Negative = TN/(TN + FP)$
4	$FPR = FP/Actual\ Negative = FP/(TN + FP)$
5	$Precision = TP/(TP + FP)$ (4.35)
6	$Recall\ rate = TP/(TP + FN)$
7	$F - 1\ Score = 2 * (Precision * Recall\ rate)/(Precision + Recall\ rate)$
8	$Accuracy = TP + TN/(TP + TN + FP + FN)$
9	$Specificity = TN/(TN + FP)$

TP = True Positive, TN= True Negative

FRR measures the risk that a biometric security system will mistakenly reject an authorized user's access request. FRR is calculated by dividing false recognitions by identification attempts.

Scores are used to demonstrate pattern-biometric template similarity (also known as weights). Higher score for resemblance. The score for a trained individual (identification) or the person against whom the pattern is proven (verification) must be above a set threshold. Theoretically, client scores (patterns from known people) should be higher than imposter scores. If so, a single threshold might be utilized to separate clients from impostors by scores. Real-world biometric technologies disprove this notion for several reasons. Imposter patterns sometimes outscore client patterns. No matter the categorization threshold, errors will occur. So, setting the barrier so high that no impostor scores exceed it, is right approach in some cases. In opposite condition the system accepts all incorrect patterns. Low-scoring client patterns would be falsely rejected. To avoid that low threshold

is required so no client patterns are wrongly rejected. Threshold is optimum point between these two points, so that false recognitions and acceptances do not occur.

Like imposter scores, client pattern scores fluctuate around a mean. Depending on the threshold, none to all client patterns are incorrectly refused. False recognition rate is the ratio of rejected to total client patterns (FRR). Its FAR value is between 0 and 1.

When client and imposter score distributions overlap, choosing a threshold value is challenging. Manufacturers only provide a single FAR number for comparing two biometric technologies, which is insufficient. The system with the lower FAR presumably has an unacceptable FRR. Even if FAR and FRR numbers are provided, threshold dependence is a concern. Assuming the criteria are adjustable, it's impossible to tell if a system with a higher FAR and lower FRR is preferable than one with a lower FAR and higher FRR.

Before implementing a solution, one should know these phrases and estimate the right requirements. FAR only gives half the story. When a biometric solution seller claims a low FAR, one should find out the FRR at this 'low' FAR. Determine if FAR and FRR are appropriate for the application. Low FAR and high FRR prohibit unwanted entry in real life. It would also require authorized users to touch the device many times before access. So, it is very important to have a clear understanding of performance matrix of Biometric systems.

Equal Error Rate

The value of FAR for which FAR and FRR are equal is known as the equal error rate (EER). That is, EER is the point where the FAR and FRR curves meet. Also, if we construct a curve of FAR vs FRR for all thresholds and connect it with a line drawn at 45 degrees from the origin, EER is the point where the two curves intersect.

Accuracy can be defined at T (threshold) if $\frac{FAR+FRR}{2}$ is the minimum for all FAR and FRR

at different thresholds.

$$\text{Accuracy} = [100 - (FART + FRRT) / 2]$$

where FART, FRRT are FAR and FRR at threshold T.

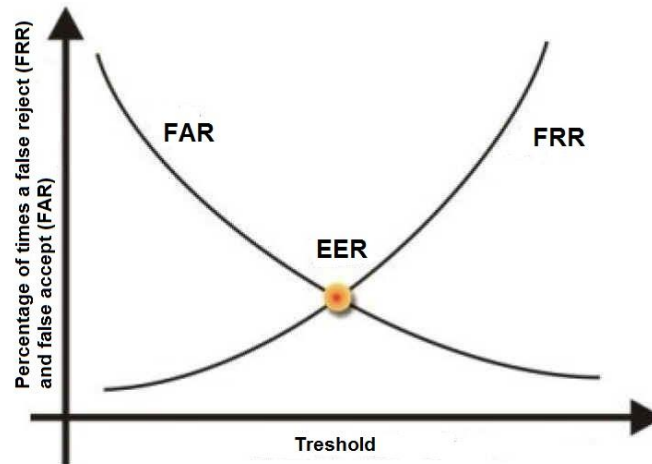


Fig. 2.16 Equal error rate [36]

The optimum threshold (T) is defined as the point at which the combination of FAR and FRR yields the maximum accuracy. It's worth noting that accuracy may not be at its peak at the EER level.

ROC Curve (Receiver Operating Characteristics)

A plot of True Positive Rate (TPR) on the y-axis against False Positive Rate (FPR) on the x-axis is the ROC curve. The ROC curve assists us in determining the threshold at which the TPR is high and the FPR is low, indicating that misclassifications are low. As a result, while determining the best probability threshold for a classification model, ROC curves should be employed. Furthermore, the cost of false positives and false negatives is not always the same in many circumstances. This is where ROC curves come in handy. AUC refers to the area beneath the ROC Curve. AUC is a total measure of a model's performance across all conceivable categorization levels.

Sensitivity

The measure of the sensitivity is the proportion of actual positives which are accurately recognized. It relates to the capacity of the test to recognize positive results.

$$\text{Sensitivity} = \frac{\text{TP}}{(\text{TP} + \text{FN})}$$

‘Where TP stands for True Positive and FN stands for False Negative’

Specificity

The evaluation of the specificity is the extent of negatives that are properly recognized. It relates to the capacity of the test to recognize negative results.

$$\text{Specificity} = \frac{\text{TN}}{(\text{TN} + \text{FP})}$$

‘Where TN stands for True Negative and FP stands for False Positive’

Accuracy

The accuracy of the suggested method is the ratio of the total number of TP and TN to the total number of biometric traits.

$$\text{Accuracy} = \frac{\text{TN} + \text{TP}}{(\text{TN} + \text{TP} + \text{FN} + \text{FP})}$$

Table 2.2 Confusion matrix

Experimental Outcome	Condition as determined by the Standard of Truth	
	Positive	Negative
Positive	TP	FP
Negative	FN	TN

2.4 Base Line Methodologies

In presented work authors have experimentally tested different methods of authentications. Adopted process was different in terms of modalities types and number, feature extraction methods, fusion types, and in terms of classification strategies. Some of these have been documented in this chapter.

Three main combinations experimented with are as follows

- (1) Multimodal Fusion: This is the standard process where multiple traits are fused at different levels. Multimodal has outperformed unimodal process.
- (2) Multilevel Multimodal Biometric Fusion: In this process, results of unimodal scores and feature level results are fused together. In experiments it has been observed that this scheme out performs normal multimodal process.
- (3) Multilevel Multi-Classification Approach: The classifier is very important because entire process depends upon its performance. Having multiple classification, ensures high reliability in end results.

During this work, many experiments of the unimodal, score level fusion, and feature level fusion have been performed. Some of these are presented below, were of multimodal multilevel classification approach. In these experiments, different feature extraction methods, classifications, fusion levels were explored.

2.4.1 Multilevel FusionType 1

With respect to this work, it has been proposed that features would be fused at multiple levels such as feature level, score level and combined level. Multiple level fusion of ear, iris and face has been performed, first at feature level and then at combined level fusion of all previous output. After acquiring ear and face features, individual matching scores were calculated for random forest and Euclidean classification methods. Combined feature sets were generated after the fusion of features generated at individual level. Joint score has been obtained when the combined feature set was matched.

2.4.1.1 Proposed Methodology

In the present work, features of ear, iris and face are fused together. IIT Delhi ear database 1.0 and near IR face database version 2.0 has been utilized for taking ear and face images respectively. Apart from that, self-made small database of 20 users has been tested by the authors. So, in this following section, proposed fusion methods and processing for input images is explained.

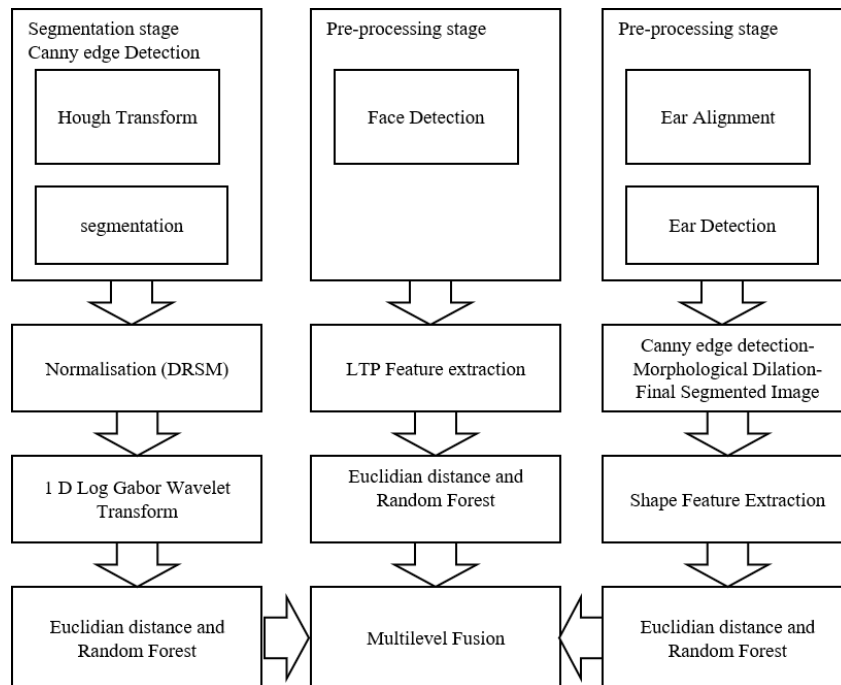


Fig. 2. 17 Fusion of iris, face & ear biometric traits

In this experiment Iris , Face and Ear were processed in parallel manner as shown in the Fig. 2.17. Features were extracted and compared with templates before and after matching. In Multimodal process, features of all three traits were fused together and compared with multimodal template stored.

At the next stage all the scores generated in previous step of unimodal and multimodal processes were fused second time. As fusion is taking place twice at different level of process pipeline, so it is called multilevel fusion. For classification purpose euclidean distance and Random forest classifier were used.

Iris Processing

A segmentation algorithm is performed on iris images, so that localization of iris area from full eye and isolation with eyelid, eyelash and reflection areas can be obtained. This segmentation is realized using the circular Hough transform, and the linear Hough transform for localizing eyelids. Thresholding is used for segregating eyelashes and reflections. Normalization of segmented iris region is done to overcome dimensional inconsistencies in-between iris region. The Dougman's Rubber Sheet model generated normalised iris images, 1D-Gabor is used to extract the unique features which include real and imaginary part. The proposed method uses these features to recognize individual's identity on different sets.

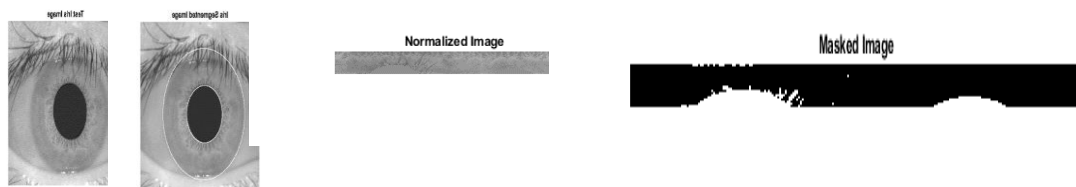


Fig. 2.18 Iris recognition process from sample image

Ear Processing

Multistep process has been used for ear recognition with accuracy of process depends upon adopted methods. So as a first step, images were acquired from the gallery and a probe image is taken and then ear extraction has been done from the probe image. After extraction, landmark detection stage comes up and then comes the normalization to feature extraction stage. Later matching has been done and then finally fusion and decision making has been achieved. The following block diagram explains the mechanism.

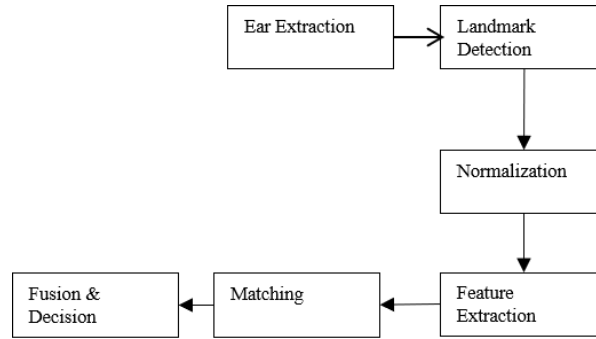


Fig. 2.19 Ear recognition process block diagram

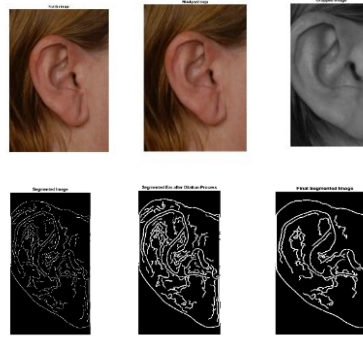


Fig. 2.20 Ear recognition process using canny edge detection

Generally, image acquisition in real time suffers from image quality and thus other several parameters would affect performance. Basically, standard reference ear images were used to align the acquired images and for alignment purpose, SURF feature descriptor was utilized. After that, by thresholding and using morphological operations, ear detection was performed. Later, hole filling and filtration process takes place once the binary images are generated from aligned images for efficient detection.

After this mechanism, segmentation methodology is applied using edge detection method [14]. Sometimes, we need to reduce the amount of data in an image while preserving the structural properties, so for that purpose, edge detection-based segmentation has been deployed. For the detection of ear edges, canny edge detection operator that uses a multistage algorithm to detect wide range of edges present in any image, has been utilized. Later, one of the famous morphological operations called dilation is applied. This dilation

which is applicable for both binary as well as gray image, gradually increases boundaries of foreground pixels. This makes the area to grow in size and making holes in that region to become smaller in size. Next stage is the feature extraction after segmentation stage in which shape features are calculated by counting of non-zero values in segmented ear image. At the final stage for ear recognition, Euclidean distance and Random Forest classifier are used [15].

Given two data points $A = [a_1, a_2, \dots, a_n]$ and $B = [b_1, b_2, \dots, b_n]$, the Euclidean distance between A and B is given by

$$dis(A, B) = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \dots + (a_n - b_n)^2} \quad (2.1)$$

Face Processing

Step by step approach for face recognition is: Image acquisition, face detection, face alignment, feature extraction and recognition. Initially test images were acquired from gallery and after that, faces were detected and cropped for further processing by the help of face detection tool. SURF feature description was applied to take care of alignment issue by comparing images with standard images. For the purpose of feature extraction, local ternary patterns (LTP) were utilized to calculate mean which can be used for score generation. From LTP features, mean was calculated and used for score generation. Again, at the final stage, for classification purpose, Euclidean distance and random forest methods were used. Basically, after thresholding step, the upper binary pattern and lower binary pattern are constructed and coded. The LTP operator is the concatenation of the code of the upper binary pattern and the lower binary pattern.

The mathematical expression of the LTP can be described as follows: where i_p, i_c, R, P and are defined as: i_p, i_c denotes the grey value of the center pixel and grey value of the neighbor pixel on a circle of radius R , respectively, and is the number of the neighbors and 't' denotes the user threshold.

$$(LTP)_{P,R} = \sum_0^{P-1} 2^P S(i_p = i_c) \quad (2.2)$$

$$s(x) = \begin{cases} 0 & \text{when } -t < x < t, \\ -1 & \text{when } x < -t, \\ 1 & \text{when } x > t \end{cases} \quad (2.3)$$

The multiple classification method gives good opportunity to select final decision. The results show that it is worthwhile to use multiple biometrics for recognitions.



Fig. 2.21 Face recognition process from sample image

Random Forest for Classification

1. For $b = 1$ to B
 - a) Draw a bootstrap sample Z^* of size N from the training data.
 - b) Grow a random-forest tree T_b to the bootstrapped data, by recursively repeating the following steps for each terminal node of the tree, until the minimum node size n_{\min} is reached.
 - i. Select m variables at random from the p variables.
 - ii. Pick the best variable/split-point among the m .
 - iii. Split the node into two daughter nodes
2. Output the ensemble of trees $\{T_b\}_1^B$
To make a prediction at a new point x :

$$\text{Regression } \hat{f}_{rf}^B(x) = \frac{1}{B} \sum_{b=1}^B T_b(x) \quad (2.4)$$

Classification: Let $\hat{C}_b(x)$ be the class prediction of the b^{th} random-forest tree. Then

$$\hat{C}_{rf}^B(x) = \text{majority vote } \{\hat{C}_b(x)\}_1^B \quad (2.5)$$

In averaging process, the bias of bagged trees is the same as that of the individual trees. An average of B random variables, each with variance σ^2 has variance $\frac{1}{B} \sigma^2$. If the variables are simply identically distributed, with positive pairwise correlation ρ , the variance of the average is

$$\rho \sigma^2 + \frac{1-\rho}{B} \sigma^2. \quad (2.6)$$

Before each split, select $m \leq p$ of the input variables at random as candidates for splitting. Typically values for m are \sqrt{p} or even as low as 1. After B such trees $\{T(x; \Theta_b)\}_1^B$ are grown, the random forest (regression) predictor is

$$\hat{f}_{rf}^B(x) = \frac{1}{B} \sum_{b=1}^B T_b(x; \Theta_b) \quad (2.7)$$

Θ_b Characterizes the b th random forest tree in terms of split variables, cutpoints at each node, and terminal-node values.

2.4.1.2 Results

Recognition by feature level fusion methodology in which concatenated feature set of ears, iris and face-based matching resulted with above 93.56% accuracy. Here, the final mean score is calculated for this combined level fusion using the score obtained by merging scores of individual matchings of face, iris and ear with score generated by feature level fusion. Thus, we get the data as S1, S2, S3, S4 as score of ear, iris, face and feature level fusion respectively. Final score 'FS' of matching is obtained by the help of mean of these four scores with 96.56% accuracy. As compared to other single level fusion-based method, this proposed method gives better accuracy, and it requires moderate computation.

In this work threshold level was kept 50% to 70 % for experimental purpose. The Ear based unimodal system has produced 92.6 % accuracy. The ROC curve is given in Fig. 2.22. The ear database is a limited part of IIT Delhi database.

In case of face biometric SURF feature descriptor is used. Using this, local ternary features and subsequently mean features were obtained. The Face unimodal has produced 91.33 % accuracy shown in ROC curve in the Fig. 2.24.

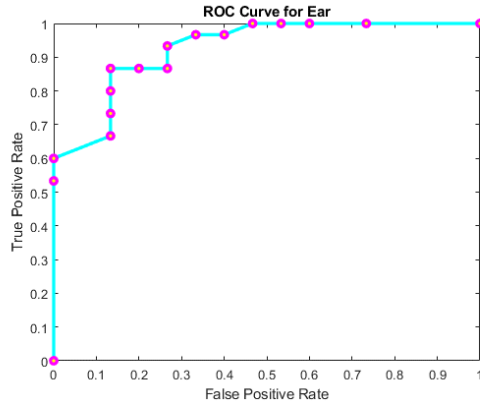


Fig. 2.22 ROC curve for ear

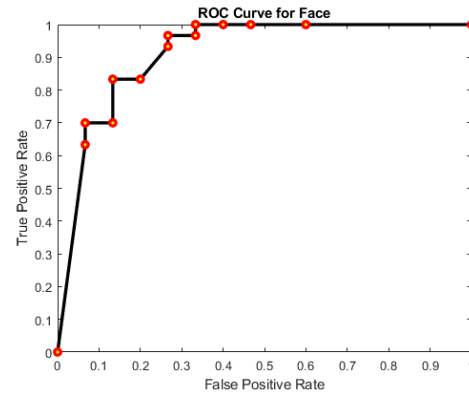


Fig. 2.23 ROC curve for face

The Fig. 2.22 illustrates the iris segmentation and the various stages involved in iris processing. Iris unimodal system has produced 93.44% accuracy as shown in Fig. 2.25.

In this work face, iris and ear features are fused together and Random Forest and Euclidian distance classification is used. The fused feature set increased the level of recognition accuracy. In comparison to accuracy obtained in various methods reported in literature, 93.56% accuracy has been obtained in feature level fusion method. The ROC curve is given in Fig. 2.26.

In next step, as per the proposed process flow, the results of individual systems for ear and face are fused again with the feature level fusion results at last stage of process. This can be called joint or combined or multilevel fusion scheme. In this multilevel fusion scheme, 96.56% accuracy has been obtained as shown in ROC curve in Fig. 2.27, with 94.33% precision and specificity as 94.54%. Recall rate was observed as 96.67%. These values (shown in Table 2.3) prove that proposed scheme outperform other unimodal and feature level fusion schemes. Comparison of different models based on ROC is shown in Fig. 2.28 whereas Fig. 2.29 depicts the comparison of performance parameters.

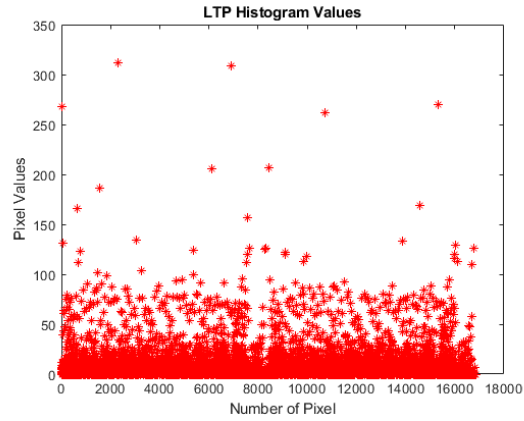


Fig. 2.24 LTP histogram values of face images

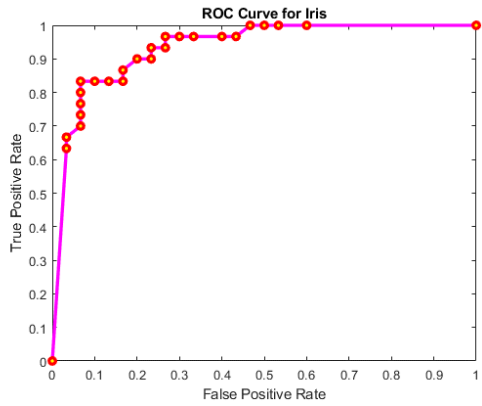


Fig. 2.25 ROC curve for iris

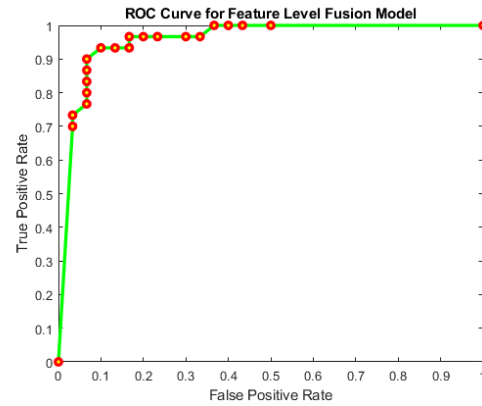


Fig. 2.26 ROC feature level fusion

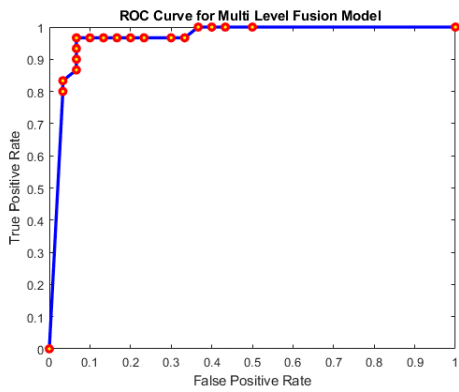


Fig. 2.27 ROC for multilevel fusion

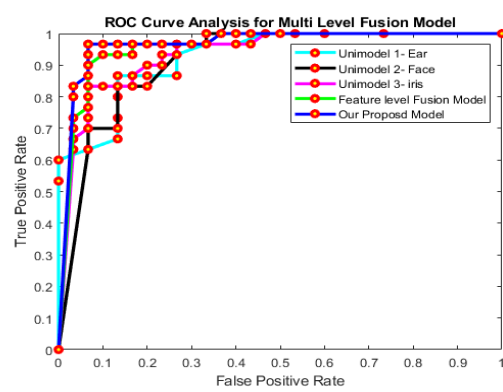


Fig. 2.28 Comparative analysis

Comparison of different models based on ROC is shown in Fig. 2.29. Various performance parameters of proposed method are shown in Table 2.3

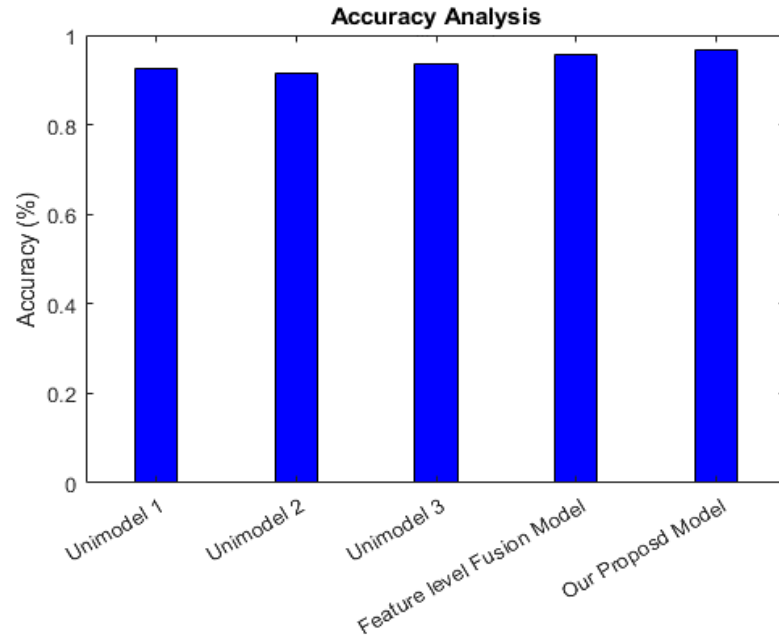


Fig. 2.29 Comparative performance analysis

Table 2.3 Performance-proposed method

Performance Parameters	Proposed method
Accuracy	96.56 %
Precision	94.33 %
Specificity	94.54 %
Recall	96.67 %

Proposed method produced 3.38%, 4.11%, 3.21% and 0.89 % more accuracy in comparison with unimodal methods like face, ear iris, and feature level fusion respectively.

2.4.2 Multimodal Multilevel Fusion Type 2

It is extended work of multilevel multimodal biometrics fusion where Harris Spatio Temporal Corner detection and Speeded Up Robust Features (SURF) extractor were used for the recognition process. MSVM and KNN classifiers were utilized for final recognition stage. For all three traits, simultaneous process goes on. The results of these individual

processes are combined at higher level of the architecture. The result presented below are SURF based only as it has outperformed HSTCP process.

2.4.2.1 Proposed Methodology

In this method, extracted features from different traits are fused together to create multimodal feature vector and it is compared with stored template for getting feature level score. In alternate process individual traits are matched with respective stored template and a matching score is generated for each trait. Respective scores of traits generate individual accuracy for unimodal systems. These scores are fused together to generate major score index for multimodal multilevel fusion.

The operational part of process is divided into three stages as (i) Segmentation stage, (ii) Feature extraction stage and (iii) Recognition stage.

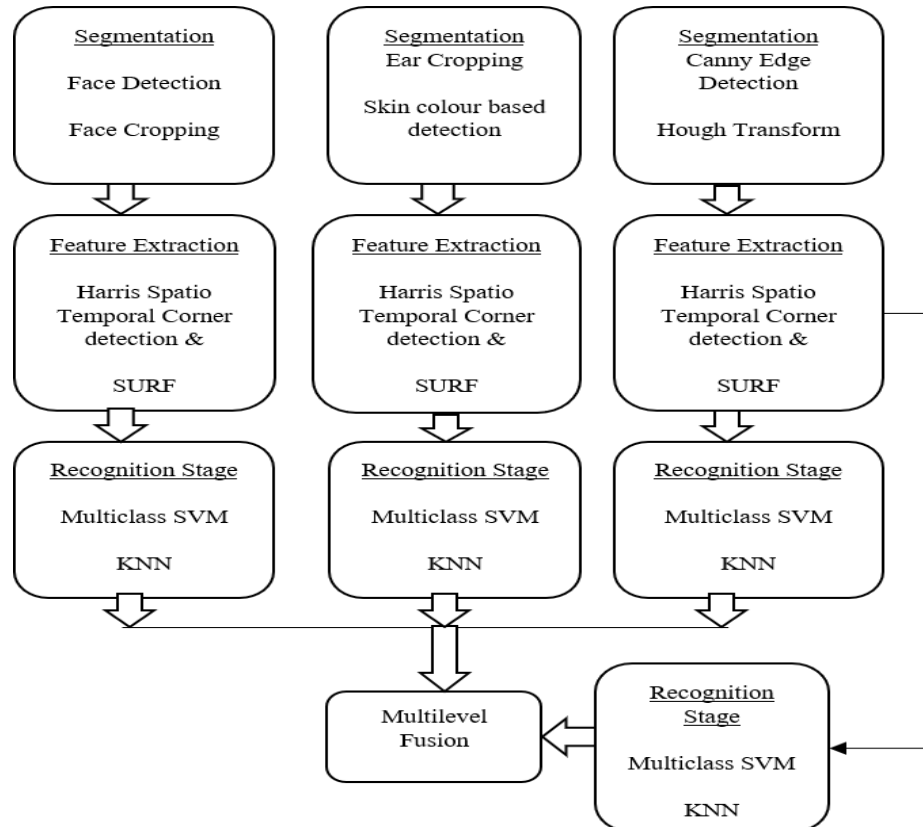


Fig. 2.30 Structural flow of multimodal multilevel fusion of face, ear and iris

I. Segmentation Stage

Segmentation is basically a mechanism which is used to partition a digital image into multiple objects. All the three input test images of face, ear and iris undergo through segmentation mechanism.

For face detection, test images of face are acquired from the gallery and then computer vision face detection tool is used to detect the face present in the test image. The detected face was then cropped for the next process of feature extraction. For ear detection, ear was cropped by automatic cropping method and then it was detected by skin color-based ear detection mechanism.

For iris detection, Canny edge detection is used for the generation of edge map, linear Hough transform is used for localizing occluding eyelids whereas circular Hough transform is used for localizing the iris and pupil regions.

The Ear has certain unique biometric features such as helix, anti-helix, tragus, antitragus and lobe. The geometric and shape-based analysis may generate unique feature set for identification. Skin color segmentation approach with standard feature extraction gives improved feature set. The universality and ease of data collection makes ear a good choice for biometric systems.

II. Feature Extraction Stage

In this stage, for the respective output image from previous stage using Harris Spatio, interest points are detected. In our research work, we have used SURF viz.(Speeded-Up Robust Features) for the temporal corner detector and strong key points detection. SURF is basically an in-plane rotation detector and descriptor in which the detector locates the key points in the image and the descriptor describes the features of the key points and then it constructs the feature vectors of the key points.

Harris Spatio-Temporal Corner Detector (HSTCP) detection algorithms are robust in detecting interest points for image in the Spatio-temporal domain. HSTCP is an extended from Harris corner detector of gray image and it detects local structure where the image values have significant variations in both the spatial domain and the temporal domain.

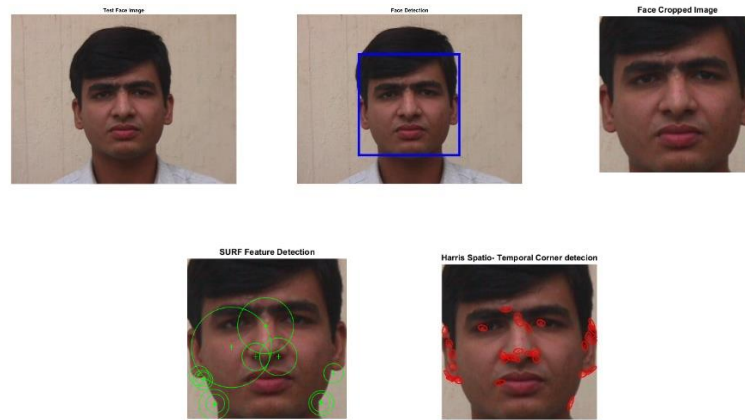


Fig. 2.31 Face processing

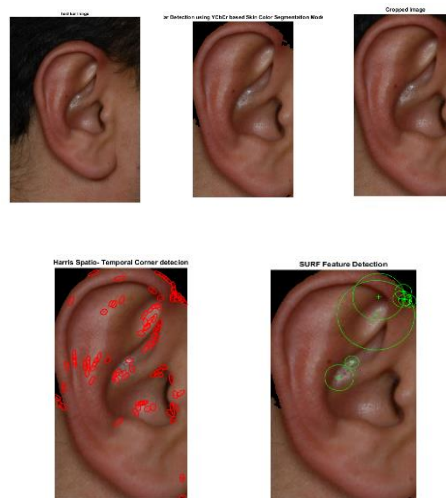


Fig. 2.32 Ear processing

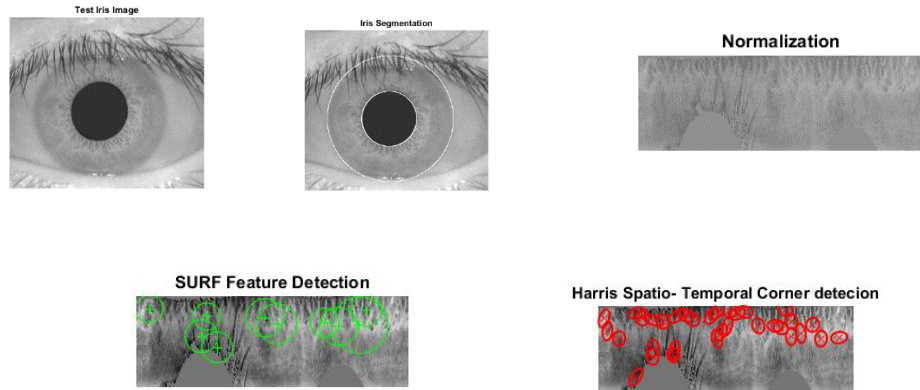


Fig. 2.33 Iris processing

SURF is a fast, reliable and robust algorithm for local, similarity invariant representation and comparison of images which involves two steps: one is feature extraction and the other is feature description. Given a point $p = (x, y)$ in an image I , the Hessian matrix $H(p, \sigma)$ at point p and scale σ , is:

$$S(x, y) = \sum_{i=0}^x \sum_{j=0}^y I(i, j) \quad (2.8)$$

$$H(\mathbf{p}, \sigma) = \begin{Bmatrix} L_{xx}(\mathbf{p}, \sigma), L_{xy}(\mathbf{p}, \sigma) \\ L_{yx}(\mathbf{p}, \sigma), L_{yy}(\mathbf{p}, \sigma) \end{Bmatrix} \quad (2.9)$$

Where $L_{xx}(\mathbf{p}, \sigma)$ etc. is the convolution of the second-order derivative of Gaussian with the image $I(i, j)$ at the point.

$$\sigma_{approximation} = \text{current filter size} \left(\frac{\text{base filter scale}}{\text{base filter size}} \right) \quad (2.10)$$

III. Recognition Stage

(A) Face Recognition

Face images were acquired from gallery and face detection tool has been used for detecting the face in test image. After that detected face was cropped for further processing. After cropping, SURF feature descriptor has been used for feature extraction. For recognition

task, SURF algorithm is selected as it has several types of invariances such as scale; translation, lighting, contrast and rotation. Object detection in images taken under different extrinsic and intrinsic settings is achievable using SURF. Mean feature were calculated from the output of SURF. Multiclass support vector machine (MSVM) and KNN based classification was performed which outperform limited class support vector machine.

(B) Ear Recognition

For this part, ear was cropped by automatic cropping method. Then skin color segmentation model, known as YCbCr based skin color segmentation model was used for detection. After detection, holes filling and filtration process was applied for enhancement of input's quality. For feature extraction, interest points in image were extracted using SURF & HSTCP feature descriptor. Just like other traits finally mean feature was calculated results were forwarded to MSVM and KNN classifiers.

(C) Iris Recognition

A segmentation algorithm is performed on iris images, so that localization of iris area from full eye and isolation with eyelid, eyelash and reflection areas can be obtained. This segmentation is realized using the circular Hough transform, and the linear Hough transform for localizing eyelids. Thresholding is used for segregating eyelashes and reflections. Normalization of segmented iris region is done to overcome dimensional inconsistencies in-between iris region. For normalization modified version of Daugman's rubber sheet model is used, where the iris is represented as rectangular block with fix polar dimensions. After this strongest feature in iris were obtained using SURF feature descriptor. Finally mean feature is calculated from SURF output. For classification, MSVM and KNN classifiers are used for iris recognition.

KNN Classifier

K-Nearest neighbour's algorithm can be used to solve classification problems.

Similarity Measures

1. Euclidean Distance: This is most commonly used distance measure. For two points (x_1, x_2) and (y_1, y_2) the Euclidean distance is given by:

$$\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (2.11)$$

2. Manhattan Distance: Also known as the city block or absolute distance, it is inspired from the structure of Manhattan city. For two points (x_1, x_2) and (y_1, y_2) the Manhattan distance is given by:

$$|x_1 - x_2| + |y_1 - y_2| \quad (2.12)$$

3. Chebyshev Distance: Also known as the chessboard or maximum value distance, for two points (x_1, x_2) and (y_1, y_2) the Chebyshev distance is given by:

$$\max(|x_1 - x_2| + |y_1 - y_2|) \quad (2.13)$$

4. Minkowski Distance: This is a generalized distance measure. All the above-mentioned distances can be obtained from the generalized formula.

$$d(p, q) = (\sum_{i=1}^n |p_i - q_i|^c)^{\frac{1}{c}} \quad (2.14)$$

When $c = 1$, Minkowski = Manhattan, $c = 2$, Minkowski = Euclidean, $c = 3$, Minkowski = Chebyshev

5. Mahalanobis Distance: To calculate the distance between two points in multivariate space, we use the Mahalanobis distance. The Mahalanobis distance is given by:

$$d(x, y) = (x - y)^T C (x - y) \quad (2.15)$$

Here x and y are the vectors of same distribution in multivariate space. C is the inverse of the covariance matrix.

- Find K neighbors that are close to the chosen data point based on the similarity measure used.
- Using majority voting from the k neighbors identify which class the data point belongs to.

2.4.2.2 Results

The experiments were performed on the system with Intel core i5 processor using MATLAB R2021a Software and its image processing toolbox as the simulation platform.

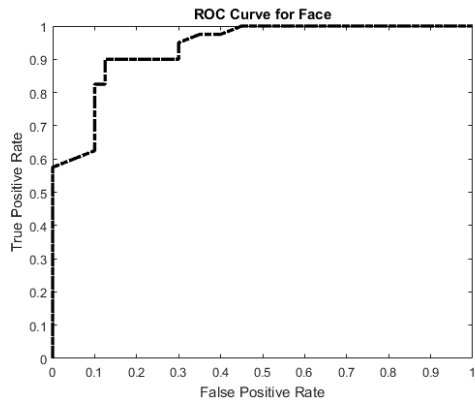


Fig. 2.34 ROC curve for unimodal face

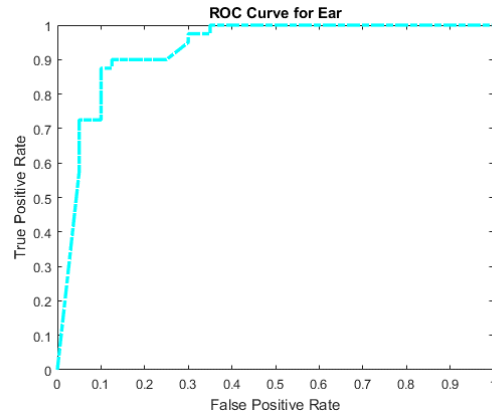


Fig. 2.35 ROC curve for unimodal ear

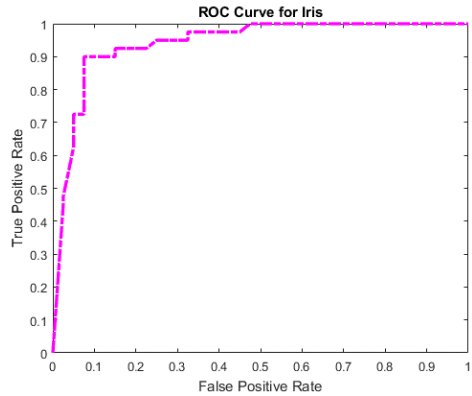


Fig. 2.36 ROC curve for unimodal iris

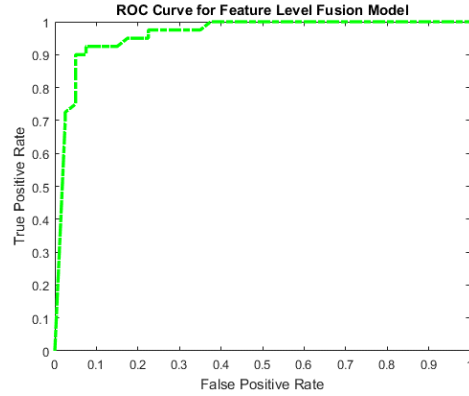


Fig. 2.37 ROC for feature level fusion

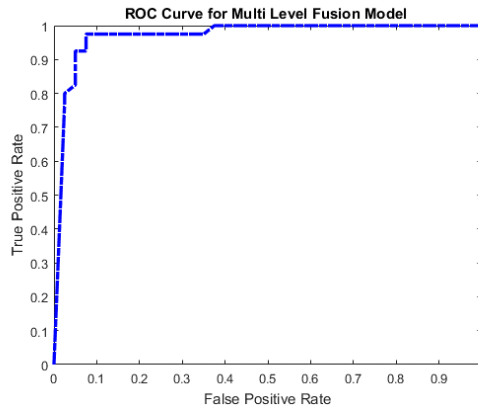


Fig. 2.38 ROC curve for multi-level fusion

For face, ear and iris, unimodal accuracy obtained is 92.4 %, 93.0% , & 94.06% respectively. For feature level fusion some improvement has been noted and result was 94.16%. In proposed method accuracy was improved significantly (95.09%) compared with unimodal accuracies, and it was 2.69%, 2.09%, 1.03% more than face, ear and iris unimodal process respectively. Fig. 2.34 to Fig 2.39 depicts above mentioned details and Table 2.4 & Table 2.5 summaries all methods.

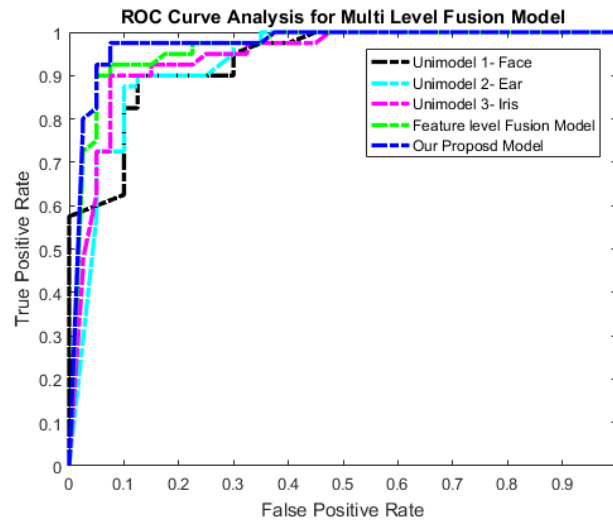


Fig. 2.39 Comparison of unimodal and proposed method

The recall rate, precision and specificity observed were 96.16%, 94.12% & % 93.03% respectively. The performance of multilevel fusion has been reported better in many experiments with constrained databases. The work with multilevel fusion of multimodal biometrics is obtained with increase in complexity but with improved accuracy. This phenomenon is verified in proposed work where accuracy has increased.

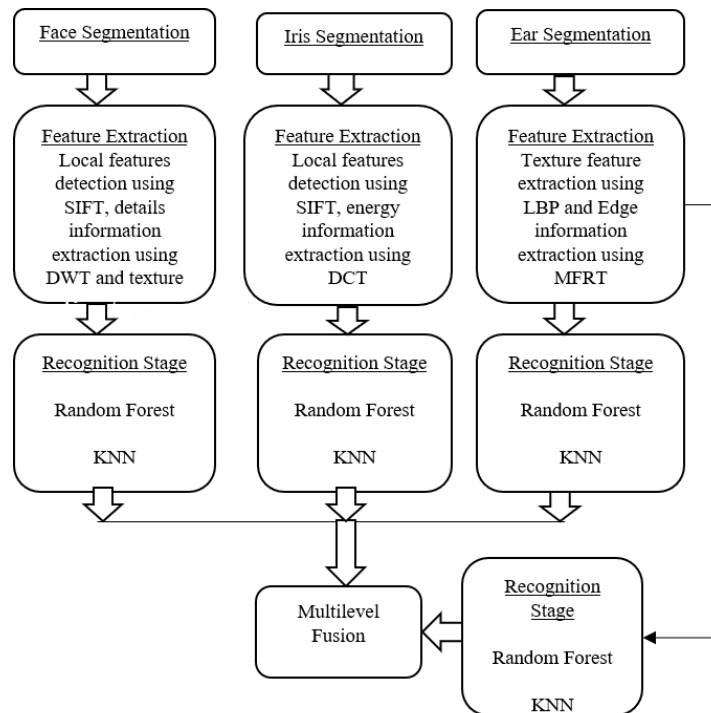
Table 2.4 Accuracy obtained for different experiments.

Parameter	Face	Ear	Iris	Feature Fusion	Proposed Multilevel Fusion method
Accuracy	92.44%	93.00%	94.06%	94.16%	95.09%

Table 2.5 Performance comparison with other methods

Performance Parameters	MBS with RNN classification	MBS with SIFT based feature extraction with KNN	Proposed MBS with SURF based feature extraction with SVM & KNN
Accuracy	90.58%	91.22%	95.09%

2.4.3 Multilevel Multimodal Biometric System Type 3

**Fig. 2.40** SIFT based multimodal biometrics classification using RF & KNN

In this process Scale Invariant Feature Transform (SIFT) algorithm is used for the detection and description of local features in face and ear. Along with SIFT, in face image, Discrete Wavelet Transform (DWT) and Local Binary Pattern (LBP) are also employed for the purpose of feature extraction. For iris image, Discrete Cosine Transform (DCT) is preferred for energy information extraction. LBP and Modified Finite Random Transform (MFRT) are preferred for feature extraction for ear image. Random Forest (RF), K- Nearest Neighbor (KNN) are used for the classification purpose.

SIFT is an algorithm which is basically used for the local feature's detection and description for the digital images. Here descriptors are used for object recognition purpose

to provide quantitative information and these descriptors are basically located as certain key points by SIFT.

2.4.3.1 Proposed Methodology

Multimodal biometric system using multilevel fusion strategy has been developed using face, ear and iris modalities. So three main stages are segmentation stage, feature extraction stage and recognition stage as described below:

I. Segmentation stage

- a) For Test Image-1, here face detection and face cropping operations are performed on the input test image.
- b) For Test Image-II, here ear cropping and than ear cropping operations are performed on the input ear image.
- c) For Test Image-III, Canny edge detection algorithm is applied for getting the necessary edge map from input iris image and than Hough Transform is performed later.

II. Feature Extraction Stage

For segmented output for face image, feature extraction is performed using SIFT, DWT and LBP. SIFT is used for local feature detection purpose, DWT is used for the extraction of information details and LBP is used for texture feature extraction. For segmented output for ear image, feature extraction is performed using SIFT and DCT. Here DCT is used for energy information extraction. For segmented output for iris image, feature extraction is performed using LBP and MFRT. This MFRT is used for edge information extraction.

III. Recognition Stage

RF and KNN are used for classification purpose for all the three modalities. These three output received and the fourth output received from the combination of individual feature extraction stages are finally combined to achieve multilevel fusion. The flow diagram in Fig. 2.40 depicts these steps pictorially.

(A) For Face:

Fig 2.41 to Fig 2.44 depicts the feature extraction steps and histogram for face.



Fig. 2.41 Face pre processing

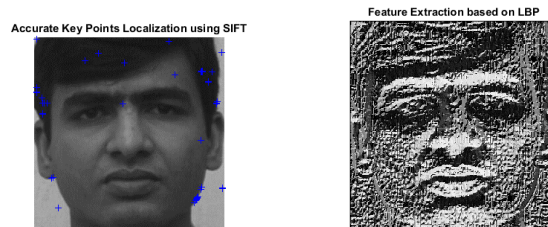


Fig. 2.42 SIFT & LBP features

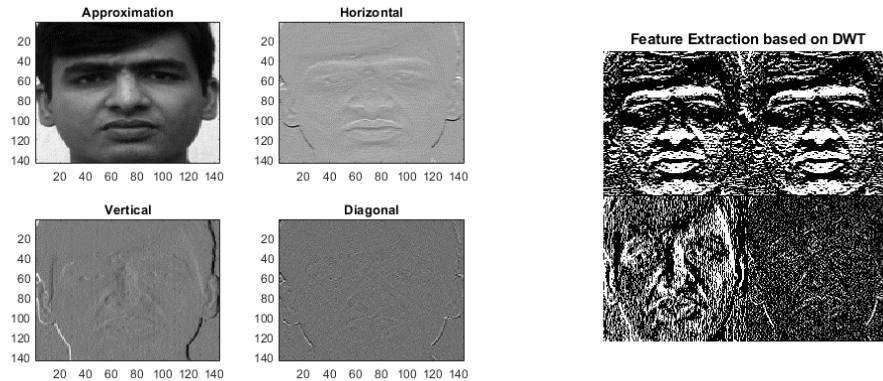


Fig. 2.43 (a & b) DWT based feature extraction

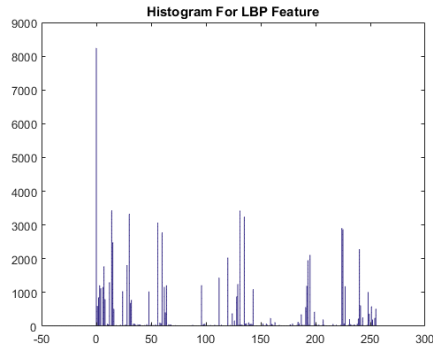


Fig. 2.44 Histogram for LBP feature

(B) For iris

Fig. 2.45 depicts feature extraction steps for iris as test image, iris segmentation, normalisation and feature extraction for iris.

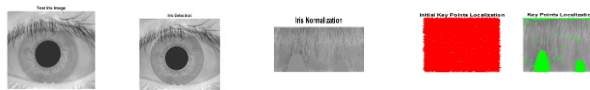


Fig. 2.45 Iris pre processing

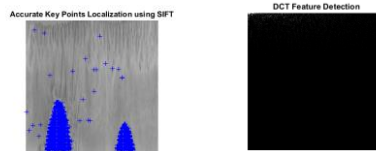


Fig. 2.46 SIFT & DCT output of iris

C) For Ear

Fig. 2.47 depicts feature extraction steps and histogram for ear as test image, ear cropping, detection, and feature extraction

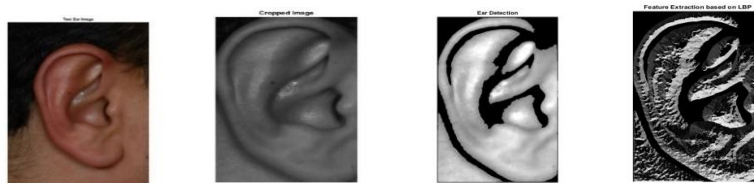


Fig. 2.47 Ear pre processing

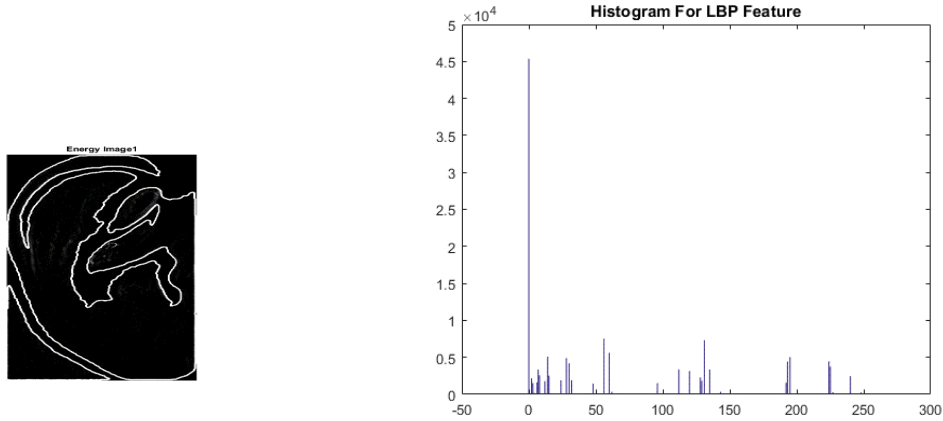


Fig. 2.48 MFRT features and LBP histogram of ear

2.4.3.2 Results

Graphs from Fig. 2.49 to Fig. 2.51 are the ROC curve for the Unimodal Face, Ear, Iris and Fig. 2.52 & Fig. 2.55 presents ROC curve for Feature level fusion and proposed method respctively. Fig. 2.55, Fig. 2.56 depicts the performence paratmetrs of the proposed method and its comparision with other unimodal methods.

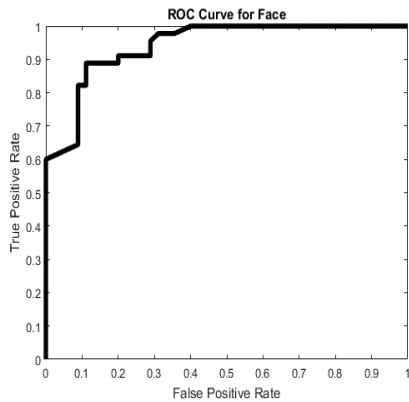


Fig. 2.49 ROC for face (unimodal)

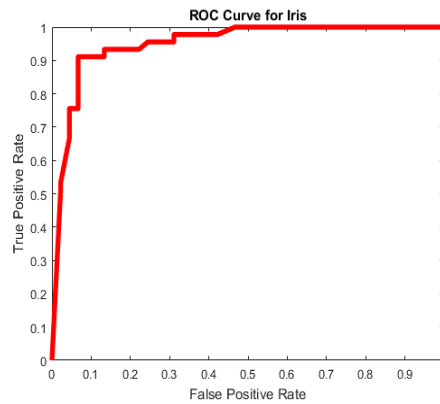


Fig. 2.50 ROC for iris (unimodal)

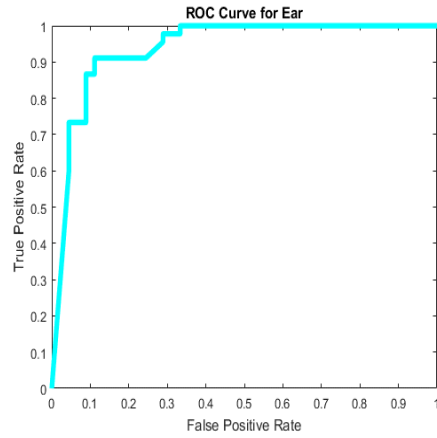


Fig. 2.51 ROC for ear (Unimodal)

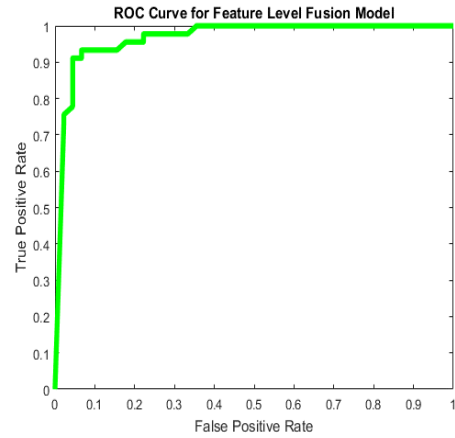


Fig. 2.52 ROC for feature level fusion

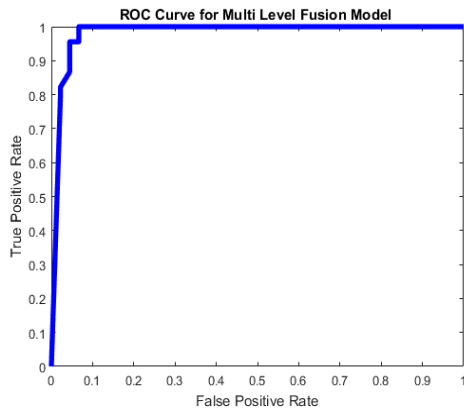


Fig. 2.53 ROC for multi level fusion

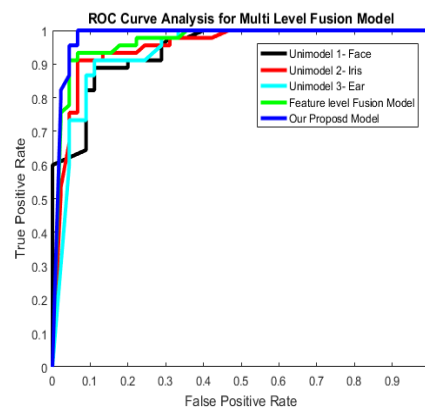


Fig. 2.54 Comparasion of multilevel fusion

Using proposed method results obtained are 98.22% accuracy, 98.00 % precision, 97.14% specificity and 98.99 % recall rate.

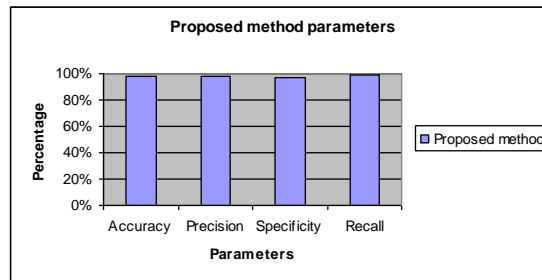


Fig. 2.55 Performence parameters of the proposed method.

Proposed method produced 4.13%, 4.39%, 3.21% and 1.49% more accuracy in comparison with unimodal methods like face, ear iris, and feature level fusion respectively.

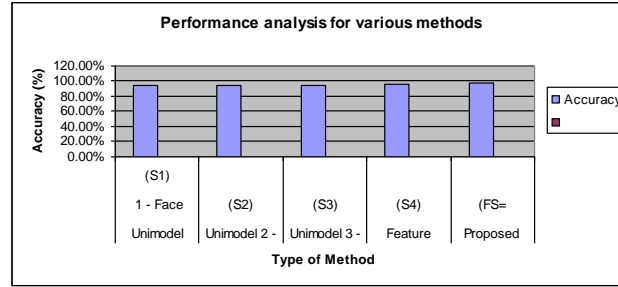


Fig. 2.56 Performance analysis for various methods based on accuracy

Section 2.5 highlights the chapter contribution and importance of multimodal biometric system for advanced authentication system design.

2.5 Summary

Thus multiple feature level and multilevel methods of biometric authentication were executed and compared. The important observations are that feature level outperforms unimodal case in terms of accuracy and improved security. At the cost of additional computation, multilevel approach makes the system more secure and reliable. Implementing traditional approaches is not compatible for advanced platform. Also in the era of convolutional neural network, some initial blocks like preprocessing, feature extraction can be left for the network itself. Application development for remote services and faster & efficient implementation of the algorithm is the way forward.

Chapter 3

Optimal Feature Level Fusion for Authentication

3.1 Introduction

In the multimodal biometric system, the information on each biometric trait is processed independently. The multimodal biometric system can be operated in three modes such as serial mode, parallel mode, and hierarchical model. In serial mode, each modality is analyzed before the next modality is inquired. It is not necessary to capture all the biometric traits at the same time overall time duration can be reduced. This type of model is also named as cascade mode. In parallel mode, the information from multiple modalities is processed together to perform recognition. Then the outputs are processed to develop the final decision. In the hierarchical operational mode, individual classifiers are equated in a tree-like structure [147]. The appropriate fusion scheme is then applied to the processed information, which can be further applied for authentication [148]. The fusion of biometrics information is present in different stages of the recognition system. There are three types of fusion techniques such as (i) feature extraction level; (ii) match score level; and (iii) decision level [149, 150].

Feature level fusion is a popular method as compared to the other fusion because all the computations are done over a single fused feature. So, the time for authentication and the amount of computation power needed is much less due to the single-level operation [151]. The feature-level fusion method is a more effective method than the score-level and decision-level method in multimodal biometric recognition since it fuses more original biometric information into a single vector before the dimensional reduction

procedure [152]. Feature-level fusion has advantages in two aspects. Firstly, it can be applied to eliminate redundant information among the multiple features and secondly it can derive effective discriminate information from multiple feature sets and develop the accuracy of recognition [153]. The fusion of multiple biometrics increased for practical applications due to enhanced recognition performance and a raised level of security [154].

3.1.1 Problem Definition

Some of the baseline or regular limitations of biometric systems observed in previous chapter are listed below.

- In the multimodal based biometric recognition, several modalities are done at various fusion levels such as sensor level, matching score level, and decision level, etc. Multimodal fusion operation at feature level is difficult to generate desirable results due to the potential incompatibility of feature spaces formed by different modalities.
- As per the published literature canonical correlation analysis method is insufficient to reveal the complex and nonlinear correlation relationship between two features sets.
- In the multimodal biometric system, the unknown relationship between the feature space of multiple modalities and the curse of dimensionality problem is a challenge.
- The integration of incompatible feature sets is not easy, and it creates significant impact in computation.
- The limitation of the feature level fusion is that feature sets from the multimodalities are neither accessible nor compatible.
- In the feature level fusion, the concatenation may yield a very large dimensional feature vector due to the presence of noisy or redundant data, thus leading to a decrease in the performance.

These are the main drawbacks of various existing biometric systems, which motivate us to do this research on the multimodal biometric system to achieve better performance.

3.2 Proposed Methodology

The multimodal biometric system is a newly emerged technique that utilizes more biometric modalities in the authentication and identification process. This enables the feature of accruing information from different modalities for enhancing the ability of recognition compared to a single biometric system. In multimodal-based biometric recognition, various works are done at various fusion levels such as sensor level, matching score level, and decision level. When equated to others, the feature level-based methods perform better in personal recognition. Feature level fusion is used to extract features from the same modality or different multimodalities. In existing feature level fusion, concatenating several feature vectors may lead to construct a relatively large feature vector. This increases the computational and storage resources demands and eventually requires more complex classifier design on the concatenated data set at the feature level space. To overcome these problems, an effective feature-level fusion method is proposed in our recommended technique. Here we are considering multimodal biometric for feature level fusion like a fingerprint, ear, and palm. The proposed method has four main processes such as preprocessing, feature extraction, optimal feature level fusion, and recognition. At first, each input image is preprocessed; the image enhancement technique is carried out for overcoming the limitations arising from occultation and background disturbance. Then the resultant output is fed to the feature extraction process. From the palm image, the features are extracted by Gabor features.

For ear image both the shape and texture features are extracted, here we are proposing a modified region growing algorithm to extract the shape features and High Magnitude Sequential Batch (HMSB) operator to extract the texture feature. For fingerprint image, HMSB operator to extract the texture feature. After extracting the features, the resultant output is fed to the next step. The next step is optimal feature level fusion, where the relevant feature is selected using the optimization technique. For selecting the optimal features, the proposed technique uses an optimum grey wolf optimization algorithm. After selecting the relevant features then, the selected features are fused. The final step of

the proposed technique is recognition. For recognition, Multi kernel support vector machine (MKSVM) is proposed. The recommended technique is implemented in the MATLAB platform.

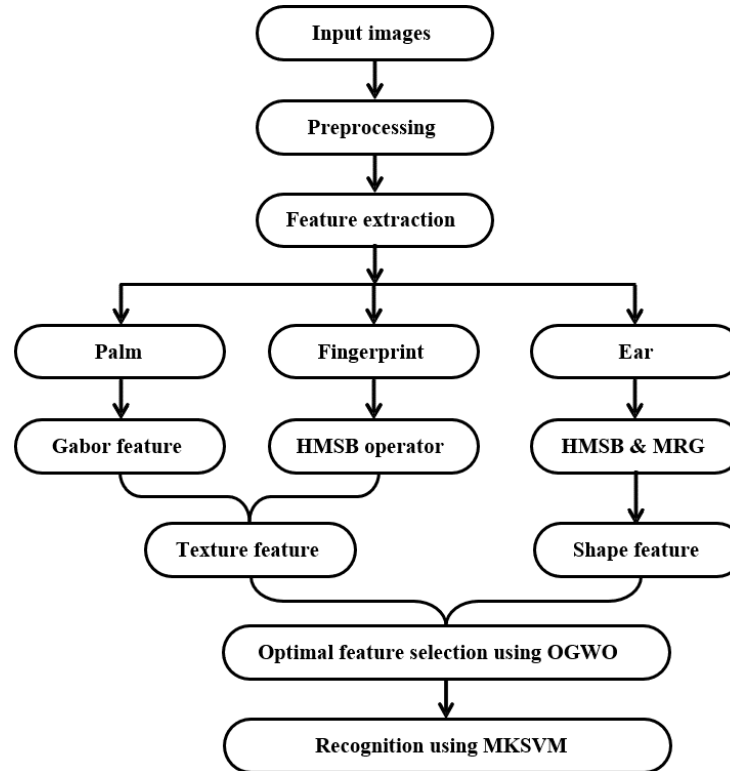


Fig. 3.1 Proposed oppositional grey wolf optimization-based biometric image recognition

3.2.1 Preprocessing

To remove the impact of noise and other factors, and to raise the quality of the image preprocessing operations have been applied on the input images. At first, each input image is preprocessed; Image enhancement technique is carried out for overcoming from the limitations such as occultation and background disturbance. Once this process is finished, the processed images are taken to the next level where feature extraction is executed for further processing.

3.2.2 Feature Extraction

From the palm image, the features are extracted by Gabor features. For ear image both the shape and texture features are extracted, here we are considering a modified region growing algorithm is utilized to extract the shape features and HMSB control to extract the texture feature. For fingerprint image, HMSB operator to extract the texture feature. After extracting the features, the resultant output is fed to the next step. The next step is optimal feature level fusion; here the relevant features are selected using the optimization technique. In this work we have extracted a palm, fingerprint and ear features for optimum feature selections process.

3.2.3 Gabor Feature Extraction from Palmprint

Here we have extracted a palm with the aid of Gabor feature because in the different palm print identification methods, Gabor filter is applied and generated better results as published in the literature. In this work, the Gabor function with certain parameters is applied which is transformed into discrete Gabor filter. To render robustness to brightness, the Gabor filter is turned to zero DC.

Gabor filter consists a series of local spatial bandpass filters. It is mostly used as a powerful texture extractor in computer vision. In the biometric literature, Gabor filter is successfully applied to exploit biometric modalities like iris [28], face [29], fingerprints [30, 31] and palm-prints [32]. A circular 2D Gabor filter is a Gaussian modulated by complex oriented sinusoidal function that captures the spatial and frequency information simultaneously [33].

In the present work, G–L fractional derivative is used as an enhancement module to enhance texture details of palm print ROI. The improved fractional derivative with respect to x direction and y direction is defined as:

$$G_x^\nu f(x, y) = \lim_{l \rightarrow 0} l^{-\nu} \sum_{k=0}^{n-1} \frac{1}{\Gamma(-\nu)} \frac{\pi(k-\nu)}{\Gamma(k+1)} f(x-kl) \quad (3.1)$$

$$G_y^\gamma f(x, y) = \lim_{l \rightarrow 0} l^{-\nu} \sum_{k=0}^{n-1} \frac{1}{\Gamma(-\nu)} \frac{\pi(k-\nu)}{\Gamma(k+1)} f(y-kl) \quad (3.2)$$

Where, γ is a real number and consider $\gamma > 0$.

The palm-print ROI $f(x, y)$ having a size of 128×128 is convolved with $q \times z$ size differential mask. The resulting response is given as

$$I^\gamma(x, y) = \sum_{q=-a}^a \sum_{z=-a}^a w(q, z) f(x+q, y+z) \quad (3.3)$$

The optimal response can be obtained at different γ orders. The images are improved in terms of contours, edges and preserves texture information.

The filters' performances were evaluated by computing the Information entropy (E_i), and Mean gradient (G_{mean}). The information entropy is calculated as given in Eq. (3.4)

$$E_i = \sum_{i,j} p(x_i, y_j) \log_2 \frac{1}{p(x_i, y_j)} \quad (3.4)$$

The higher value of E_i shows that the image contains more visual information.

Mean gradient (G_{mean}) is used to evaluate the image clarity. It reflects the method's ability to create contrast between small details.

The mean gradient is calculated as given in Eq. (3.5)

$$G_{mean} = \frac{1}{PQ} \sum_{x=1}^P \sum_{y=1}^Q \sqrt{\left(\left(\frac{\partial f(x, y)}{\partial x} \right)^2 + \left(\frac{\partial f(x, y)}{\partial y} \right)^2 \right)} / 2 \quad (3.5)$$

Circular 2D Gabor filter is expressed as follows:

$$G_{\sigma, \mu, \phi}(x, y) = gu_\sigma(x, y) \cdot \exp[2\pi j\mu(x \cos \phi + y \sin \phi)] \quad (3.6)$$

Where, $gu_\sigma(x, y)$ is Gaussian function defined as

$gu_\sigma(x, y) = \frac{1}{2\pi\sigma^2} \exp[-(x^2 + y^2)/2\sigma^2]$, $j = \sqrt{-1}$, σ is standard deviation $[0, 1]$, μ is frequency in $[0, 0.5]$ and ϕ is the orientation $[0^\circ - 180^\circ]$. The complex form of Gabor

filter $G_{\sigma, \mu, \phi}(x, y)$ can be decomposed in terms of real part, $R_{\sigma, \mu, \phi}(x, y)$ and imaginary part (for edge detection), $I_{\sigma, \mu, \phi}(x, y)$ as given in Eq.(3.7), Eq.(3.8), Eq.(3.9):

$$G_{\sigma, \mu, \phi}(x, y) = R_{\sigma, \mu, \phi}(x, y) + jI_{\sigma, \mu, \phi}(x, y) \quad (3.7)$$

$$R_{\sigma, \mu, \phi} = gu_{\sigma}(x, y) \cdot \cos[2\pi\mu(x \cos \phi + y \sin \phi)] \quad (3.8)$$

$$I_{\sigma, \mu, \phi} = gu_{\sigma}(x, y) \cdot \sin[2\pi\mu(x \cos \phi + y \sin \phi)] \quad (3.9)$$

To make the illumination response insensitive, the DC component of the Gabor filter is removed as given in Eq. (5),

$$\bar{G}_{\sigma, \mu, \phi}(x, y) = G_{\sigma, \mu, \phi}(x, y) - \frac{\sum_{i=-k}^k \sum_{j=-k}^k G_{\sigma, \mu, \phi}(i, j)}{(2k+1)^2} \quad (3.10)$$

Finally, Gabor transform with robust illumination is defined as

$$O(x, y; \sigma, \mu, \phi) = I(x, y) \otimes \bar{G}_{\sigma, \mu, \phi}(x, y) \quad (3.11)$$

Where, $(2k+1)^2$ denotes the size of the Gabor filter, $I(x, y)$ is the ROI and symbol \otimes is convolution operator. The literature demonstrated that Gabor filter provides the accurate recognition when filter parameters (orientation, variance, center frequency) are suitably chosen [34]. Therefore, amplitude calculation involves the process of both real and imaginary parts

$$E_{\sigma, \mu, \phi}(x, y) = \sqrt{R_{\sigma, \mu, \phi}^2(x, y) + I_{\sigma, \mu, \phi}^2(x, y)} \quad (3.12)$$

The adjusted Gabor filter is convoluted with sub-images. Sample points in the filtered image are coded into two bits by using certain inequalities. By applying this coding method, only the phase information about the palm print image is stored in the feature vector. In the matching process, the two-palm print images are equated by calculating the hamming distance. Here, each feature is conceived as two feature matrices that are real and imaginary. A normalized Hamming distance is applied for palm print matching. For perfect matching, the hamming distance is zero. It provides robustness against varying

brightness and contrast for images. However, they lack the ability of orientation selectivity and representation of main principal lines and wrinkles in the palm print.

3.2.4 HMSB based texture feature extraction from Fingerprint

Here we have extracted a fingerprint kind of texture feature with the aid of HMSB operator and the texture is communicated as the utility of the spatial adjustment in pixel quality i.e., dim qualities, which are helpful in various capacities and have been engaging the devoted convergence of a few examiners [154]. Gordon & Pathak [155] have recognized an effective surface representation prepare [156] known as the selective LBP administrator (Salton) [157] in similar manner.

3.2.5 Shape and Texture Feature Extraction from Ear

Here the texture feature extracted based on HMSB and shape feature-based Modified Region Growing algorithm.

Here we have extracted an ear with the aid of a modified Region growing method, which is one of the popular segmentation methods. In this modified method, we admit an HMSB operator along with region growing. This method initiates with seed pixels and grows the region by adding the neighboring pixels established because of threshold value. When the growth of a region stops, another seed pixel that does not belong to any other region is selected and the process is repeated. The region growing is stopped when all pixels belong to some regions. Region growing segmentation is particularly applied for the delineation of small, simple structures such as tumors and lesions. The various limitations of applying this method are

1. Sometimes, manual interaction is needed to choose the seed point.
2. Sensitive to noise so it produces holes or over-segmentation in the extracted regions. The discontinuity in the extracted image can be removed by applying the Homotopic Region Growing algorithm.

3. Unless image has had a threshold function applied, a continuous path of points related to color may exist, which connects any two points in the image. So, we need find a single threshold to be taken. By using this region-merging algorithm, we will extract texture-based features then it will be optimized by using a fruit-fly algorithm.

Consider the biometric image $A^{in}(x, y)$, which cannot be fed directly as the input for the suggested technique since it is having various types of noises and unwanted things. To remove the noise from the image, the input image $A^{in}(x, y)$ is passed through a Gaussian filter to minimize the noise and get a better image. Passing the image through the Gaussian filter also raises the image quality. After that, we convert the image from the RGB to the Gray model $A(x, y)$, which makes the image fit for region growing process. After the gray level conversion, we extract the region of interest (ROI) region from the image because not all the area of the images is having information required for the texture analysis. Therefore, a Modified region growing (MRG) operation has been used on biometric images to extract the region of interests (ROIs) which contain the abnormalities, excluding the unwanted portion of the image. The modified region growing is a three-step process which comprises of (i) gridding, (ii) selection of seed point, (iii) applying region growing to the point. Procedure for ROI region extraction using a modified region growing procedure is given in Table 3.1.

Table 3.1 Procedure for modified region growing

<p>Input: Input image $A^{in}(x, y)$</p> <p>Output: Regions</p> <p>Start</p> <p>Step 1: Get the input image $A^{in}(x, y)$</p> <p>Step 2: Remove the noise from the image using the Gaussian filter.</p> <p>Step 3: Convert the RGB image into the gray image we obtain $A(x, y)$</p> <p>Step 4: Find the gradient of the Image $A(x, y)$ in both x-axis (A_x) and y-axis (A_y)</p> <p>Step 5: Combine the gradient values applying the formula.</p> $G = \frac{1}{1+(A_x^2+A_y^2)}$ <p>to get the gradient vector \mathcal{G}.</p> <p>Step 6: Convert Gradient vector values from in radians, to degrees to get the orientation values of the pixels of the image.</p> <p>Step 7: Spilt the Image A into Grids G_i.</p> <p>Step 8: Set the intensity threshold T_1^h and the orientation threshold T_0^h.</p> <p>Step 9: for each Grid do</p> <ol style="list-style-type: none"> a) Find the histogram (denoted as Hist) of every pixel H_j in the grid G_i b) Find the most frequent histogram of the G_i^{th} grid and denote it as $Freq_{Hist}$ c) Select any pixel H_j representing to the $Freq_{Hist}$ and assign that pixel as the seed point SP having intensity I_p and orientation O_p. d) For the neighboring pixel having intensity I_N and orientation O_N, check for intensity constraint $I_p - I_N \leq T_1$ and the orientation constraint $O_p - O_N \leq T_0$ e) If both the constraints are satisfied and met, the region is grown to the neighboring pixel. The region is not grown to the neighboring pixel in the other case. <p>Stop</p>
--

3.3 Optimal Feature Level Fusion Gray Wolf Optimization (OGWO)

For selecting the optimal features, the proposed technique uses a Gray Wolf optimization algorithm (GWO). After selecting the relevant features then, the selected features are fused. GWO is developed based on the behaviors of the wolf. The GWO mimics the initiative progression and chasing component of dim wolves. The dim wolves adequately encase a Canidae's segment predecessors and are esteemed as the head predators seeing their course of action at the sustenance's nourishment grouping. They typically delineate a prejudice to detail appropriate as a get together. The pioneer speaks to a male and a female, set apart as alpha, which is for the larger part division in allegation of captivating legitimate assortment screening differing highlights, for instance, the chasing, resting area, time to wake, etc. The determinations arranged by the alpha are acknowledged on to the gathering. The Beta locations to the second grade in the pecking exhibit of the dark wolves. They are, in a general sense, optional wolves that successfully suggested a couple of sponsorship to the alpha in the assortment improving or proportional gathering utility. The omega is the littlest division of the dark wolf pack an immense undertaking as a substitution introduces into the further premier wolves nearly on each event and is allowed to incorporate only the small remains charming after a great banquet by the pioneer wolves. A wolf is established as optional or as delta so as often as possible if it does not fit in with the gathering of alpha, beta, or omega. Because of the reality that these delta wolves needed to reverence the alphas and betas, they have a prospering high extent over the omegas. In our technique, the alpha (α) is esteemed as the most suitable accumulation by a standpoint to recreating judiciously the group pecking arrange of wolves though imagining the GWO. In this way, the second and the third most great arrangement are beta (β) and delta (δ) freely. The leftover cheerful arrangement is seen to be the omega (ω). In the GWO approach the chasing (improvement) is directed by then α , β , δ and ω . The systematic procedure of GWO based feature selection is explained below.

Step 1: Initialization: Solution initialization is an important process for all optimization problems. Initially, the solution is randomly generated. The extracted features are given to the input of feature selection. The parameters of GWO namely, a , A , and C also initialized.

Step 2: Fitness Evaluation: After the solution initialization, the fitness of each solution is calculated. The maximum accuracy is considered as the fitness of this process. The fitness function is given in Eq. (3.13)

$$Fitness = \max(accuracy) \quad (3.13)$$

Step 3: Separate the Solution Based on Fitness: Based on the fitness the solution is divided. The first-best fitness is denoted as d_α , the second-best fitness is d_β , and the third finest fitness results d_δ .

Step 4: Encircling Prey: The tracking is aimed at by, β , δ and ω to tag the length of these three.

$$d(t+1) = d(t) + \vec{A} \cdot \vec{K} \quad (3.14)$$

$$\vec{K} = |\vec{C} \cdot d(t+1) - d(t)| \quad (3.15)$$

$$\vec{A} = 2\vec{a}r_1 - \vec{a} \text{ And } \vec{C} = 2r_2 \quad (3.16)$$

Where; t depicts the iteration number, $d(t)$ corresponds to the prey position, A and C depicts the coefficient vector, ' A is linearly decreased from 2 to 0, r_1 and r_2 '.

Step 5: Hunting: We presume that the alpha (best competitor arrangement), beta, and delta incorporate the upgraded data about the plausible position of the casualty with a specific end goal to mirror precisely the following exercises of the dim wolves. Because a result, we amass the soonest three finest results fulfilled up to now and needed the further investigate the middle person (counting the omegas) to change their posture because of the course of action of the finest investigate arbiter. For replication, the novel result is $d(t+1)$ unsurprising by the formulae expressed underneath.

$$\vec{K}^\alpha = |\vec{C}_1 \cdot d_\alpha - d|, \quad \vec{K}^\beta = |\vec{C}_2 \cdot d_\beta - d|, \quad \vec{K}^\delta = |\vec{C}_3 \cdot d_\delta - d| \quad (3.17)$$

$$d_1 = d_\alpha - \vec{A}_1 \cdot (\vec{K}^\alpha), \quad d_2 = d_\beta - \vec{A}_2 \cdot (\vec{K}^\beta), \quad d_3 = d_\delta - \vec{A}_3 \cdot (\vec{K}^\delta) \quad (3.18)$$

$$d(t+1) = \frac{d_1 + d_2 + d_3}{3} \quad (3.19)$$

It can be tried that the closing area would be in an easygoing position encompassed by a circle by the area of alpha, beta, and delta in the investigate hole. By means, alpha, beta, and delta figure the area of the casualty, and further wolves modernize their area unpredictably in the district of the casualty.

Step 6: Attacking Prey (exploitation) and Search for Prey (exploration): Examination and usage are guaranteed by the versatile evaluations of ‘a’ and ‘A’. The versatile estimations of confinement n and A grant GWO to proficiently transformation among examination and use. By declining ‘A’, half of the cycles are consistent to examination ($|A| \geq 1$) and the further half are concentrated on use ($|A| < 1$). The GWO contains two premier restrictions to be acclimated (A and C). However, we admit saved the GWO calculation as easy as achievable through the littlest number of the agent to be acclimated the technique will be tenacious anticipating the best precision is procured. Finally, the finest characteristic is selected and supply to the extra system.

3.3.1 Recognition by Multi-kernel SVM

The multi-kernel SVM for dealing with multi-sources has two situations. First, one is applying the kernel for all sources. In this case, fusion is twofold. The second one is that each input has a relevant kernel to obtain the result. The kernel function is determined as follows,

$$K(X_i, X_j) = \langle \phi(X_i), \phi(X_j) \rangle \quad (3.20)$$

Where,

$X_i, X_j \rightarrow$ input vector

$\phi \rightarrow$ is a map to transform source data from input space to feature space.

The kernel functions have various forms such as linear, polynomial, or Gaussian etc. The scalar value found from the kernel function is equal to the dot product of transformed vectors in feature space. The kernel function is applied to evaluate the distance or describe the similarity of high dimensions to some degree.

The decision function in classification of each volume with the best parameter set has a form of

$$f_n(X) = \sum_{m=1}^M \alpha_m y_m K_n(X, X_m) + b \quad (3.21)$$

α_m - is a weight series

y_m - is the label of sample X_m

M - sample points

n^{th} kernel is used in learning.

b - constant coefficient.

In a two-class multi-kernel-based segmentation problem, synthesize Eq. (3.21) and the number of sources N , the final decision function is determined as the

$$f(X) = \sum_{n=1}^N \beta_n f_n(X) + b_t \quad (3.22)$$

$f_n(X)$ -decision function is defined as in Eq. (3.22) relevant to a certain volume; X is the input data needed to classify.

β_n -is another weight series to show the effect of each volume to find the l result.

b_t - is another scalar coefficient similar to b .

The smaller β_n is, less proportion $f_n(X)$ takes in the final decision. When $\beta_n = 0$, the corresponding $f_n(X)$ does not influence the process of clustering.

In our work we use Gaussian kernel functions is chosen to form multi-kernel.

$$K_q(X, X_i) = \exp\left(-\frac{\|X-X_i\|^2}{2\sigma_q^2}\right) \quad (3.23)$$

q - the number of kernels

σ_q - representing the standard deviation of each kernel. Different parameters σ construct various kernels. This method deals with multi-input data and produces accurate results. The performance of our research can be estimated and established on accuracy.

3.4 Result and Discussion

The proposed method uses the Oppositional grey wolf optimization algorithm and LQ algorithm. Our experiments are implemented in MATLAB platform.

3.4.1 Dataset Description

In this work, following three types of datasets have been utilized:

- (1) Ear Database: IIT Delhi Ear Database version 1.0,
- (2) Palm Database: CASIA Palm Print Image Database (or CASIA-Palm print for short)
- (3) Fingerprint Database: CASIA Fingerprint Image Database Version 5.0 (or CASIA-FingerprintV5)

3.4.2 Evaluation Metrics

The performance of the proposed methodology is analyzed in terms of accuracy, sensitivity, and specificity. Fig.3.2 shows our experiment result of the proposed work is described below.

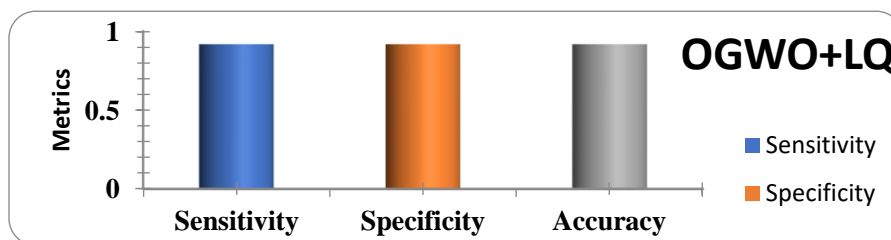


Fig. 3.2 Performance measures for the proposed methodology

3.4.3 Comparative Analysis

The proposed technique uses an oppositional grey wolf optimization algorithm and an LQ algorithm for optimal selection. After that, the relevant selected features are fused. The final step of the proposed technique is recognition, for recognition Multi kernel support vector machine (MKSVM) is proposed. The comparison outcomes of proposed and existing methods are presented in the following Table 3.2.

Table 3.2 The evaluation measures result of the proposed and existing methods

Measures	Sensitivity	Specificity	Accuracy
Proposed OGWO+LQ	0.906	0.912	0.91667
Existing OGWO+L	1	0.58333	0.79167
Existing OGWO+Q	1	0.66667	0.83333
Existing GWO+LQ	0.91667	0.83333	0.875
Existing GWO+L	0.83333	0.58333	0.70833
Existing GWO+Q	0.91667	0.58333	0.75

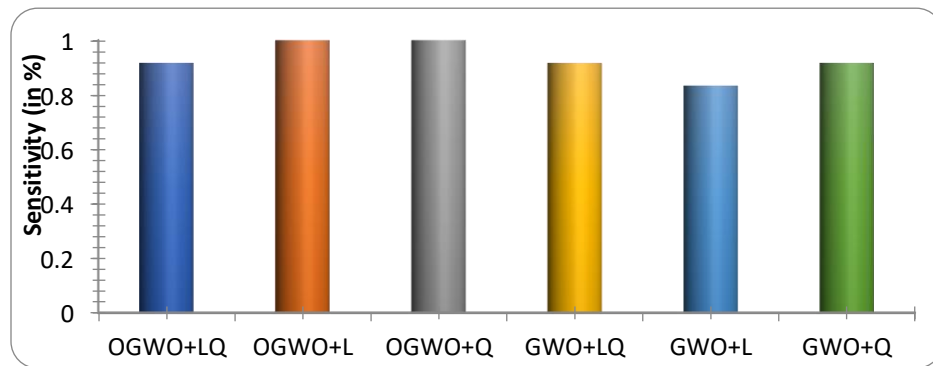


Fig. 3.3 Graphical representation for comparison of proposed and existing sensitivity measures

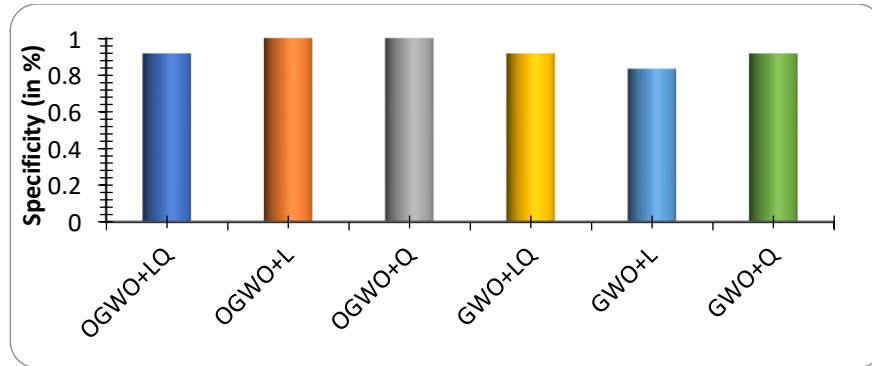


Fig. 3.4 Comparison of proposed and existing specificity measures

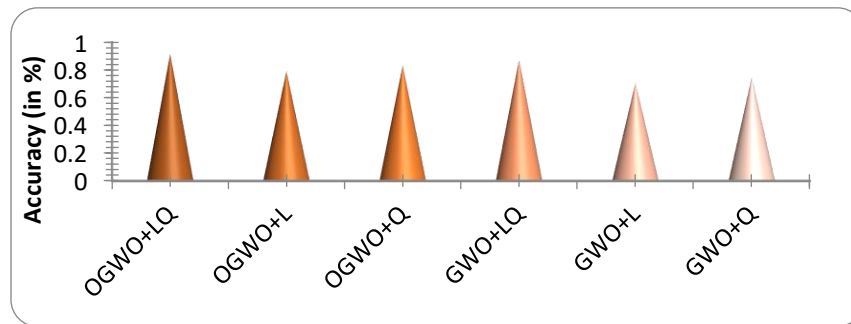


Fig. 3.5 Comparison of proposed and existing accuracy measures

In the proposed OGWO+LQ technique is equated to the existing technology to prove our suggested technique would give better performances. Graphical representations for comparison of the proposed and existing methods are present in the above figure. Here we have taken a comparison measure such as Sensitivity, Specificity, and accuracy. In the proposed OGWO+LQ technique, the sensitivity measure is 0.906. The sensitivity measures of existing OGWO+L, OGWO+Q, GWO+LQ, GWO+L, GWO+Q are 1%, 1%, 0.91667%, 0.83333% and 0.91667% percentages. In the proposed OGWO+LQ method, a specificity measure is 0.912. The specificity measures of existing OGWO+L, OGWO+Q, GWO+LQ, GWO+L, GWO+Q are 0.58333%, 0.66667%, 0.83333%, 0.58333%, 0.58333%. In the proposed OGWO+LQ method, the accuracy measure is 0.91667%. The accuracy measures of existing OGWO+L, OGWO+Q, GWO+LQ, GWO+L, GWO+Q are 0.79167%, 0.83333%, 0.875%, 0.70833% and 0.75%. When compared to existing methods and the proposed method, the proposed OGWO+LQ technique gives better

sensitivity, specificity, and accuracy results effectively. The analysis based on the ROC curve is presented in Fig. 3.6 and 3.7.

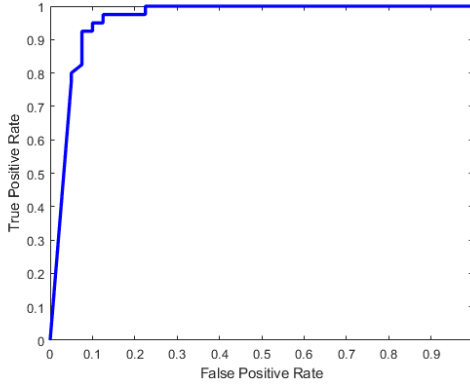
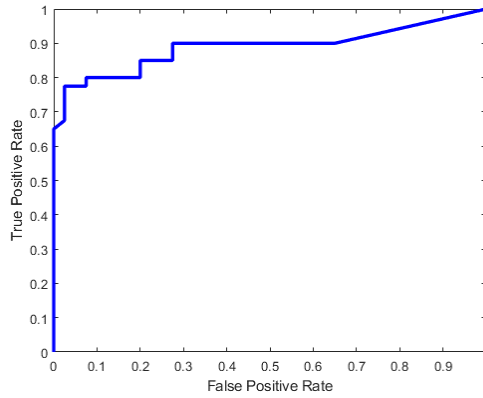
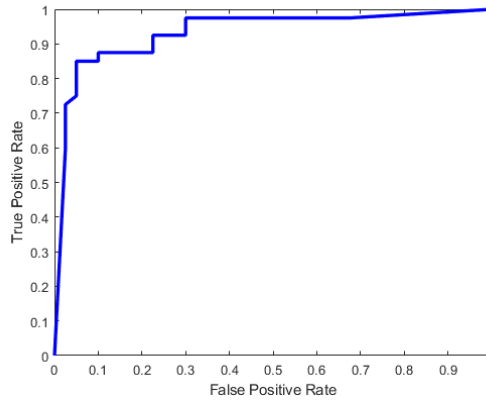


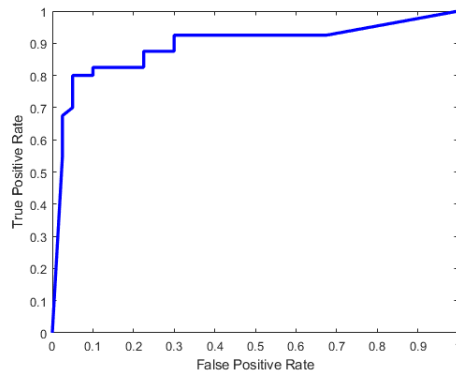
Fig. 3.6 ROC curve of OGWO+LQ



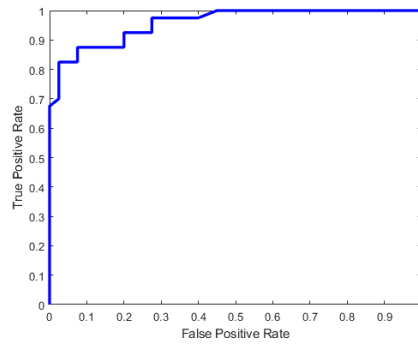
(a) ROC curve-GWO+L



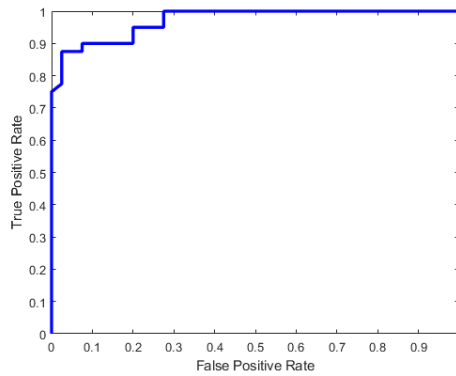
(b) ROC curve -GWO+LQ



(c) ROC curve-GWO+Q



(d) ROC curve -OGWO+L



(e) ROC curve-OGWO+Q

Fig. 3.7 ROC curve analysis for existing researches (a) GWO+L (b) GWO+LQ (c) GWO+Q (d) OGWO+L (e) OGWO+Q

From the ROC curve, it is observed that the proposed technique obtained 91.6% accuracy when compared to existing research. The oppositional learning behavior of the grey wolf optimization algorithm and optimal feature selection will lead to better performance in the recognition process. The other methods produced accuracy such as 79%, 83%, 87.5%, 70%, 73% etc. Thus, the proposed methodology will be used in human authentication process effectively.

3.5 Summary

In the proposed methodology, we suggested an effective feature-level fusion method for multimodal biometric recognition system. We considered the multimodal biometric feature level fusion of a fingerprint, ear, and palm. In the proposed method, we did four main processes such as preprocessing, feature extraction, optimal feature level fusion, and recognition. We used a modified region-growing algorithm for extract the shape features and we used HMSB operator for extracting texture features. Moreover, we selected the relevant features with the help of the optimization technique. For selecting the optimal feature, we used the OGWO+LQ algorithm. In final we proposed recognition, for the recognition we used the Multi kernel support vector machine (MKSVM) algorithm. The experimental results and comparative analysis demonstrate that the proposed method effectively gives better sensitivity, specificity, and accuracy results than other existing methods. In proposed method accuracy is almost 15% to 18% higher than the other reported methods.

Optimal feature level fusion method has shown better results, so it shows path forward to implement similar process for live authentication. Continuous authentication has several applications of high importance, particularly in pandemic time education and business world are facing problems for candidate recognition during the continuous user authentication is a challenge for multimodal biometrics systems in limited resources and constrained environment. As it was observed in the literature, E proctoring through multimodal biometric system is a great application area, so it is taken as next step of our work.

Chapter 4

Continuous Biometric User Authentication for e-Proctoring

4.1 Introduction

Client confirmation depends on "something the client knows" for many years [134]. This strategy is exceptionally famous in personality validation, yet research has demonstrated that PINs and passwords don't offer adequate insurance. A validated record and secret key are needed when you sign in to the standard login framework. Nevertheless, beneath this validation system, the mechanism can just distinguish the client from the login data. It is obscure who is utilizing the framework [135]. One disadvantage of the personal computer (PC) validation framework is that when a client leaves a space at short stretches, others can sign into the PC and access a particular archive or demonstrate that they are approved clients to reestablish data. Such security weaknesses are not satisfactory in applications containing delicate information, such as bank monetary records or individual client data, military, industry, and exchange protection [136], [137]. Clients can physically bolt the screen before leaving or reusing the terminal to forestall this harm under typical login conditions. That causes many issues for the client, particularly when the client is occupied with doing different things. Clients may sometimes skip the log-off cycle to try not to rehash the log-off and re-login measure. So, data assurance is no more there. Nonetheless, these terms don't exist in the uninvolved persistent verification framework [138].

Biometrics is the way toward recognizing an individual dependent on their physical or social guidelines. The manual process incorporates facial, unique marks, iris, and

calculations, while biometric designs incorporate voice, mark, keystroke, and walk. Biometric validation has numerous points of interest over conventional verification techniques [139]. Replicating biometric properties is hard, and also sharing is not possible, so it is considered exceptionally secure. Be that as it may, for current biometric frameworks, psychological information can cause time-dependent changes in the characteristic and information levels [140]. In expansion, the exactness of these biometric frameworks is influenced by enormous intra-class contrasts, uniqueness, and non-all-inclusiveness [141]. Actual biometrics needs sumptuous equipment to be reliable and vivacious against falsification assaults, accordingly, increasing the cost the expense of the gadget.

At last, a great deal of these biometric frameworks needs active client participation, which aggravates the clients [142]. Biometrics-based error proof confirmation and approval (BIA: Biometric Information Assurance) systems save the customer's reference Biometric Templates (BT). If the template is breached, it will allow the spillage of customer's private data, allowing competitors to use their custom characters and evade system settings [143], [144]. For example, in 2015, inadvertently the federal government wing of personnel disclosed the 5.6 million fingerprints of government personnel, which created the information risk. Encryption and cancellable biometric kind of methods are used to guarantee BTs security. Security factors for biometric template, is the challenging field of study [145]. Some key challenges are: (a) The BTs should be unscrambled to facilitate required level of security, (b) The translating keys related with the mixed designs ought to be taken care of very efficiently (reasonably and really), and (c) Encrypted configurations may be compromised if the system is accessed by unwanted users. Since BTs require lifetime maintenance, it is essential to pick structures and frameworks that are suitable for a significant long an ideal opportunity (for instance, 40-50 years), which can be outstandingly good lifespan [146].

Despite the choice to complete the test with a mystery word, security components should be viewed when offering web tests. It isn't absurd to acknowledge that one subject can

purposely give their mystery key to another, considering that the other individual will take the appraisal for the understudy. An evaluation system should be set up to choose if an understudy is enlisted for the assessment. While writing online tests from a remote setup, it is very tough to identify the person who is appearing in the exam [147]. One way to achieve this is with a biometric structure that controls the components of keystrokes of the up-and-comer appearing for the appraisal. The characteristics of keystroke components are assorted for each individual and are considered solely and aggregately. By assessing flight time or the advancement time from one "key" to another event, the customer can make a profile by forming a "signature". By differentiating the components of pressing this key analyzer button, we can choose if the customer is enlisted [148], [149].

The front-most favorable position of the online test is that far away up-comers can do it, the appraisal of the suitable reactions is completely mechanized in the multiple-choice question (MCQ question), and other article-type questions can be surveyed actually or thus, dependent upon the nature and necessities of the requests. Additionally, online tests can be taken at whatever point, while the non-attendance of test undertakings doesn't occur in standard test conditions, duplication, and course of action. The cost of online tests is particularly less stood out of normal test conditions [150]. In web learning, experimentation is gotten together with educating and learning segments. Under online tests, there is no quick contact with understudies, instructors, or chiefs, so prosperity is critical in a web learning environment. The possibility of the online learning environment will shield them from various security perils. Online tests, an essential piece of the learning environment, are extraordinarily participatory, provoking character changes and more dangerous attacks [151]. One of the critical targets of understudy affirmation is to ensure certifiable participation with solitary understudy during the online appraisal. A customer ID and a considerable mystery word are insufficient to check an understudy's character on the web. Online decisions and the online security cycle can help you with murdering coercion. We use biometrics to help the decency of security control, approval, and online assurance measure. Understudy e-perception uses fingerprints and cameras to thwart cheating and

misleading with [152, 153]. The main objectives of the proposed technique are given as follows:

- To design and analyze a technique to authenticate E-examinations continuously and transparently.
- To design an optimal biometric system for examiner authentication.
- To analyze the reliability of the biometric system through different test vectors.
- To propose a hybrid optimization technique to enhance the security in online examination platforms.

4.2 Problem Methodology

We proposed a fitting practical adjustment and grouping calculation to recognize facial pictures utilizing diverse informational indexes in past work. Energetic and LTP capacities are being used to extricate capacities from de-noised facial images. Firefly's high-level streamlining calculation is used to choose ideal picture vectors. By selecting the ideal highlights from the chosen highlights, you will get a bunch of extraordinary highlights. These ideal properties are depicted utilizing the deep belief network (DBN) classifier. The DBN classifier accomplished 98.92% exactness in the ORL data set, 97.92% precision in the essential data set, and 98.12% exactness in the FASSEG data set.

As a rule, the primary test in web-based learning is the trouble in guaranteeing the believability of distance inspectors in taking on web tests. Nonetheless, in numerous nations, the decision of online motivation is restricted. Furthermore, limitless validation is significant for some online applications that need to follow a client's personality all through a meeting, not just toward the start of the session [154-155]. Such a persistent ID can be accomplished utilizing any biometric test; Traditional client accreditations, such as username and secret phrase check, approve "login" confirmation techniques [156]. Online sampling methods can lead to cyber-attacks, leading to unauthorized access to DoS or sensitive information. To prevent such attacks and avoid the risk of a fake person, we used intelligent techniques of continuous detection. We offer an effective continuous biometric

user authentication system for online tests (CBUA-OE) for further improvement. The main contributions of the proposed CBUA-OE technique are as follows:

- First, we extract different soft and spectral-domain biometric features such as mouse, touch, and keystroke from the real-time data set. We developed a modified wolf optimization (MWO) algorithm to compute matching score values with the help of different soft biometric weights.
- Second, we illustrate an optimal feature fusion (OFF) algorithm to fuse optimal weight features of multiple biometric responses.
- Then, a hybrid Lookup Based Convolutional Neural Network (LCNN)-Slap swarm optimization-based classifier used for continuous user authentication classifies examiners' presence and interaction.

4.2.1 System Model of Continuous Biometric User Authentication System for Online Examinations (CBUA-OE) Technique

We used a specific method for generating tester contacts/contacts from biometric data statistics in this segment. Fig. 4.1 shows the process of the planned procedure. User verification is a significant safety confront in an online exam environment. Login is complicated because it is impossible to identify users in an online test visually. In this method, feature extraction is used to calculate the value of the usable point using different soft biometric weights using the MWO method. Several issues need to be considered to create an effective biometric authentication system, especially the acquisition and Classification of features. During taxonomy, individual biometric data are converted into the highest vectors of nature. Biometric Taxonomy is a classification process where traits are classified as per the features. Features are extracted by different methods and at different levels of processing. In the feature-level method, extraction happens before matching and directly from the pre-processed traits. So, the information content is more and so as the dimension of feature sets. Here “nature” indicates the unique characteristics of the particular modality.

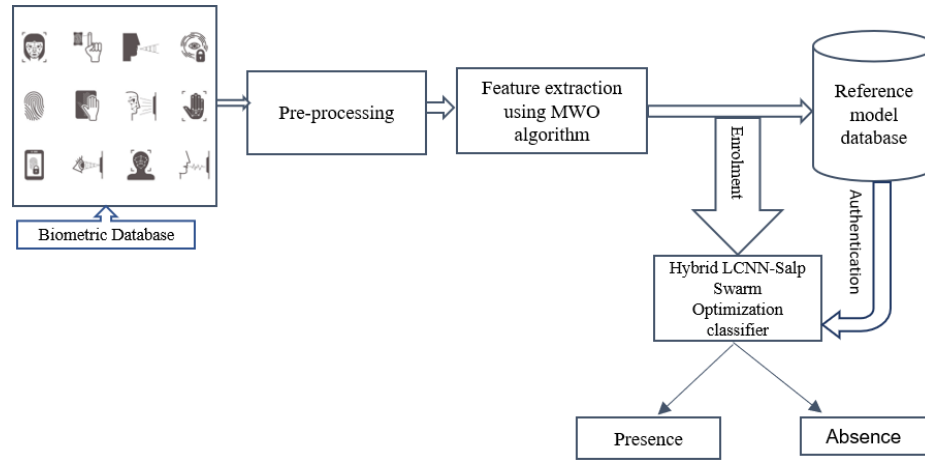


Fig. 4.1 Proposed efficient continuous biometric user authentication system for online examination

The proposed biometric authentication scheme consists of five stages: data collection, data denomination, feature extraction, optimization, and classification.

4.3 Continuous Biometric User Authentication System for Online Examinations (CBUA-OE)

4.3.1 Optimal Feature Extraction Using the Modified Wolf Optimization Algorithm

This section discusses the modified wolf optimization (MWO) algorithm employed for the optimal selection of feature sets. Here, the initial sets of features are fixed randomly. The best set of features can be found by reducing the range and position of the sensitive spot. The modified wolf optimization algorithm is designed by combining two new optimization methods, the Weil optimization algorithm and the Gray Wolf optimization algorithm.

4.3.1.2 Objective Function

Here, the modified wolf optimization algorithm is employed to pick the finest feature set required for the authentication by system. Here, the objective function of the proposed optimization problem is to select the maximum number of compatible feature sets. So the fused feature set can be of optimized size. The objective function is given as follows,

$$\max(L) = \min(B_R(n), \sum_{j=1}^A S_R^j(n)) \quad (4.1)$$

Where L is the maximum length between feature sets s required, $B_R(n)$ represent the number of features sets in having large distance; $S_R^j(n)$ number of datasets close A^j ; and $j = 1, 2, \dots$ Adenotes the number of sets. To achieve the above objective, the following conditions must be satisfied.

$$S_R^j(n) \geq L^j \quad (4.2)$$

$$B_R(n) \geq \sum_{j=1}^A S_R^j(n) \geq \sum_{j=1}^A L^j \quad (4.3)$$

In MWO the heads are made up of males and females known as Alpha, responsible for hunting, sleep, waking, and other decisions. Choices made by Alpha are permissible on the board. Beta and Delta Gray refer to classes two and three in the specific structure of wolves. In short, they are extra wolves with enough help to set up an alpha or group presentation [160]. The remaining wolves represent omega, which is a small group of gray wolves. Alpha, beta, and deltoid mega wolves are responsible for monitoring (optimization) the MWO process.

4.3.1.3 Mathematical Model of Modified Wolf Algorithm (MWO)

In MWO, the most significant inspiration is to surround a prey by leadership through α , β and δ , which can be methodically obtained as below:

$$M(k+1) = M(k)C.P \quad (4.4)$$

In the above equation (5), P can be given as,

$$P = |V \cdot M_l(k) - M(k)| \quad (4.5)$$

Here, M represent the modified wolf position, M_l is the prey position, C, V are the coefficient vectors, and the number of iterations is defined by ' k '. The coefficient vectors C and V can be obtained by the equation below:

$$C = 2c \cdot w_1 - c \quad (4.6)$$

$$V = 2 \cdot w_2 \quad (4.7)$$

where 'c' will be linearly decreased from 2 to 0 and w_1 and w_2 are the random vectors from [0,1]. The parameter 'c' is updated in every iteration within the range from 2 to 0 according to,

$$c = 2 - k \left(\frac{2}{K} \right) \quad (4.8)$$

At this point 'K' denotes the total number of iterations allowed. The updating of wolf's position based on first three best solutions can be obtained as beneath:

$$M_1 = |M_\alpha(k) - C_1 \cdot P_\alpha| \quad (4.9)$$

$$M_2 = |M_\beta(k) - C_2 \cdot P_\beta| \quad (4.10)$$

$$M_3 = |M_\delta(k) - C_3 \cdot P_\delta| \quad (4.11)$$

where, P_α , P_β and P_δ are obtained as follows:

$$P_\alpha = |V_1 \cdot M_\alpha - M| \quad (4.12)$$

$$P_\beta = |V_2 \cdot M_\beta - M| \quad (4.13)$$

$$P_\delta = |V_3 \cdot M_\delta - M| \quad (4.14)$$

Based on the above equations (12), (13) and (14), the explanation for next iteration will be:

$$M(k + 1) = \frac{(M_1 + M_2 + M_3)}{3} \quad (4.15)$$

The process of updating of wolf position takes place incessantly until the maximum iteration is achieved. The working function of the proposed optimal feature selection is given in Algorithm 1.

4.3.2 Weight Fusion Using Optimal Feature Fusion (OFF) Algorithm

Competition between team members for food leads to distribution and group mergers. A food deficit group is divided into OFF all the groups that show a wholly separated social structure when there is a food deficit group. OFF has seven phases: Startup Phase, Local Leader Phase, Global Leader Phase, Local Leader Learning Phase, Global Leader Learning Phase, Local Leader Completion Phase, and Global Leader Phase. These steps are described as follows:

4.3.2.1 Formation

An initial N optimal feature fusion algorithm population is generated, where FFO_i represents the i th FFO in the population. Each FFO_i is initialized as shown in equation (16):

$$FFO_{ij} = FFO_{minj} + \varphi * (FFO_{maxj} - FFO_{minj}) \quad (4.16)$$

Where FFO_{minj} and FFO_{maxj} are lower and upper limits of the search space in the j th dimension, and φ is a consistently dispersed random number in the range [0, 1]. When φ is 0, it will produce FFO_{minj} as optimum solution. If it is 1, the optimum value will deviate from the minimum level.

4.3.2.2 Feature Phase

All spider monkeys, which are feature vectors in our case, are updated at this point based on the experience of their local leader and local team member—the FFO Exercise checks to update the FFO level to a new level. If the exercise is high, the FFO will not change its position. Here, the level update process is given in equation (17):

$$FFO_{newij} = FFO_{ij} + \varphi * (LL_{kj} - FFO_{ij}) + \psi * (FFO_{rj} - FFO_{ij}) \quad (4.17)$$

Where FFO_{ij} is the j th dimension of i th FFO , LL_{kj} represent the k th local leader of that group, and FFO_{rj} is r th FFO chosen illogically within k th group in j th size such that $r \neq i$ and ψ is consistently distributed random number in the range [-1,1]

4.3.2.3 Environmental Phase

Each FFO uses the knowledge of a global leader and the experience of neighboring points to update their position and find the best solution. The level update equation at this stage is as follows:

$$FFO_{newij} = FFO_{ij} + \varphi * (GL_j - FFO_{ij}) + \psi * (FFO_{rj} - FFO_{ij}) \quad (4.18)$$

Where GL_j is the location of group leader in j^{th} dimension and $j = \{1, 2, 3, \dots, D\}$ chosen randomly.

4.3.2.4 Region-Wise Phase

Unless any local leader is reorganized to a point known as the local leader limit, all members of this group will renew their positions through random startups or using the experience of a global leader. This is the confusion ratio given in equation (4.19):

$$FFO_{newij} = FFO_{ij} + \varphi * (GL_j - FFO_{ij}) + \psi * (FFO_{rj} - LL_{ij}) \quad (4.19)$$

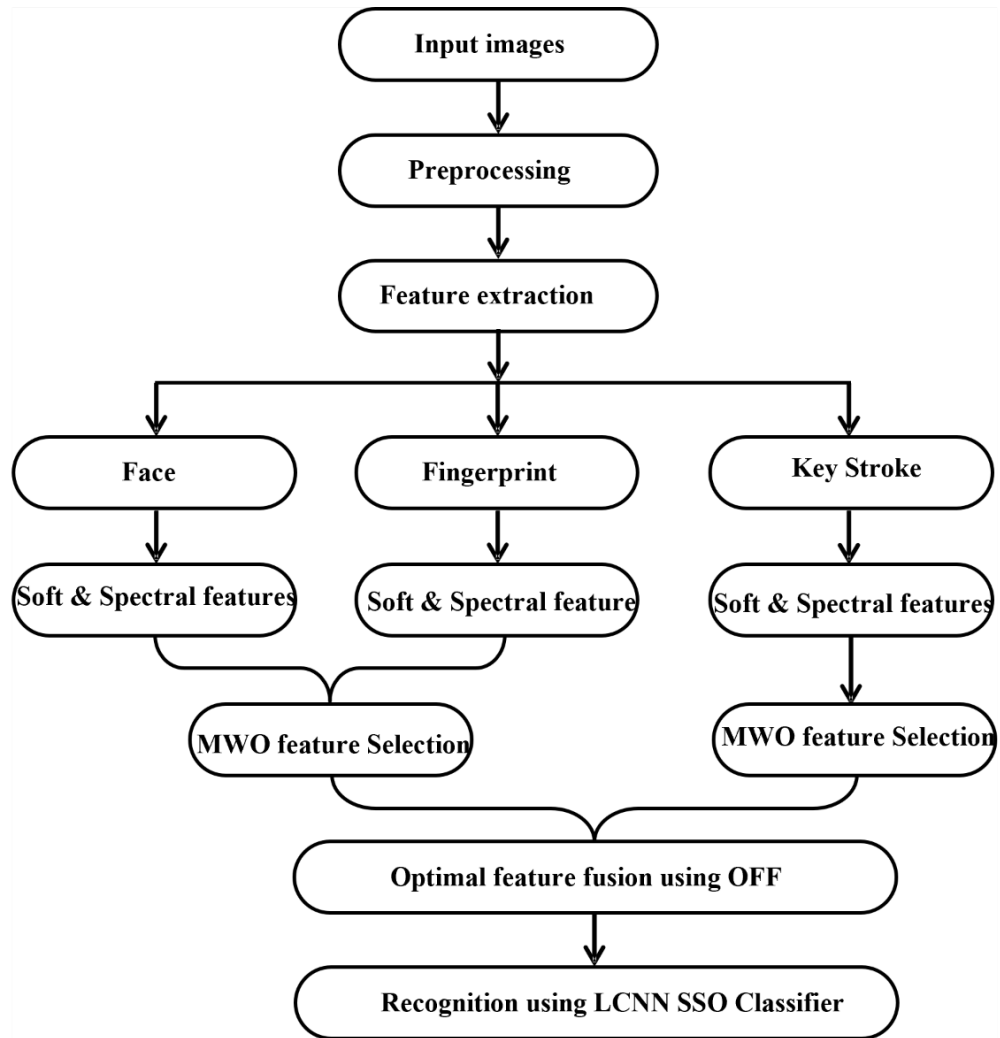


Fig. 4.2 Proposed model for fusion of face, fingerprint and keystroke for efficient continuous biometric user authentication system for online examination

4.3.3 Action Classification Using Lookup based Convolutional Neural Network-Salp Swarm Optimization (LCNN-SSO) based Classifier

Here hybrid LCNN-Salp swarm optimization (LCNN-SSO) based classifier with an optimum feature level fusion layer is presented. LCNN SSO classifier is tested with different dataset and then looking at the recognition rate obtained it was selected. The main motivation behind SSO is the salp's swarming nature while navigating and foraging in oceans. The algorithm has been tested on various numerical optimization functions to observe and confirm its effective responses in finding the optimal solutions. The outcome of the mathematical functions confirms that the SSO is capable of improving the random solutions and converges to the optimum. This is certainly not a significant issue when versatility to show different sorts of affiliations.

LCNN is an accurate model for convolutional neural network architecture despite being fast and compact. It enables efficient inference and training. Training of LCNN consists of co-learning a new set of vectors (dictionary) with some linear combinations. Here Convolutions are encoded using LCNN, a lookup-based convolutional neural network that is trained to cover the space of CNN weights. The size of the lexicon naturally reflects a range of efficiency and accuracy trade-offs. In ImageNet challenge LCNN offered 3.2 times speedup with 55.1% top-1 accuracy using AlexNet architecture [154].

4.3.3.1 Mathematical Dependencies

At duration point t , the terms mini pack $Y_t \in R^{n \times d}$ (n : number of models, d : number of information sources). When layer l ($l = 1, \dots, T$) is $G_t^{(l)} \in R^{n \times h}$ have secured condition (h : number of units hidden), the yield variable layer $Out_t \in R^{n \times q}$ A confirmed layer is affirming work h_l for level l . We could figure the checked condition of level l as already, utilizing Y_t information, for each following layer, the confirmed condition of the past level is utilized in its domicile.

$$G_t^1 = h_1(Y_t, G_{t-1}^1) \quad (4.20)$$

At long last, the yield of the yield layer is essentially picked the confirmed condition of camouflaged layer L. We utilize the yield work G to point out as,

$$Out_t = k_1(G_t^{(L)}) \quad (4.21)$$

Specifically, we can pick a standard LCNN, the long short-term Memory (LSTM), to fathom the model. An LCNN is a class of fake neural systems that builds up the standard feed-forward neural structure with coasts in affiliations.

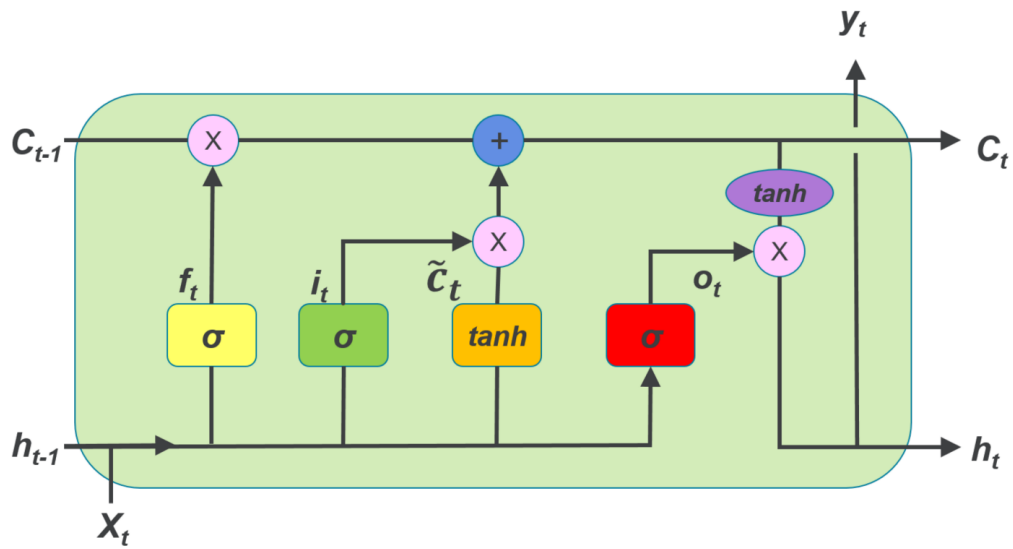


Fig. 4.3 Long short-term memory (LSTM) architecture

Right now, the structure can show dynamic passing leads. Given improvement information, where y_t denotes the information at j^{th} duration point, an LCNN fortifies it is sporadic confirmed g_t input state by

$$g_t = \begin{cases} 0, & \text{if } t = 0 \\ \phi(g_{t-1}, y_t), & \text{otherwise} \end{cases} \quad (4.22)$$

Where ϕ denotes limitation of nonlinear, sigmoid limit. Then again, the LCNN required a yield y . For specific tasks, for example, hyper-powerful picture course of action, this only needed one yield, i.e., y_t . In the standard LCNN type, which upload standard of the dull covered state is ordinarily executed as seeks after:

$$g_t = \varphi(Uy_t + Vg_{t-1}) \quad (4.23)$$

U and V are the terms in grids dedicating present movement and the foundation of unpredictable secured units at the past advancement. An LCNN can display a likelihood allotment all through the going with a fragment of the social occasion information, given its current input state h_t by getting a stream over game-plan information of length variable. Let 'y₁, y₂ and y_n' be the social occasion likelihood, which can be deteriorated into A broken layer with standard repetitive concealed units, which registers a weighted direct total of wellsprings of information and in this manner applies a nonlinear utmost. Abnormally, based on LSTM dreary layer makes a memory part b_t at duration t. The beginning units of LSTM can be figured by

$$g_t = out_t \tanh(b_t) \quad (4.24)$$

Where $\tanh(\bullet)$ denotes digression hyperbolic capacity and out_t is the yield door that decides the piece of the content memory that will be uncovered. The output gate is refreshed by

$$g_t = out_t \tanh(b_t) \quad (4.25)$$

Where $\sigma(\bullet)$ is a sigmoid key limit and term U mean weight, e.g., Uoi denotes information yield cross weight section and addresses Woc the storage-yield weight framework. The storage cell b_t is revived by including a new substance of storage cell b_t by discarding portions of the present memory content. The data entryway modifies how much the new memory information is added to the memory cell. How much substance of the current memory cell is ignored is picked by the neglected door g_t . The conditions that figure these two entryways are according to the accompanying:

$$j_t = \sigma(U_j j y_t + U_{jh} h_{t-1} + U_{jc} b_{t-1}) \quad (4.26)$$

$$p_t = \sigma(U_p j x_t + U_{ph} h_{t-1} + U_{pc} b_{t-1}) \quad (4.27)$$

The failure of the numerically simulated artificial neural network for signal classification leads to the enslavement of the initial centers of the slaves and significantly affects the final groups of nodes. When clustering data objects, the most accurate cluster group results can be achieved by selecting a data model for accurate data distribution, selecting cluster centers correctly, and updating the clustering center.

The idea of selecting cluster centers for the optimal LCNN classifier: (a) k rejects isolated points (k data objects and other objects that are rapidly added to the Euclidean distance of the cluster data); (b) Calculate all the centers of the remaining objects, (c) select the point closest to the center of the first point, and (d) repeat to find the point farthest from the starting point (the selected focus will no longer participate in the selection of the next center). The hierarchical system moves on to the next step until a point is found.

(a) Initialization Step:

In the introduction step, the k-implies initial calculation,

$$\mu_j^{(1)} = \{y_p: y_p \in X, \mu_i^{(1)} \neq y_p, 1 \leq j \leq k, j \neq i\} \quad (4.28)$$

Different types of SSO can use different methods. However, the most ordinary method is to use random data points to extract each centroid.

(b) Assignment Step:

Presently, every data point is named to a bundle whose centroid gives minimal squares in the gatherings.

$$c_i^{(t)} = \{y_n: \|y_n - \mu_i^{(t)}\|^2 \leq \|y_n - \mu_j^{(t)}\|^2, \forall j, 1 \leq i \leq k\} \quad (4.29)$$

Over time, y_n is only assigned to one cluster with every j one i . Overtime y_n can be assigned to an additional cluster, which reduces the cluster size of the squares.

(c) Update Step:

During the update phase, based on the cluster period centroids K - or c calculates the following iterations.

$$\mu_j^{(k+1)} = \frac{1}{|c_j^{(k)}|} \sum_{x_i \in c_j^{(k)}} x_i \quad (4.30)$$

For a given threshold $\xi \geq 0$, if $|\mu_j^{(k+1)} - \mu_j^{(k)}| \leq \xi, \forall j, 1 \leq j \leq t$, stops the iteration process. Otherwise, it returns to the work position and drops sets $k = k+1$.

Algorithm 1 Action classification using LCNN-SSO Classifier

Input: Received Biometric Data to be processed

Output: Classified output

Begin

Generate the initial solution randomly

Evaluate each individual in the population $f(x)$ based on error rate

Find the best solution from the population

While (stop when criteria satisfied)

for $i = 1$ to n do

for $j = 1$ to n do

if $(f(x_j) < f(x_i))$

Calculate signal by eq.

Calculate the distance between each biometric data i and j by eq.

Move all data (x_i) to the best solution (x_j) by eq

end if

end for j

end for i

Moves the best solution randomly

Find the best solution from the new population

end While

Return Best

End of the algorithm

4.4 Results and Discussion

In this section, we will evaluate the effectiveness of specific CBUA-OE devices with existing advanced equipment. Data availability for performance evaluation is a significant obstacle to developing user authentication systems [185]. Having relevant datasets is very important for the system design. If the data set characteristics are having good correlation with the classification problem, then getting good authentication performance is possible. The entire exercise is analyzed in terms of performance matrix parameters. Essential parameters are accuracy, specificity, sensitivity and equal error rate.

4.4.1 Dataset Descriptions

The proposed CBUA-OE technique is analyzed through multi-biometric data's is fingerprint, face, and keystroke. All the dataset used are cited at [162-172]. Following are the names (1) ORL Database (2) Yale Face Database (3) FASSEG Dataset (4) KEY STROKE -Kevin Killourhy and Roy Maxion (5) Finger Print FVC 2004 datasets 1 and 2



Fig. 4.4 Sample image from ORL database

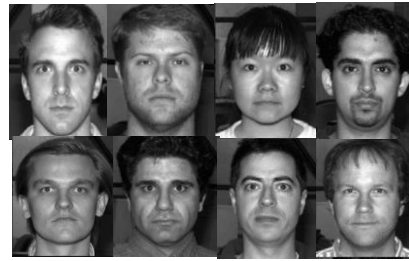


Fig. 4.5 Sample image from YALE database



Fig. 4.6 Sample image from FASSEG database

All these above-mentioned datasets have been used for the experiment purpose. After the multiple biometric data fusion, we compute the consolidated dataset, consisting of fingerprint, face, and keystroke of users. In our study we have considered the user's keystrokes like press-release, release-press, and hold duration. Fig. 4.7 shows the fused database of three different users (are U1, U2, and U3).

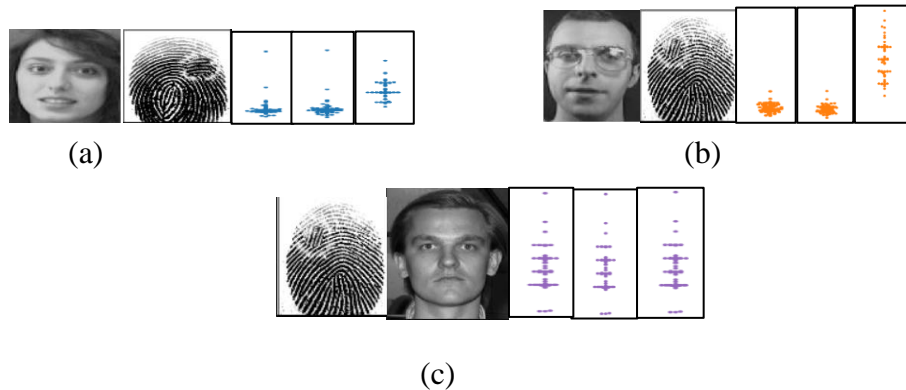


Fig. 4.7 Sample data with fused format of (a) User 1 face, fingerprint, and keystrokes (b) User 2 face, fingerprint, and keystrokes (c) User 3 face, fingerprint, and keystrokes

4.4.2 Performance Analysis of Proposed and Existing Techniques

In this section, we examine and contrast the presentation of the planned LCNN-SSO classifier with existing classifiers SVM, KNN, LSTM and DBM in terms of different presentation metrics: accuracy, precision, recall, etc. F-1 score, sensitivity, and specificity. Table 4.1 describes the performance of classifiers over different datasets are ORL, YALE, and FASSEG. The proposed LCNN-SSO classifier's performance is compared with the existing classifiers are SVM, KNN, LSTM, and DBN. Fig.4.8 clearly shows the accuracy comparison of planned and current classifiers for three databases: ORL, YALE, and FASSEG. It depicts that the average accuracy of the proposed LCNN-SSO classifier is 14% higher than the existing state-of-art classifiers. For the ORL database, the accuracy of the proposed LCNN-SSO classifier is 28%, 20%, 8%, and 0.2% higher than the SVM, KNN, LSTM, and DBN classifier, respectively. For the YALE database, the accuracy of the proposed LCNN-SSO classifier is 32%, 17%, 9.3%, and 1.3% higher than the SVM,

KNN, LSTM and DBN classifier, respectively. For the FASSEG database, the accuracy of the proposed LCNN-SSO classifier is 31%, 14%, 5.8%, and 1% higher than the SVM, KNN, LSTM, and DBN classifier, respectively.

Fig. 4.9 clearly shows the precision judgment of planned and existing classifiers for three databases: ORL, YALE, and FASSEG. It depicts the average precision of the proposed LCNN-SSO classifier is 15% higher than the current state-of-art classifiers. For the ORL database, the precision of the proposed LCNN-SSO classifier is 37.2%, 25.7%, 5.7%, and 1.5% higher than the SVM, KNN, LSTM, and DBM classifier, respectively. For the YALE database, the precision of the proposed LCNN-SSO classifier is 31%, 23.4%, 6.2%, and 0.2% higher than the SVM, KNN, LSTM and DBN classifier. For the FASSEG database, the precision of the proposed LCNN-SSO classifier is 15.6%, 19.5%, 7.1%, and 1.8% higher than the SVM, KNN, LSTM and DBN classifier.

Fig. 4.10 clearly shows the recall judgment of planned and existing classifiers for three databases: ORL, YALE, and FASSEG. It depicts the average recall of the proposed LCNN-SSO classifier is 11% higher than the current state-of-art classifiers. For the ORL database, the recall of the proposed LCNN-SSO classifier is 32%, 20%, 4.3%, and 0.3% higher than the SVM, KNN, LSTM and DBN classifier. For the YALE database, the recall of the proposed LCNN-SSO classifier is 28%, 18%, 3.2%, and 1.2% higher than the SVM, KNN, LSTM and DBN classifier. For the FASSEG database, the recall of the proposed LCNN-SSO classifier is 9.6%, 9.5%, 6.2%, and 0.8% higher than the SVM, KNN, LSTM, and DBN classifier, respectively.

Fig. 4.11 clearly shows the F-1 Score judgment of planned and existing classifiers for three databases: ORL, YALE, and FASSEG. It depicts the average F-measure of the proposed LCNN-SSO classifier is 18% higher than the current state-of-art classifiers. For the ORL database, the F-measure of the proposed LCNN-SSO classifier is 9.5%, 7.7%, 7.2%, and 1% higher than the SVM, KNN, LSTM, and DBN classifier, respectively. For the YALE database, the F-measure of the proposed LCNN-SSO classifier is 9.6%, 17.7%, 9.67%, and 7.3% higher than the SVM, KNN, LSTM, and DBM classifier, respectively. For the

FASSEG database, the F-measure of the proposed LCNN-SSO classifier is 6.57%, 5.2%, 3%, and 2.8% higher than the SVM, KNN, LSTM, and DBN classifier, respectively.

Fig. 4.12 clearly shows the sensitivity judgment of planned and existing classifiers for three databases: ORL, YALE, and FASSEG. It depicts the average sensitivity of the proposed LCNN-SSO classifier is 4.2% higher than the existing state-of-art classifiers. For the ORL database, the sensitivity of the proposed LCNN-SSO classifier is 6.07%, 4.9%, 4.5%, and 2.6% higher than the SVM, KNN, LSTM and DBM classifier. For the YALE database, the sensitivity of the proposed LCNN-SSO classifier is 5.3%, 4%, 1.56%, and 1.5% higher than the SVM, KNN, LSTM and DBN classifier. For the FASSEG database, the sensitivity of the proposed LCNN-SSO classifier is 8%, 6%, 5.5%, and 0.7% higher than the SVM, KNN, LSTM and DBN classifier.

Fig. 4.13 clearly shows the specificity assessment of planned and existing classifiers for three databases: ORL, YALE, and FASSEG. It depicts the average specificity of the proposed LCNN-SSO classifier is 3% higher than the current state-of-art classifiers. For the ORL database, the specificity of the proposed LCNN-SSO classifier is 3.4%, 3%, 3.2%, and 2.6% higher than the SVM, KNN, LSTM and DBN classifier. For the YALE database, the specificity of the proposed LCNN-SSO classifier is 2%, 1.9%, 4.4%, and 4.8% higher than the SVM, KNN, LSTM and DBN classifier. For the FASSEG database, the specificity of the proposed LCNN-SSO classifier is 5.5%, 4.3%, 4%, and 2% higher than the SVM, KNN, LSTM and DBNM classifier.

Fig. 4.14 shows the false positive rate (FPR) assessment of planned and existing classifiers for three databases ORL, YALE, and FASSEG. It highlights the average FPR of the proposed LCNN-SSO classifier is at least 35% lower than the existing state-of-art classifiers. For the ORL database, the FPR of the proposed LCNN-SSO classifier is 44%, 41%, 42%, 37% lower than SVM, KNN, LSTM, and DBN classifier, respectively.

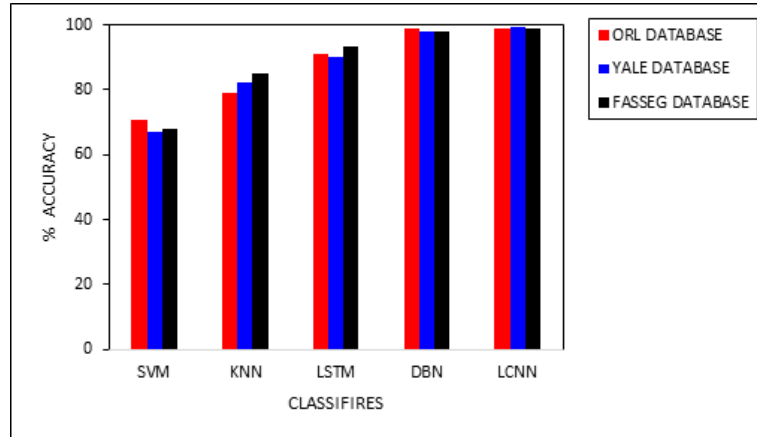


Fig. 4.8 Accuracy of proposed and existing classifiers

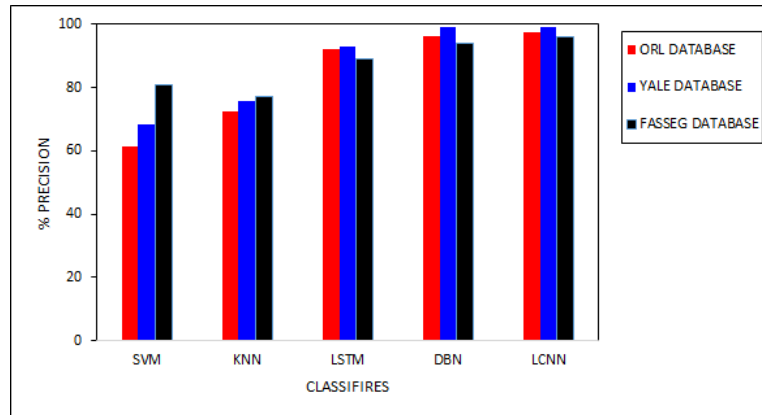


Fig. 4.9 Precision of proposed and existing classifiers

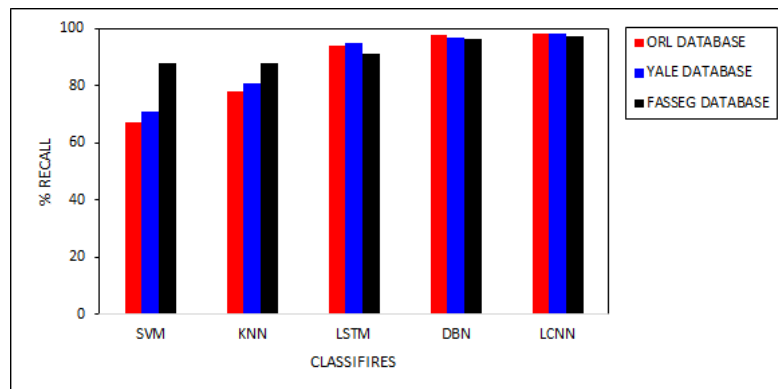


Fig. 4.10 Recall of planned and obtainable classifiers

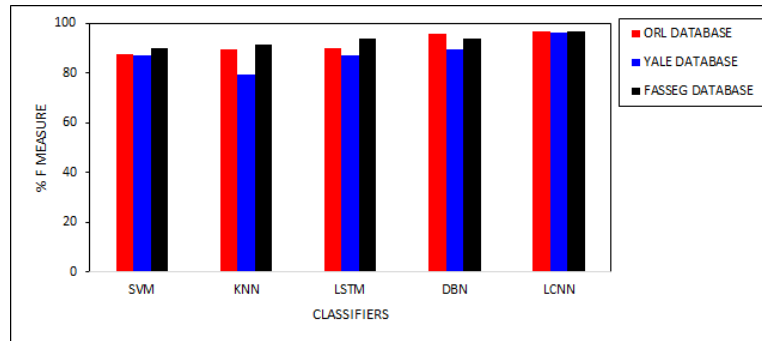


Fig. 4.11 F-measure of planned and obtainable classifiers

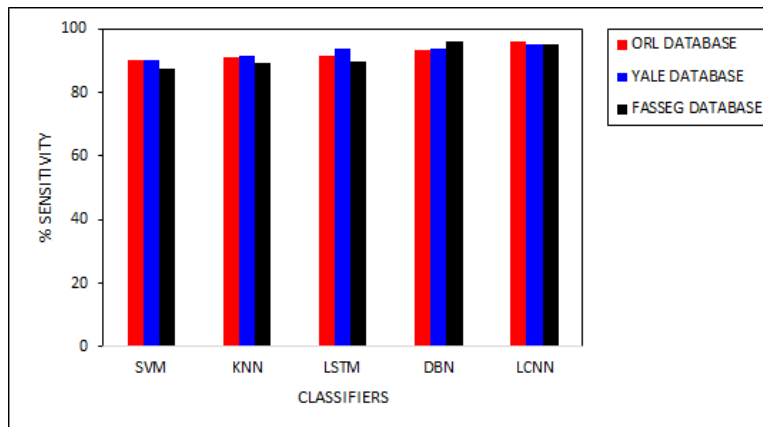


Fig. 4.12 Sensitivity of planned and obtainable classifiers

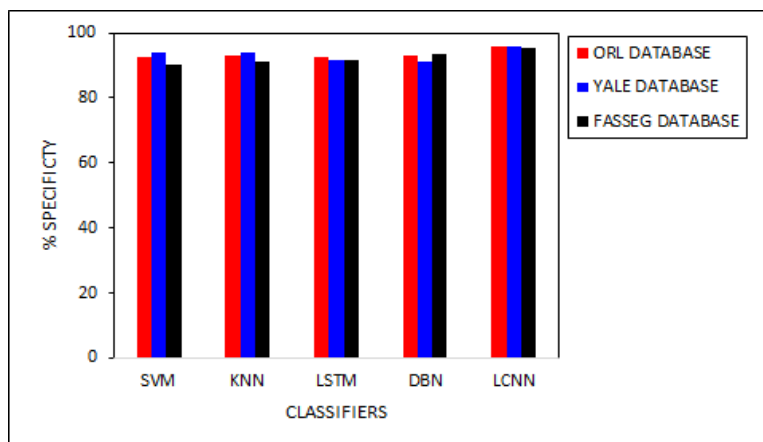


Fig. 4.13 Specificity of planned and presented classifiers

Table 4.1 Performance comparison of classifier for different databases

Metrics (%)	ORL database					YALE database					FASSEG database				
	SVM	KNN	LSTM	DBN	LCNN	SVM	KNN	LSTM	DBN	LCNN	SVM	KNN	LSTM	DBN	LCNN
Accuracy	70.98	79	90.87	98.92	99.12	67.20	82.30	90	97.92	99.24	68.18	84.98	93.22	98.12	99.13
Precision	61.2	72.39	91.90	96	97.5	68.27	75.88	92.93	98.90	99.12	80.83	77.09	89	94.09	95.87
Recall	67.09	78.01	93.98	97.96	98.25	70.92	80.80	94.98	97	98.17	87.93	88	91.25	96.5	97.29
F-measure	87.52	89.3	89.78	95.82	96.78	87.12	79.35	87.12	89.35	96.45	90.13	91.45	93.78	93.83	96.47
Sensitivity	90.13	91.25	91.58	93.45	95.96	90.13	91.45	93.78	93.83	95.27	87.52	89.3	89.78	95.82	95.08
Specificity	92.48	92.89	92.67	93.27	95.78	93.78	93.83	91.45	91.07	95.67	90.13	91.25	91.58	93.45	95.39
False Positive	7.52	7.11	7.33	6.73	4.22	6.22	6.17	8.55	8.93	4.33	9.87	8.75	8.42	6.55	4.61

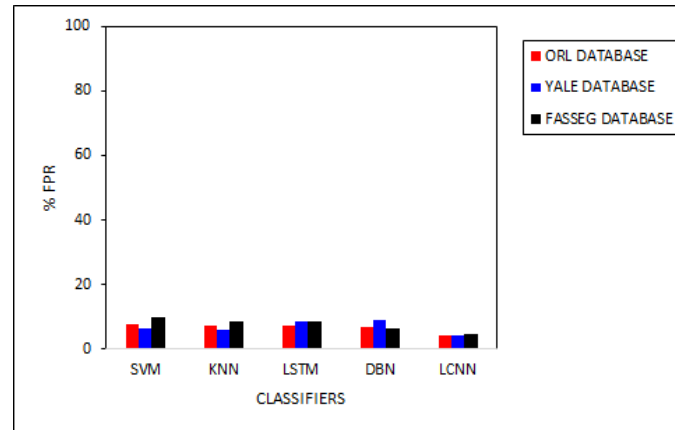


Fig. 4.14 False positive rate of planned and presented classifiers

For the YALE database, the FPR of the proposed LCNN-SSO classifier is 31%, 30%, 49%, and 52% lower than the SVM, KNN, LSTM and DBN classifier. For the FASSEG database, the specificity of the proposed LCNN-SSO classifier is 54%, 48%, 44%, and 30% higher than the SVM, KNN, LSTM and DBN classifier.

Table 4.2 describes the presentation contrast of planned CBUA-OE and obtainable techniques: FO-DBN, IILBDL and IKLDA+PNN in terms of accuracy, precision, recall, F-measure, sensitivity, and specificity. For the ORL database, the accuracy of the proposed CBUA-OE technique is 7%, 7%, and 14% higher than the existing FO-DBN, IILBDL, and IKLDA+PNN techniques, shows in Fig. 4.15. The precision of the proposed CBUA-OE technique is 17%, 1.6%, and 21.3% higher than the existing FO-DBN, IILBDL, and IKLDA+PNN techniques. The recall of the proposed CBUA-OE technique is 4.7%, 13.9%, and 6.6% higher than the existing FO-DBN, IILBDL, and IKLDA+PNN techniques. The F-1 of the proposed CBUA-OE technique is 5.4%, 4%, and 6% higher than the existing FO-DBN, IILBDL, and IKLDA+PNN techniques. The sensitivity of the proposed CBUA-OE technique is 1.7%, 0.4%, and 5.5% higher than the existing FO-DBN, IILBDL, and IKLDA+PNN techniques. The specificity of the proposed CBUA-OE technique is 2.8%, 1.4%, and 3% higher than the existing FO-DBN, IILBDL, and IKLDA+PNN techniques. All performance measures of proposed method are higher than existing methods for ORL database.

Metrics (%)	CBUA-OE	FO-DBN	IILBDL	IKLDA+PNN
Accuracy	99.12	84.92	91.31	91.48
Precision	97.5	80.37	79.02	78.05
Recall	98.25	84.26	83.50	80.25
F-measure	96.79	90.61	87.05	85.68
Sensitivity	95.96	90.60	90.22	89.02
Specificity	95.78	92.82	91.50	90.20
False Positive Rate	4.22	7.18	8.5	9.8

The FPR of the proposed CBUA-OE technique is 42%, 51%, and 57% lower than the existing FO-DBN, IILBDL, and IKLDA+PNN methods.

Table 4.3 describes the performance comparison of proposed CBUA-OE and existing techniques are FO-DBN, IILBDL, and IKLDA+PNN in terms of accuracy, precision, recall, F-measure, sensitivity, and specificity. The proposed CBUA-OE technique's accuracy is 8.7%, 14.6%, and 15% higher than the existing FO-DBN, IILBDL, and IKLDA+PNN for the YALE database show in Fig. 4.16.

Metrics (%)	CBUA-OE	FO-DBN	IILBDL	IKLDA+PNN
Accuracy	99.24	84.35	98.82	77
Precision	99.12	83.93	81.20	80.20
Recall	98.17	85.92	84.90	83.02
F-measure	96.45	85.73	83.15	80.50
Sensitivity	95.27	92.29	90.12	89.80
Specificity	95.67	92.53	91.25	90.12
False Positive Rate	4.33	7.47	8.75	9.88

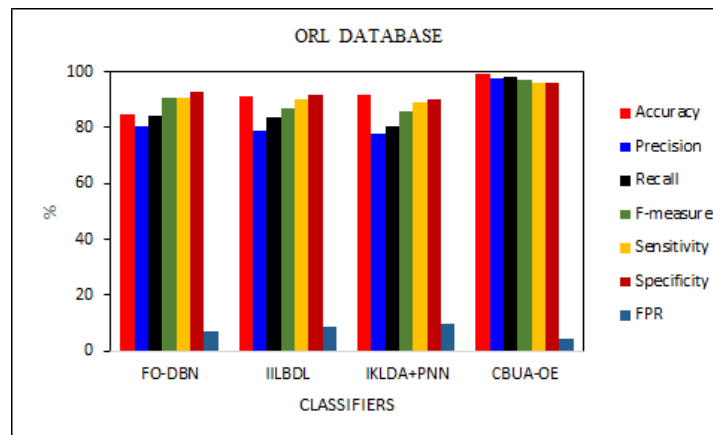


Fig. 4.15 Comparative analysis of planned and existing techniques for ORL database

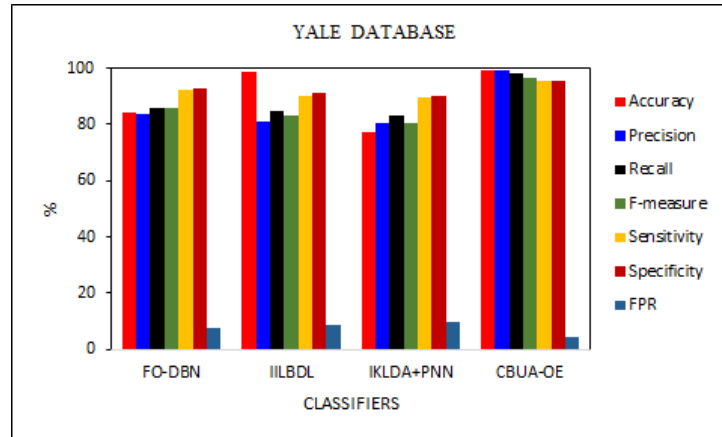


Fig. 4.16 Comparative analysis of planned and existing techniques for YALE database

The precision of the proposed CBUA-OE technique is 4.5%, 3.3%, and 15.2% higher than the existing FO-DBN, IILBDL, and IKLDA+PNN techniques. The recall of the proposed CBUA-OE technique is 15.4%, 1.1%, and 12.5% higher than the existing FO-DBN, IILBDL, and IKLDA+PNN techniques. The F-measure of the proposed CBUA-OE technique is 5.4%, 4%, and 6% higher than the existing FO-DBN, IILBDL, and IKLDA+PNN techniques. The sensitivity of the proposed CBUA-OE technique is 2.6%, 2.3%, and 3.1% higher than the existing FO-DBN, IILBDL, and IKLDA+PNN techniques. The specificity of the proposed CBUA-OE technique is 2.6%, 1.3%, and 3.2% higher than the existing FO-DBN, IILBDL, and IKLDA+PNN techniques. The FPR of the proposed CBUA-OE technique is 43%, 51%, and 57% lower than the existing FO-DBN, IILBDL, and IKLDA+PNN methods.

Table 4.4 describes the performance comparison of proposed CBUA-OE and existing techniques are FO-DBN, IILBDL, and IKLDA+PNN in terms of accuracy, precision, recall, F-measure, sensitivity, and specificity.

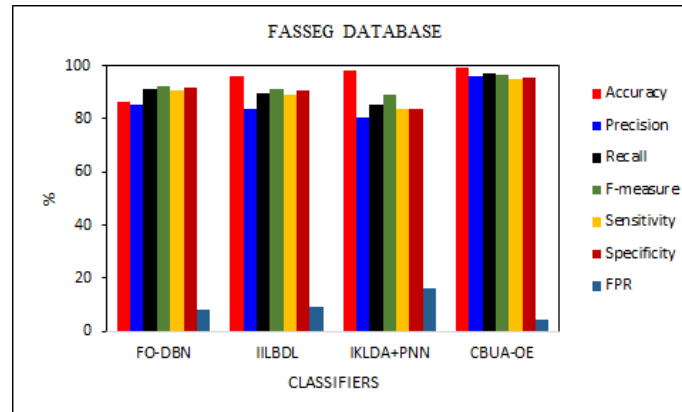


Fig. 4.17 Comparative analysis of planned and existing techniques for FASSEG database

For the FASSEG database, the accuracy of the proposed CBUA-OE technique is 12.2%, 10%, and 13% higher than the existing FO-DBN, IILBDL, and IKLDA+ PNN methods shows in Fig. 4.17. The precision of the proposed CBUA-OE technique is 5.4%, 2.1%, and 11% higher than the existing FO-DBN, IILBDL, and IKLDA+PNN techniques. The recall of the proposed CBUA-OE technique is 6.1%, 1.4%, and 6.5% higher than the existing FO-DBN, IILBDL, and IKLDA+PNN techniques. The F-measure of the proposed CBUA-OE technique is 3.3%, 1.1%, and 4.3% higher than the existing FO-DBN, IILBDL, and IKLDA+PNN techniques. The sensitivity of the proposed CBUA-OE technique is 7.8%, 1.7%, and 4.7% higher than the existing FO-DBN, IILBDL, and IKLDA+PNN techniques. The specificity of the proposed CBUA-OE technique is 8.7%, 1.1%, and 4% higher than the existing FO-DBN, IILBDL, and IKLDA+PNN techniques. The FPR of the proposed CBUA-OE technique is 46%, 52%, and 72% lower than the existing FO-DBN, IILBDL, and IKLDA+PNN methods.

Metrics (%)	CBUA-OE	FO-DBN	IILBDL	IKLDA+PNN
Accuracy	99.13	86.12	96.10	98.12
Precision	95.87	85.25	83.45	80.60
Recall	97.29	90.92	89.56	85.32
F-measure	96.47	92.29	91.25	89.24
Sensitivity	95.08	90.60	89.00	83.52
Specificity	95.39	91.60	90.52	83.54
False Positive Rate	4.61	8.4	9.48	16.46

4.5 Summary

The study proposed a continuous biometric user authentication system for online examinations for recognizing the normal and abnormalities. Here, MWO algorithm features are utilized to extract the features from the facial data. Then, the OFF algorithm selects the most favorable feature vectors and combines the facial images from different records. A set of most overriding discriminative features is obtained by choosing the optimal features from the extracted features. These best features are passed through the LCNN-SSO classifier. From simulation results, we observe that the average accuracy of the proposed CBUA-OE technique is 9.3%, 12.7%, and 11.7% higher than the existing techniques with ORL, YALE, and FASSEG databases. The average precision of the proposed CBUA-OE technique is 13.3%, 7.6%, and 8.4% higher than the existing techniques with ORL, YALE, and FASSEG databases. The average recall of the proposed CBUA-OE technique is 11.73%, 9.6%, and 4.6% higher than the existing techniques with ORL, YALE, and FASSEG databases. The average F-measure of the proposed CBUA-OE technique is 5.4%, 6.5%, and 2.9% higher than the existing techniques with ORL, YALE, and FASSEG databases, respectively. The average sensitivity of the proposed CBUA-OE technique is 2.5%, 2.7%, and 4.7% higher than the existing techniques with ORL, YALE, and FASSEG databases, respectively. The average specificity of the proposed CBUA-OE technique is 2.4%, 2.42%, and 4.6% higher than the existing techniques with ORL, YALE, and FASSEG databases. The average FPR is 35%, 39%, 42%, and 48 % lower than the existing FO-DBN, IILBDL, and IKLDA+PNN techniques, establishing the superior performance of the proposed method on all performance measures.

Chapter 5

Modified VGG 16 based Multimodal Biometric

Authentication

Biometric recognition technology has become a common part of daily life due to the global demand for information compliance and safety regulations. This work proposes a new multimodal biometric person identification system based on a deep learning algorithm for identifying people using their face, iris, and ear biometrics. The proposed multimodal biometric systems system's architecture is built on convolutional neural networks (CNNs), which extract features and use a SoftMax classifier to categorize images. Three CNN models—one for the face, one for the ear, and one for the iris—were integrated to create the system. The CNN model was created using the Modified Visual Geometry Group-16(MVGG16) model, which was implemented based on the changing of input's size and number of kernels and reducing of convolutional layers from the conventional pre-trained model of VGG16. Categorical cross-entropy was utilized as the loss function and the Adam optimizer was employed. Image augmentation and dropout techniques were used to prevent overfitting. Feature-level fusion techniques were used to fuse the CNN models. The SDUMLA-HMT dataset, a multimodal biometrics dataset, and IIT Delhi ear database, were used in a number of experiments to empirically assess the performance of the proposed system. The outcomes showed that employing three biometric traits in biometric authentication systems produced better outcomes than using two or even one biometric traits. The outcomes also demonstrated that proposed feature-level fusion strategy, which achieved an accuracy of 100 percent, comfortably surpassed other state-of-the-art methodologies.

5.1 Research Methodology

This section gives the overview of proposed system structure, proposed system modules and proposed methods. Proposed system structure includes unimodal and multimodal system, preprocessing, feature level fusion and classification. Proposed methodologies include image augmentation, deep learning (DL), convolutional neural networks (CNN), transfer learning, Visual Geometry Group-16 (VGG16) and Modified Visual Geometry Group-16 (MVGG16) architecture.

5.1.1 Overview of Proposed Method

Unimodal System

The suggested unimodal biometric structure relies on face, ear, and iris biometrics is depicted in Fig. 5.1. In preprocessing, contrast enhancement using adaptive histogram equalization (AHE), image scaling, and image augmentation, are used to gather more informative data and expand the training set. The major section of the biometric system is the features extraction stage, which enables the extraction of the relevant data from the face, ear, and iris modalities. After preprocessing in both the training and testing stages, deep MVGG16 architectures are used to extract features. SoftMax classifier classified the training and testing features. The user identity/class can be determined using these results.

Multimodal System

This study suggests a face, ear, and iris-based multimodal biometric system based on MVGG16-CNN. Fig. 5.2 illustrates the overall layout of the proposed approach. The user's face, ear, and iris images are first taken. The multimodal system, which consists of three fused MVGG16 CNNs for face, ear, and iris recognition at the feature level, is then utilized to determine the user's identity. The model then generates the user identity.

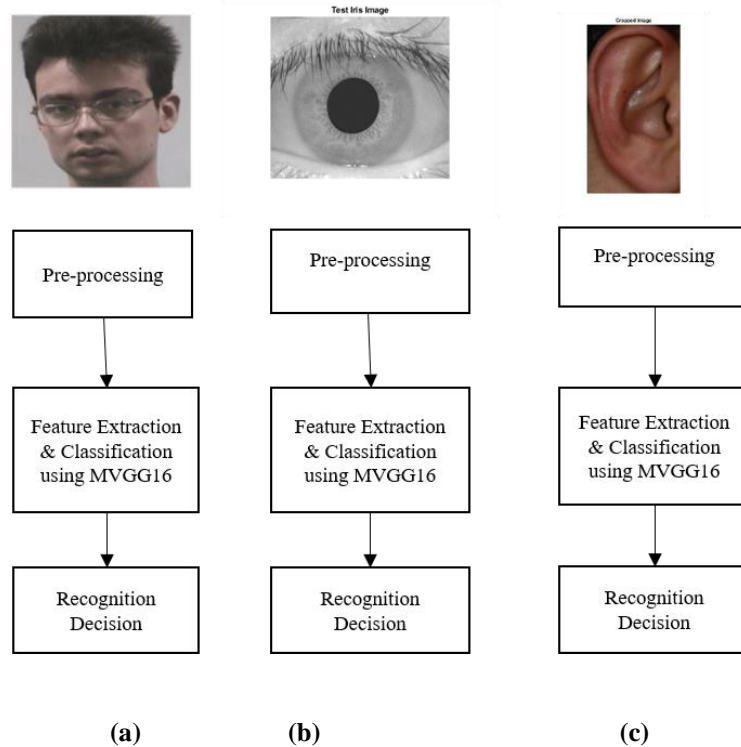


Fig. 5.1 Unimodal recognition (a) face (b) ear (c) iris

Fig. 5.3 illustrates the suggested multimodal MVGG16- CNN model's framework and the feature-level fusion method used in this study. The proposed system is as follows:

- Images of the face, ears, and irises are first taken from the SDUMLA-HMT multimodal biometric dataset and IIT Delhi dataset.
- The images are then subjected to preprocessing steps like image scaling, contrast enhancement via adaptive histogram equalization to extract more information, and data augmentation to extend the dataset.
- Each biometric attribute is put into its MVGG16 model. By changing the input size, the number of kernels, and the convolutional layers of the original VGG16 model, the MVGG16 model was developed.
- Fusion of features is conducted at feature level fusion, so, the features are integrated prior to the SoftMax classifier. The user identity is the final output of the fused model. The following subsections provide descriptions of the model's various components.

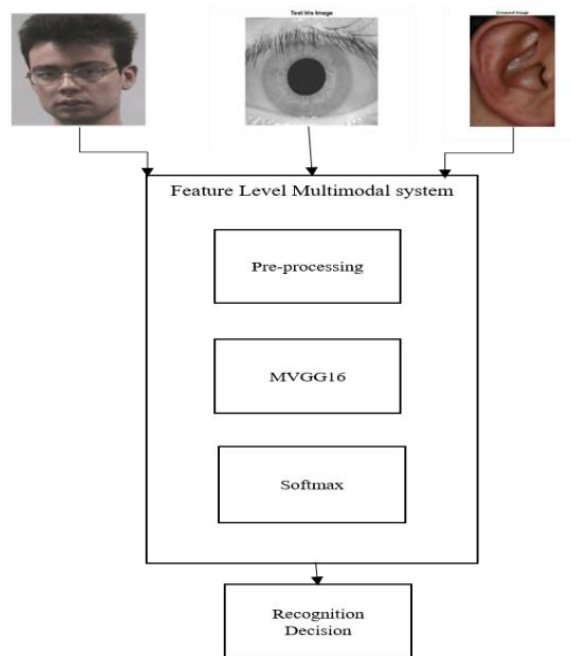


Fig. 5.2 Proposed multimodal architecture

5.1.2 Modules used in CNN based Multimodal Architecture

A. Preprocessing

Image scaling, contrast enhancement and data augmentation are the three preprocessing methods used. Generally, dataset has different size of images but deep learning models are developed for standard image size so dataset images are resized from different size into same size. Multi- biometrics face, iris and ear images are reduced to 128×128 pixels.

Sometimes, images have uneven brightness and poor contrast. Due to this, the contrast enhancement technique was developed. It uses an Adaptive Histogram Equalization (AHE) method to improve the visual quality of an image without losing image data. This method of enhancement function, which derives from transformation function, is applicable to all surrounding pixels. Finally, data augmentation was used to increase the size of training data and thereby decrease over fitting issues.

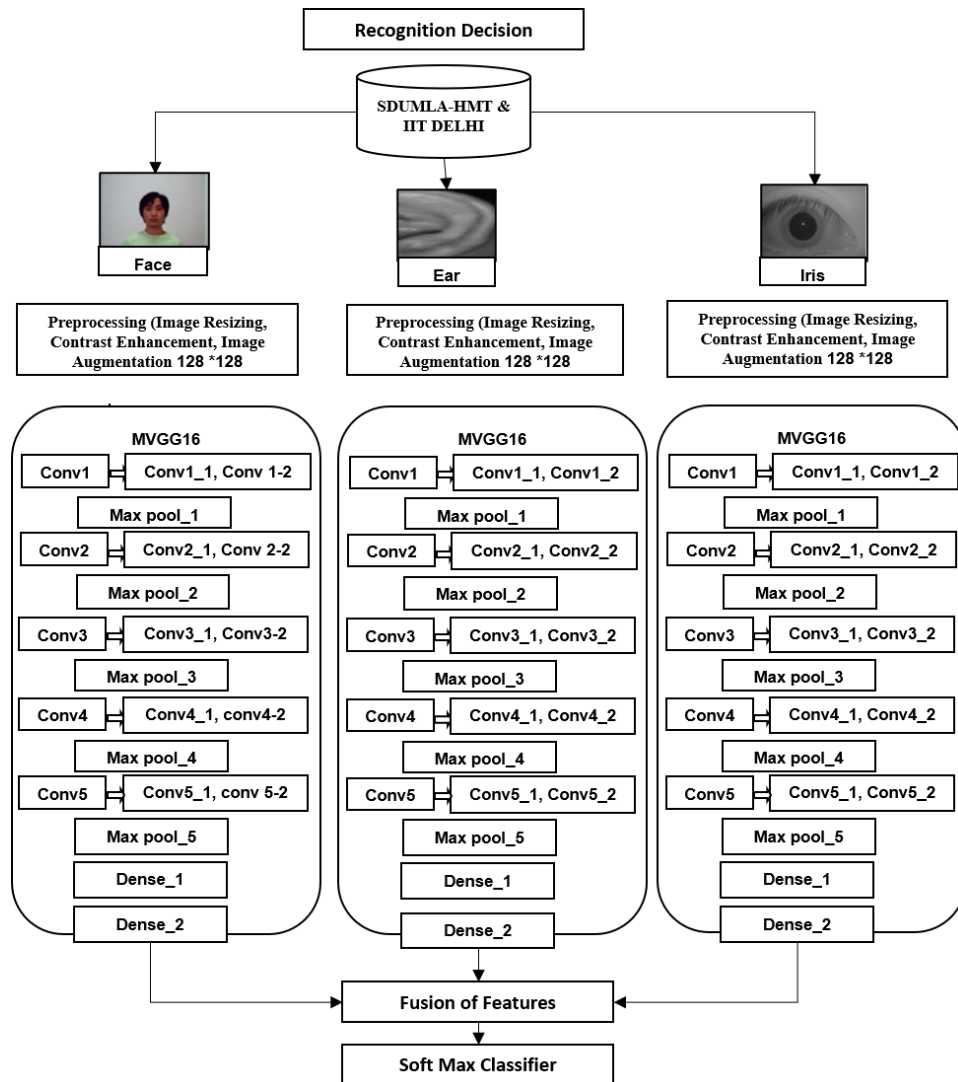


Fig. 5.3 Multimodal MVGG16 CNN model's framework (with feature level fusion)

Rotation, shearing, zooming, breadth and height shifting have been employed to increase the number of iris and ear images. Additionally, rotation, shearing, zooming, width and height shifting, and horizontal flipping were employed to enlarge the faces in the images.

B. Feature Level Fusion

Fusion at the feature level involves combining features that are representative of several attributes. In the training phase of this fusion strategy, the model picked up on the merged features. The results of the face, ear, and iris CNN models are fused at the second fully connected layer. As a result, the features vectors from the second fully connected layer of the three MVGG16 CNN models combine to form a single vector, which is represented by the following definition:

$$F = f_f | e_f | i_f \quad (5.1)$$

where f_f denotes features that were extracted from the face image, e_f denotes features that were retrieved from the ear image, and i_f denotes features that were extracted from the iris images. The Softmax classifier subsequently labels the image based on the similarity score and determines the individuals identify after receiving the resultant vector (F).

C. Classification

One of the crucial processes in deep learning architectures is the classification of the retrieved features. Softmax approach is now a common classifier in Alex Net, Google Net, VGG16 model, and other Deep Learning systems. Softmax activation function is employed in the last layer of the network. When employed in the DL models, it is used to analyze the output of the FC layer for N classes and is called as multinomial logistic regression. In this study, SoftMax is used to classify the face, ear, and iris features in the presented unimodal and multi - modal biometric recognition systems.

Softmax: It starts with a vector of n real numbers, where n is the number of classes, and normalizes the input into a vector of values that follow a probability distribution, the sum of which is 1. The neural network (NN) model can accept as many classes as the output values, which range from 0 to 1. The target class is determined by calculating the probabilities of each class among all feasible classes using the SoftMax classifier. Each element in the vector is subjected to the exponential function by the SoftMax classifier,

which then normalizes these values by dividing by the sum of all the exponentials according to the following formula:

$$S_x = \frac{\exp(f_x)}{\sum_{y=1}^M \exp(f_y)} f_x = \sum_z a_z b_{zx} \quad (5.2)$$

where $0 \leq S_x \leq 1$ and $\sum_{x=1}^M S_x = 1$

A SoftMax layer function's entire input is provided by f_x and it represents the values from the nodes of the output layer. The parameters a and b , respectively, reflect the activations and weight of an FC layer. M is the number of classes.

5.1.3 Proposed Methods

(a). Data Augmentation

Large training datasets are required for deep learning because more images can lead to better training accuracy. In comparison to a powerful algorithm with a little amount of data, even a weak method with a large volume of data can become more accurate [193]. Class imbalance is another concern, if there are significantly more samples of one class than the other during binary classification training, the final model will be biased. When there are an equal or balanced number of samples in each class, deep learning algorithms operate at their best.



Fig. 5.4 Image Augmentation. The transformations are: crop & pad, elastic distortion, scale, piecewise affine, translate, horizontal flip, vertical flip, rotate, perspective transformation, and shear, in that order from left to right and from top to bottom [158]

Image augmentation is one method of expanding the training dataset without adding new images. The process of image augmentation modifies the original images. This is

accomplished by applying various processing techniques, such as rotations, flips, translations, zooms, addition of noise and etc. [194]. Various samples of images following image augmentation are shown in Fig. 5.4.

The amount of pertinent data in the dataset can also be increased with the aid of data augmentation. We may provide pertinent features and patterns through augmentation, effectively improving overall performance. Additionally, data augmentation aids in avoiding overfitting. Overfitting is the process of a network learning a multi-variate function, such as the perfect modelling of training data. Overfitting occurs when a network remembers the noises and patterns from training data sets and begins to closely imitate them. When a model attempts to fit more data than is necessary and catch every datapoint that is provided to it, this is known as overfitting. As a result, it begins to extract erroneous and noise-filled data from the dataset, which lowers the model's performance. An overfitted model can't generalize well and doesn't perform well with the test or unknown dataset. It is referred to as having low bias and large variance in an overfitted model.

By providing the model with additional varied data, data augmentation tackles the overfitting issue [195]. Due to the diversity of the data, the model's generalization is enhanced, and variance is decreased.

Data augmentation, however, cannot eliminate all biases found in a limited dataset [195]. Higher training time, transformation compute costs, and additional memory costs are some further drawbacks of data augmentation.

(b) Deep Learning (DL)

Since its initial introduction in 2006, the deep learning concept has generated a significant amount of research and commercial attention. The ILSVRC competition has drawn many of the top artificial intelligence (AI) companies in the world since it was founded.

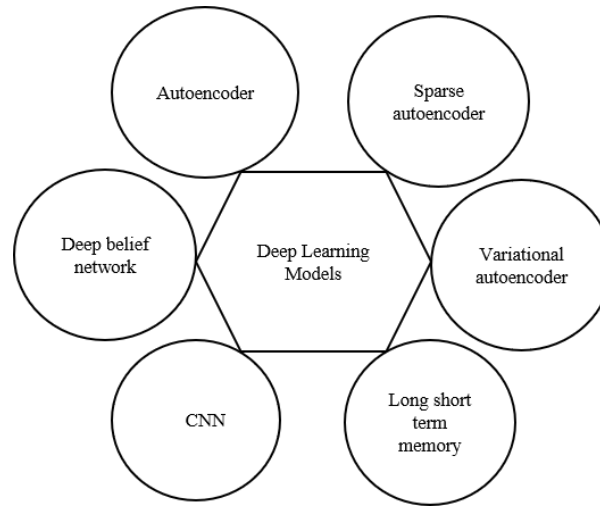


Fig. 5.5 Various deep learning methodologies

Functions of DL - While conventional machine learning techniques manually extract fundamental features from a dataset and feed those features into a particular ML algorithm, deep learning approaches automatically derive high-level features from the original dataset. Low, middle, and high-level characteristics will be retrieved in order to detect or categorize the dataset. As a result, using non-linear functions on the original data as inputs to produce abstracted outcomes is an essential component of deep learning systems. The readily accessible deep learning methods need large datasets and sophisticated processing tools. Deep learning techniques have used powerful GPUs& TPUs to shorten learning times and boost classification performance in recent years.

In computer vision, natural language processing, and automatic audio classification, deep learning architectures (shown in Fig. 5.5) like CNN, Recurrent Neural Networks (RNN), Deep Belief Network (DBN), Deep Stacking Networks (DSN), and Long Short-Term Memory (LSTM) are used to address problems with big datasets. Fig. 5.6 illustrate the uses of deep learning algorithms. Similar to ANN, CNN draws its inspiration from the network of connected biological neurons. Both algorithms have neurons that have biases and weights [197,198]. Although both CNN and ANN have layers, there are significant distinctions between the two networks' topologies. In a standard ANN algorithm, the layer

structure is one-dimensional and all layers are fully-connected. CNN, on the other hand, has a layer of two or three-dimensional neurons that encompass width, height, and depth.

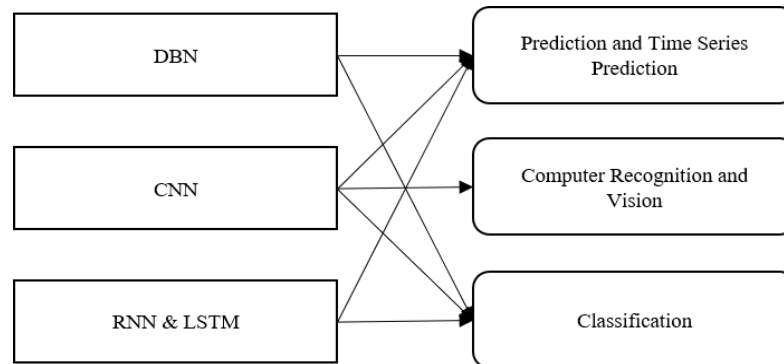


Fig. 5.6 Key applications of deep learning

(c). Convolutional Neural Networks (CNN)

The most basic deep learning algorithm, CNN, is very good at detecting patterns in images. In a nutshell, CNN takes an image's features and reduces its dimensions without losing any of the original image's qualities. Fig. 5.7 shown the general architecture of CNN. The main layers that make up a CNN are convolutional layer, pooling layer, Relu and fully-connected layer. The two steps that CNNs take during learning are typically feature extraction and classification. During the feature extraction stage, convolution is applied to the input data using a filter or kernel. Afterward, a feature map is made. During the classification stage, the CNN calculates a probability that the image belongs to a particular class or label. Since CNN automatically learns features rather than requiring manual feature extraction, it is especially helpful for classifying and recognizing images. Additionally, CNN can be retrained and deployed in a different domain via transfer learning. Transfer learning enhances classification performance, as has been demonstrated [196]. Spatial and Temporal correlation in input data can be explored using CNN. This capacity makes it suitable for image segmentation, classification, object detection and localization tasks.

Input Image- Input to CNN are images of different dimensions and it can be grey or color with RGB color plane details. The role of convolution process is to reduce the dimension of image which is convenient to process.

Convolutional layer-The convolutional layer, is the brain of a convolutional neural network. Different convolutional kernels filters (3x3x1, 2x2x1, 5x5x1 etc.) can be used to obtain different attributes. As shown in Fig. 5.8, set of weights (kernel values) are multiplied by the input in a linear process called convolution. A dot product is created by multiplying the filter by a portion of the input that is the size of the filter. To obtain a single value, the dot product is then combined together. Same step is repeated by shifting the kernel all over the image by a specific value like 1 or 2. This shifting parameter is known as stride. The resultant matrix would be of different size than original input image. To maintain the original size zero padding can be used. The output dimension can be calculated by Eq. 5.4 to 5.6.

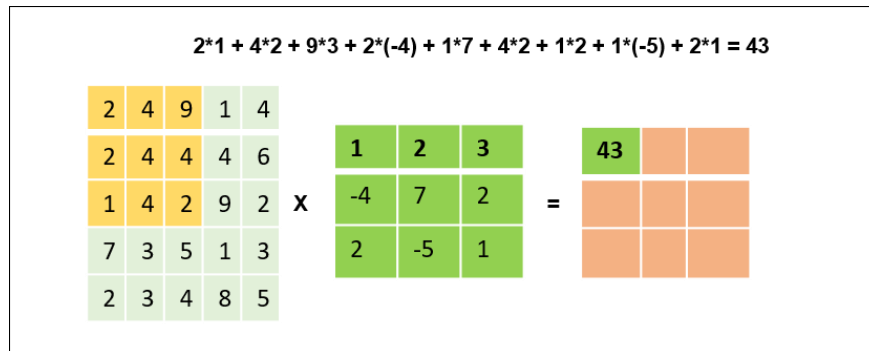


Fig. 5.7 Convolution operation example with kernel size 3x3

In case of RGB images, depth of the Kernel filter is same as of input image that is 3. The objective of the convolution layer is to extract features from the input image. Initial layers extract low level features and with increasing depth of added layers, the architecture adapts high level features.

Table 5.1 Hyper parameters of CNN

Hyperparameter	Description
Kernel/Filter Size	Kernel size at each convolution layer
Kernel/Filter count	Number of kernels at convolution layer
Stride	Amount by which we slide filter in horizontal and vertical direction when performing a convolution operation
Padding	Parameters used to preserve size and not to lose information of the training data
Epoch	Number of learning iterations
Learning rate	Amount of change in weight that is updated during training
Layer depth	Number of layers constituting an entire network
Batch size	Group size to divide the training data into several groups
Neuron count	Number of nodes in a fully connected layer
Loss function	Function to calculate error/loss
Activation function	Activation function at each node (Relu, sigmoid, softmax)

The convolution process can be expressed by the equation

$$map_{i,j}^{x,y} = f\left(\sum_m \sum_{h=0}^{H_i-1} \sum_{w=0}^{W_j-1} k_{i,j}^{h,w} map_{i-1,m}^{(x+h),(y+w)} + b_{i,j}\right) \quad (5.3)$$

here $k_{i,j}^{h,w}$ is the value at position (h, w) of kernel connected to the m^{th} feature in $(i - 1)^{th}$ location. H_i and W_j are the height and width of the kernel, $b_{i,j}$, is the bias of j^{th} feature in $(i - 1)^{th}$ layer.

To protect the corner information, zeros are padded in out periphery of the input image. It can also produce output, which is of input size. This process is known as Padding, which keeps the output image the same size after the convolution operation. To retain the dimensions, zero padding is applied to more layers so that more and better features are extracted.

Pooling layer – The pooling layer is typically employed between convolutional layers to reduce the complexity and network calculation; as a result, the input size is decreased in all depth sections through the subsampling operation, preventing overfitting during network training. The pooling method reduces the input's spatial size, which leaves the depth dimension unchanged. The two most popular kinds of pooling operations are max

pooling and average pooling shown in the Fig. 5.9. Equations (5.4), (5.5) and (5.6) allow for the breadth and height of the output in the pooling layer. W_n , H_n , P and D_n in both equations stand for the input's width, height, padding value and depth, respectively. K stands for the kernel size, L stands for the stride size (the amount by which kernels are shifted on the input image).

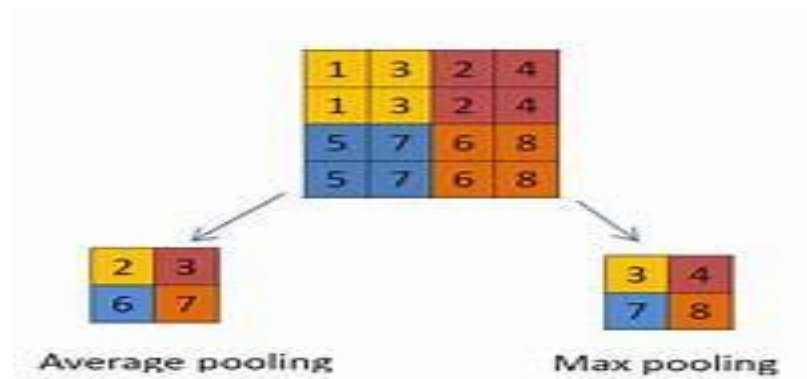


Fig. 5.8 An example for max pooling and average pooling

$$W_n = \frac{W_o - K + 2P}{L} + 1 \quad (5.4)$$

$$H_n = \frac{H_o - K + 2P}{L} + 1 \quad (5.5)$$

$$D_n = D_o \quad (5.6)$$

Fully Connected Layer - Fully connected layers (FC Layers) make up the final layers of the CNN architecture, are positioned before the output layer. All the output values of previous layers are flattened and then connected with FC layer. The flattened vector undergoes some fully connected layers depending upon the architecture.

If all the neuron output is connected, network may get into overfitting. Where it starts imitating training dataset very closely. To overcome this issue, some neurons inputs are dropout from the training process. Which results in less network parameters and avoids overfitting. Alternate method to control overfitting is to use global max pooling or global

average pooling layers instead of fully connected layer. Global pooling creates one extracted feature to reach relevant category of the classification model from the last convolution layer. In global averaging pooling layer, the average of each feature map from the convolution layer is computed and fed to the softmax layer, rather than building a fully connected layer on the top of feature maps.

Activation Function- The activation function is one of the most crucial elements of the CNN model. They are employed to discover and approximation any type of continuous and complex link between network variables, and the network gains nonlinearity as a result. It determines which model information should shoot ahead and which should not at the network's end.

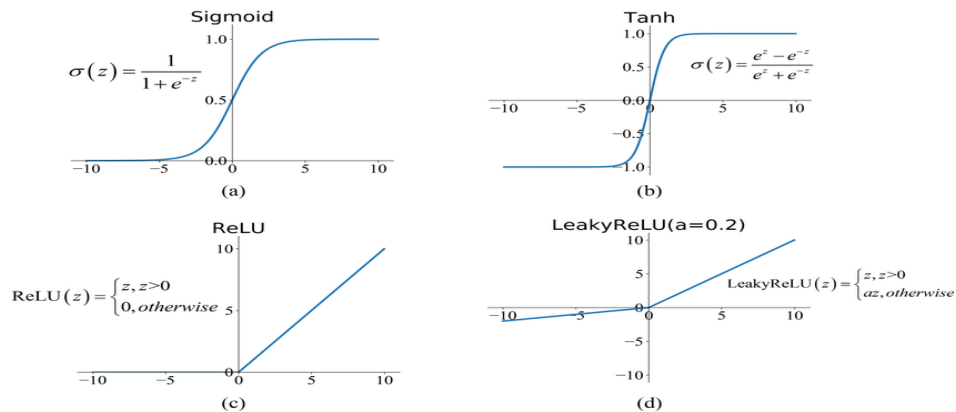


Fig. 5.9 Different activation functions used in CNN

The ReLU, Softmax, tanh, and Sigmoid functions are a few examples of regularly used activation functions. Each of these operations has a particular use. Sigmoid and softmax functions are preferred for a CNN model for binary classification, and softmax is typically employed for multi-class classification. Due to its short calculation size and quick training time, the SoftMax classifier is commonly utilized in the output layer to solve multi-classification issues [196]. To put it simply, activation functions in a CNN model decide whether or not to activate a neuron. It determines via mathematical processes whether the input to the work is significant or not for prediction.

The output of the previous layer's activation is applied to the output of a rectified linear unit (ReLU), which enhances the CNN by applying an element-wise activation function, such as sigmoid. ReLU function shown in Eq. (5.7). Value v is the net value.

$$\text{Relu}(v) = \max(0, v), \quad (5.7)$$

Optimization

There are different types of optimization methods for the CNN learning to converge to the desired output. Such as (1) Stochastic gradient descent (SGD) (2) Adaptive gradient (3) RMS prop (4) Adaptive moment estimation (ADAM). In this work we have used SGD and ADAM optimizer.

Stochastic gradient descent

The SGD algorithm is an improved version of the Gradient Descent (GD) algorithm that addresses some of the challenges of GD algorithm. GD consumed alot memory to load the full dataset of n -points at once to compute derivative. The SGD algorithm computes the derivative one point at a time. The process of updating a parameter(θ) for each training example is given by

$$\theta = \theta - \alpha \frac{\partial}{\partial \theta} f(\theta; x_i, y_i) \quad (5.8)$$

where α is the learning rate, f is the cost function. The size of the steps we take to reach a (local) minimum is determined by α .

Adaptive moment estimation

Adaptive moment estimation (ADAM) is for adaptive learning rates. It stores the decaying average of gradients m_t . The decaying gradients and squared gradients are calculated by the equations

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \quad (5.9)$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \quad (5.10)$$

m_t & v_t are the estimate of means and uncentered variance of gradient. The parameter update by Adam optimizer is given by

$$\theta_{t+1} = \theta_t - \bar{m}_t \left(\frac{\alpha}{\sqrt{\bar{v}_t^2 + \epsilon}} \right) \quad (5.11)$$

(d) Transfer Learning

A pre-trained model is used as the basis for a new model in the machine learning technique known as transfer learning. Simply expressed, an optimization that enables quick progress when modeling the second task is applied to a model that was trained on the first, related task. One can attain considerably better performance than training with only a modest quantity of data by applying transfer learning to a new task. It is uncommon to train a model from scratch for tasks linked to image or natural language processing because transfer learning is so widespread. Transfer learning structure shown in Fig 5.10.

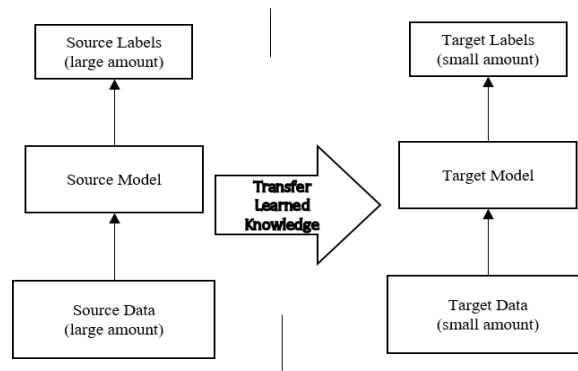


Fig. 5.10 Transfer learning mechanism

In proposed work, modified architectures of VGG 19-CNN have been used. Basic operational and structural details of VGG and MVGG16 are given below.

5.2 Visual Geometry Group -VGG

The Visual Geometry Group at the University of Oxford and Google DeepMind together created VGGNet CNN [7]. The network design of the VGGNet, which is depicted in Fig. 5.12 is characterized by 3×3 convolutional kernels and 2×2 pooling layers. The concept of a much deeper network with much smaller filters has been utilized as a VGG network. VGGNet-16 and VGGNet-19 are the two most popular VGGNet versions [198]. In ImageNet, the VGG16 model achieves top-5 test accuracy of about 92.7%. A dataset called ImageNet has over 14 million images that fall into almost 1000 classes. It was also among the very well models submitted at ILSVRC-2014. It significantly outperforms

AlexNet by substituting a number of 3×3 kernel-sized filters for the huge kernel-sized filters. The VGG19 model has the same idea as the VGG16 model, with the exception that it supports 19 layers. The numbers "16" and "19" refer to the model's weight layers (convolutional layers). In comparison to VGG16, VGG19 contains three extra convolutional layers.

5.2.1 Architecture of VGG

Input: The VGGNet input standard dimension is 224×224 . Sometime central window of 224×224 is cropped from the input image for processing.

Convolutional Layers: Convolutional layers consist of fix size, kernel of 3×3 with varying depth to cater the need of required features. The next component is a Rectified linear unit activation function (ReLU), is a piecewise linear function that, if the input is positive, outputs the input; otherwise, the output is zero. To maintain the good spatial resolution after convolution, the convolution stride is kept at 1 pixel.

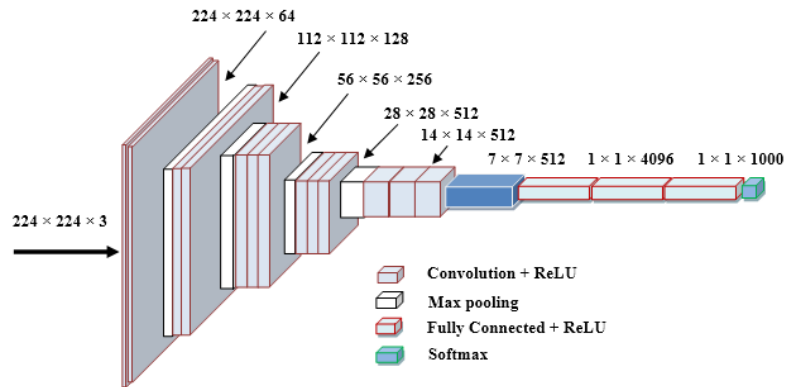


Fig. 5.11 VGGNet architecture [125]

Pooling & Relu Layers: In VGG architecture max pooling is used to reduce the data size. Normally maxpooling of 2×2 with stride of 2 is used after convolution layers. The VGG network's layers uses ReLU to incorporate nonlinearity factor in the process.

Fully Connected Layers: The VGGNet contains three layers with full connectivity. The first two layers each have 4096 channels, while the output layer has 1000 nodes with one node for each class.

(i) VGG19

In this version there are 19 weight layers, including the same 5 pooling layers and 16 convolutional layers with 3 fully connected layers each. Architecture of VGG19 shown in Fig. 5.13. There are two fully connected layers with 4096 channels each in both variations of the VGGNet, followed by a further fully connected layer with 1000 nodes with to predict 1000 classes. Softmax layer is used as the final fully connected layer for categorization.

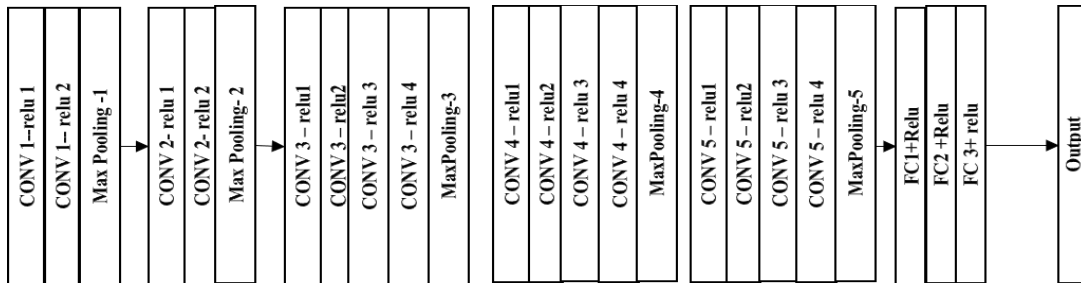


Fig. 5.12 Architecture of VGG19

(ii) VGG16

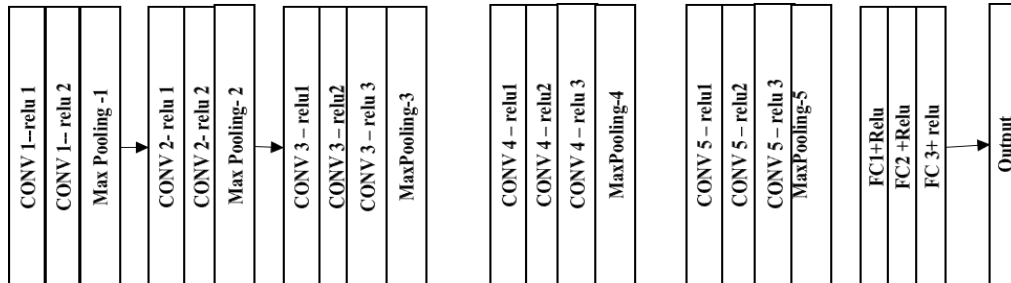


Fig. 5.13 Architecture of VGG16

Table 5.2 VGG16 network model

VGG16 Layers	Number of Kernels	Kernel Size	Output Size	Trainable Parameters
Conv1_1	64	3 × 3	64 × 224 × 224	1792
Conv1_2	64	3 × 3	64 × 224 × 224	36928
Maxpool_1		2 × 2	64 × 112 × 112	0
Conv2_1	128	3 × 3	128 × 112 × 112	73856
Conv2_2	128	3 × 3	128 × 112 × 112	147584
Maxpool_2		2 × 2	128 × 56 × 56	0
Conv3_1	256	3 × 3	256 × 56 × 56	295168
Conv3_2	256	3 × 3	256 × 56 × 56	590080
Conv3_3	256	3 × 3	256 × 56 × 56	590080
Maxpool_3		2 × 2	256 × 28 × 28	0
Conv4_1	512	3 × 3	512 × 28 × 28	1180160
Conv4_2	512	3 × 3	512 × 28 × 28	2359808
Conv4_3	512	3 × 3	512 × 28 × 28	2359808
Maxpool_4		2 × 2	512 × 14 × 14	0
Conv5_1	512	3 × 3	512 × 14 × 14	2359808
Conv5_2	512	3 × 3	512 × 14 × 14	2359808
Conv5_2	512	3 × 3	512 × 14 × 14	2359808
Maxpool_5		2 × 2	512 × 7 × 7	0
Flatten			25088	0
Dense_1			4096 × 1 × 1	102764544
Dropout_1			4096 × 1 × 1	0
Dense_2			4096 × 1 × 1	16781312
Dropout_2			4096 × 1 × 1	0
Dense_3			2 × 1 × 1	8194
Total params: 134,335,400, Trainable params: 134,335,400, non-trainable params:0				

Architecture of VGG16 shown in Fig. 5.14 and VGG16 model are presented in Table 5.1. This indicates that the VGG16 network is quite large, with a total of over 138 million parameters. A 224×224 RGB image serves as the input to the VGG-based convNet. The preprocessing layer subtracts the mean image values derived for the complete ImageNet training set from the RGB image with pixel values in the range of 0-255. After preprocessing, the input images are run through such weight layers. Convolution layers are stacked and then applied to the training images. In the VGG16 architecture, there are a total of 13 convolutional layers and 3 fully connected layers. Instead of having huge filters,

VGG uses smaller 3×3 deeper filters. The result is that it now has the same effective receptive field as if there were just one 7×7 convolutional layer.

Working of VGG16

- The first two layers are convolutional layers with 3×3 filters. The first two layers employ 64 filters, resulting in a volume of $224 \times 224 \times 64$ due to the employment of the same convolutions. Filters are always three by three with a one stride.
- After that, a pooling layer with a max-pool of 2×2 size and a stride of 2 was employed to reduce the volume's height and breadth from $224 \times 224 \times 64$ to $112 \times 112 \times 64$.
- The next two convolution layers have 128 filters each. The new dimension as a result is $112 \times 112 \times 128$.
- Volume is decreased to $56 \times 56 \times 128$ once the pooling layer is employed.
- Following the addition of two further convolution layers with 256 filters each, the size is decreased to $28 \times 28 \times 256$ using downsampling.
- A max-pool layer separates two more stacks, each of which has three convolution layers.
- $7 \times 7 \times 512$ volume is flattened into a Fully Connected (FC) layer with 4096 channels and a softmax output of 1000 classes after the last pooling layer.

5.2.2 Modified VGG16 (MVGG16)

As indicated in Table 5.3 and Fig. 5.15 (b), this study suggests MVGG16, a modified VGG16 that decreases the depth of the VGGNet network. To prevent both under and overfitting issues during training, the suggested network architecture minimizes the number of parameters by decreasing the network depth in comparison to the original VGG16. By doing feature extraction with two consecutive small convolutional kernels rather than a single large one, the original VGG16 convolutional architecture was preserved. This decreases the number of parameters while maintaining the VGG16 perceptual effects, which speeds up training.

Architecture of VGG16 shown in Fig. 5.15 (a) and VGG16 model are presented in Table 5.1. This indicates that the

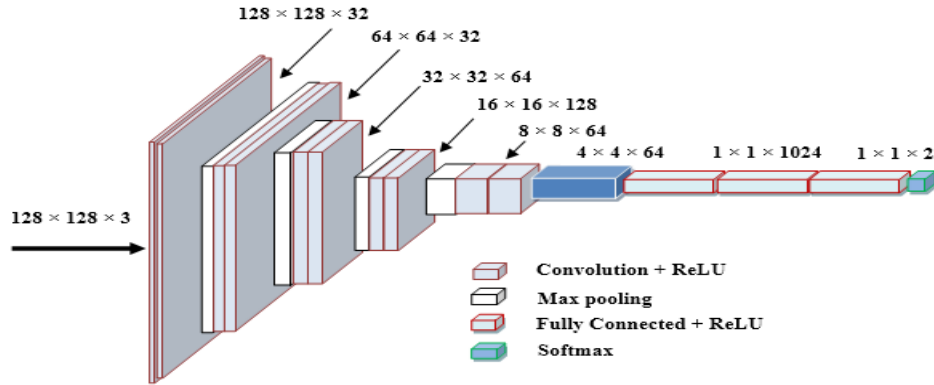


Fig. 5.14 MVGG16 architecture [126]

By modifying the input size, number of kernels, and convolutional layer from the original VGG16, MVGG16 has been constructed. MVGG16 Architecture is depicting in Fig. 5.16. There are 13 weight layers in MVGG16, which including 10 convolutional layers, 5 pooling layers, 3 fully connected layers. The MVGG16 algorithm was developed as a way to reduce the number of parameters while maintaining the network's depth, avoiding overfitting and underfitting problems, and speeding up computation.

Table 5.3 Comparison between VGG16 and MVGG16

Parameters	VGG16	MVGG16
Depth	16	13
Input Size	224x224	128x128
No. of kernels	64, 128, 256, 512 & 512	32, 32, 64, 128 & 64
Size of Kernels	3x3	3x3
No. of Fully Connected Layers	3	3
Total params	134,335,400	2,515,425

Working of MVGG16 - Initially, the input image size was modified to 128×128 , and the hidden layer was separated into five blocks, each of which contained two convolutional

layers and a pooling layer. Each convolutional layer extracted features using 32 randomly generated 3×3 convolutional kernels, and the pooling layer reduced the image with 2×2 convolutional kernels. Blocks 3-5 had convolutional kernels that were 3×3 in size, however there were 64, 128, and 64 kernels in each block, respectively. When compared to VGG16, the number of parameters needed was decreased by decreasing the number of convolutional kernels. After that, the pooling layer reduced the size of the image, the flattened layer reduced feature mappings to one dimension, and three fully connected layers combined output features into classification. Table 5.3 illustrate the comparison between VGG16 and MVGG16.

For calculating parameters following formula has been used

$$(m * n * d + 1)k \quad (5.12)$$

Where m is width and n is height of the filter, d is the depth of the input and k is the length of the next stage filter set.

Table 5.4 Proposed MVGG16 network model

MVGG16 Layers	Number of Kernels	Kernel Size	Output Size	Parameters
Conv1_1	32	3 × 3	32 × 128 × 128	896
Conv1_2	32	3 × 3	32 × 128 × 128	9248
Maxpool_1		2 × 2	32 × 64 × 64	0
Conv2_1	32	3 × 3	32 × 64 × 64	9248
Conv2_2	32	3 × 3	32 × 64 × 64	9248
Maxpool_2		2 × 2	32 × 32 × 32	0
Conv3_1	64	3 × 3	64 × 32 × 32	18496
Conv3_2	64	3 × 3	64 × 32 × 32	36928
Maxpool_3		2 × 2	64 × 16 × 16	0
Conv4_1	128	3 × 3	128 × 16 × 16	73856
Conv4_2	128	3 × 3	128 × 16 × 16	147585
Maxpool_4		2 × 2	128 × 8 × 8	0
Conv5_1	64	3 × 3	64 × 8 × 8	73792
Conv5_2	64	3 × 3	64 × 8 × 8	36928
Maxpool_5		2 × 2	64 × 4 × 4	0
Flatten			1024	0
Dense_1			1024 × 1 × 1	1048576
Dropout_1			1024 × 1 × 1	0
Dense_2			1024 × 1 × 1	1048576
Dropout_2			1024 × 1 × 1	0
Dense_3			2 × 1 × 1	2048
Total parameters: 2,515,425, Trainable parameters: 2,515,425, non-trainable params:0				

5.3 Results and discussions

5.3.1 Dataset

IIT Delhi and SDUMLA-HMT [200] datasets were both used in this study. First, the face, ear, and iris unimodal recognition model is trained and tested using two datasets. The proposed feature level fusion based multimodal system's performance is then assessed using SDUMLA-HMT and IIT Delhi.

The multimodal biometrics dataset SDUMLA-HMT was produced by a group of machine learning and applications researchers at Shandong University. A variety of biometric information, including the face, iris, finger vein, fingerprint, and gait, was collected from 106 individuals and stored by SDUMLA-HMT. Images for 61 men and 45 women between the ages of 17 and 31 can be found in SDUMLA-HMT. The dataset comprises of various images for each subject's five biometric characteristics. From the 106 subjects, 1060 iris images were collected by SDUMLA-HMT. The University of Science and Technology of China created an intelligent iris capture system that was used to capture the iris images. The eye and the equipment were within a 6 cm to 32 cm range during the capture operation. Regarding facial images, SDUMLA-HMT has 8904 distinct images of 106 persons in various poses, with various facial emotions, illuminations, and accessories.

The IIT Delhi ear database has 375 images from 125 different people, for a total of three images per person. The resolution of each of these images is 272 x 204 pixels. Fig. 5.15 provide samples of face, ear and iris images from SDUMLA-HMT and IIT Delhi database.

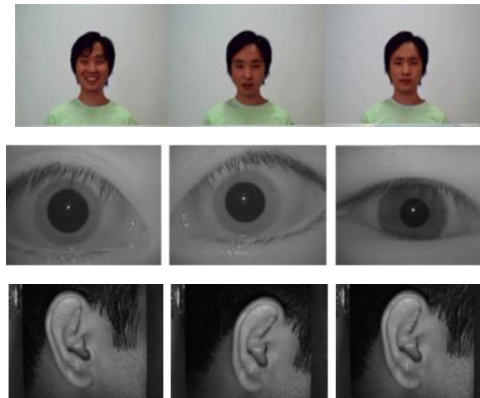


Fig. 5.15 Sample images of face, ear and iris traits from SDUMLA-HMT and IIT Delhi database

In SDUMLA-HMT and IIT Delhi, the images of each subject (class) were sorted at random into training, validation, and testing sets using various percentages (90:15:5), (80:10:10), (60:20:20) and (70:20:10). The data for each subject is finally divided into 60:20:20, 60 percent for training, 20 percent for validation, and 20 percent for testing, because this ratio produced the greatest results.

For training, validation, and testing, the dataset images were divided into three folders, each of which comprises samples for each subject. The validation set was used to assess the final model fit using only the forward pass, while the training set was utilized to train and fit the deep learning model utilizing continuous forward and backward passes through it.

5.3.2 Performance Metrics

The major goal of this work is to suggest a reliable multimodal recognition system that makes use of the face, ear and iris modalities. We employed false acceptance rate (FAR), false rejection rate (FRR), EER, and accuracy metrics to assess the effectiveness of the suggested biometric recognition designs. A key performance parameter in biometric recognition systems is the recognition time.

5.3.3 Experimental Parameters

In this analysis, all the training and testing stages were implemented using a system with Intel® Core™i7-64720 HQ CPU @ 2.60 GHz (4cores) and NVIDIA GeForce GTX 1650. MATLAB® R2021a and Microsoft Windows 10 Pro 64-bit.

Experiments on Unimodal framework

Numerous factors were taken into account during training the unimodal model, including the optimizer, mini batch size, learning rates, epochs, and dropout values. After adjusting the learning rate to 0.0004 and choosing a batch size of 64 and 70 epochs, it was discovered that the best model was produced. Prior to the classifier, the dropout layer with a rate of 0.5 was introduced. For optimization and loss function, the Stochastic Gradient Descent

with Momentum (SGDM) and categorical cross-entropy methods were used. Training hyper parameters of MVGG16 model for unimodal demonstrate in Table 5.5.

Experiments on Multimodal Framework

The optimizer, mini batch size, learning rate, epoch, and dropout values were main hyper parameter of the variables taken into consideration during training the multimodal model. Training parameters of MVGG16 model for unimodal and multimodal demonstrate in Table 5.5. The best model was created after the learning rate hyper parameter was changed to 0.0003, a batch size of 128 was selected, and 70 epochs were used. The dropout layer with a rate of 0.3 was introduced before the classifier. The categorical cross-entropy approach and Adaptive Moment Estimation (ADAM) were utilized for optimization and loss function.

Table 5.5. MVGG16 training parameters for unimodal and multimodal

Parameters	MVGG16 Unimodal	MVGG16 Multimodal
Input Size	128x128	128x128
Optimizer	SGDM	ADAM
Epochs	70	70
Learning Rate	0.0004	0.0003
Mini Batch Size	64	128
Dropout	0.5	0.3

5.3.4 Result Analysis

In this section, we discuss the result from our experiments and a performance assessment of our suggested approach.

Unimodal recognition- To extract informative features, CNN network (MVGG16) re-training is used, as shown in Section 5.2. The extracted data was then classified using Softmax classifier. A biometric system was carried out using the SDUMLA-HMT and IIT Delhi datasets, respectively, for the face, ear, and iris. The performance findings of the

proposed face, ear, and iris unimodal recognition systems are summarized in Table 5.6. For the face, ear, and iris, respectively, MVGG16 achieved recognition times of 1.98 seconds, 2.22 seconds, and 2.54 seconds with accuracy rates of 100 percent (EER of 0 percent), 95 percent (EER of 0.49 percent), and 90 percent (EER of 1.00 percent).

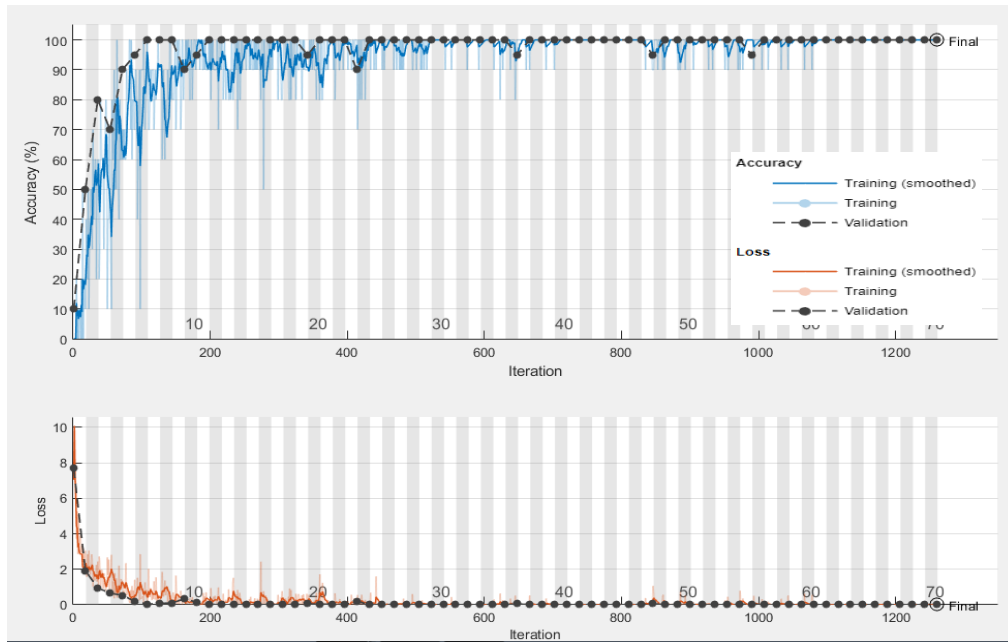


Fig. 5.16 Training process- accuracy and loss – unimodal face

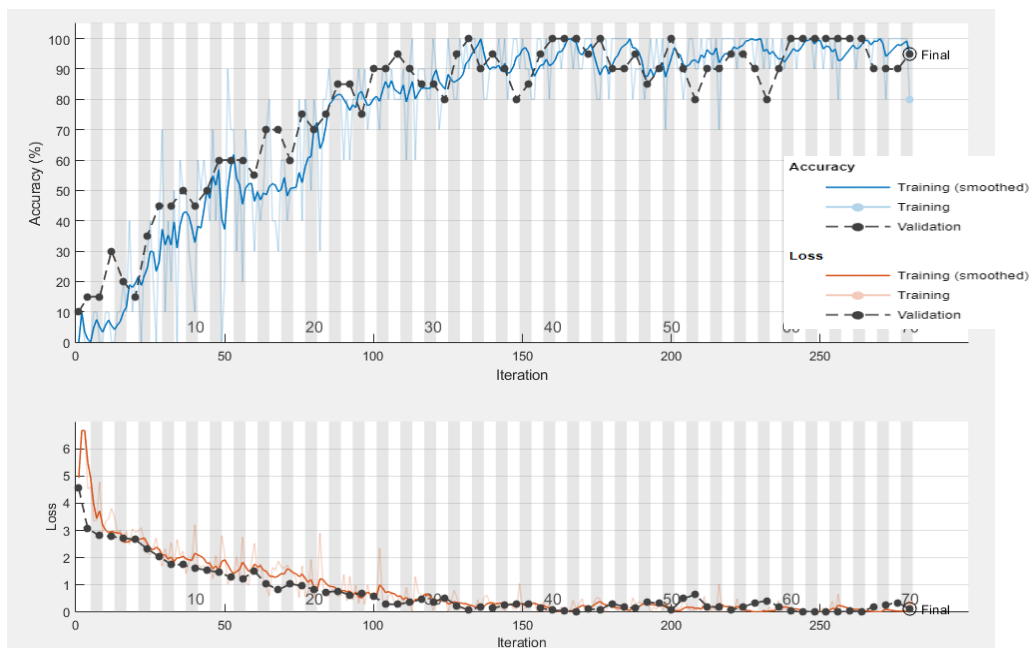


Fig. 5.17 Training process- accuracy and loss – unimodal ear

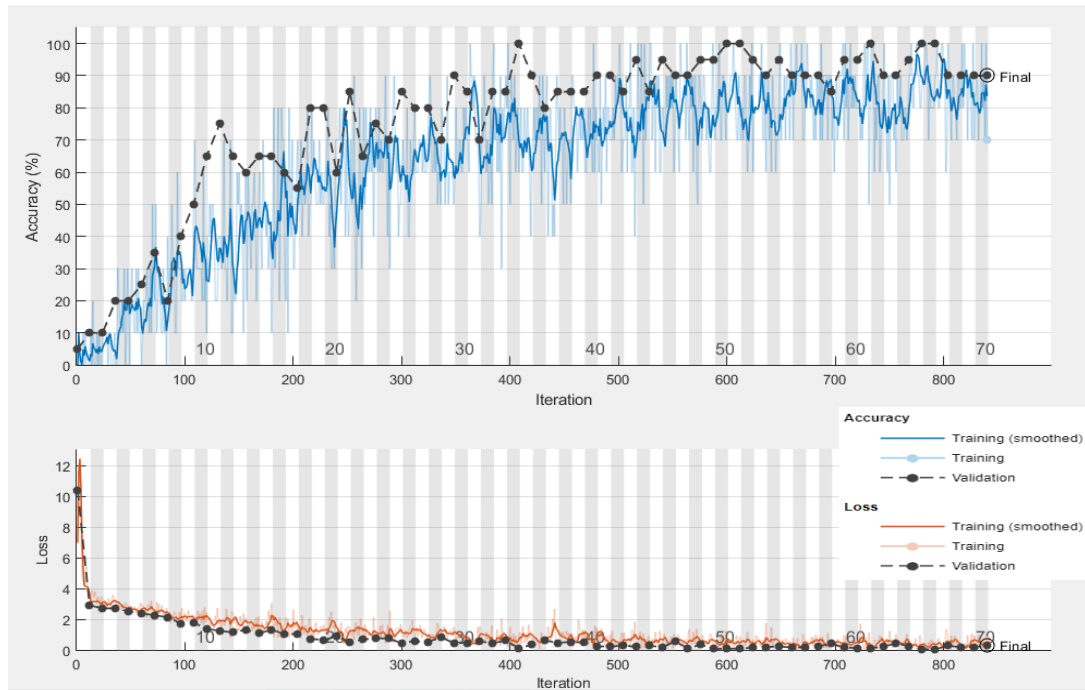


Fig. 5.18 Training process- accuracy and loss – unimodal iris

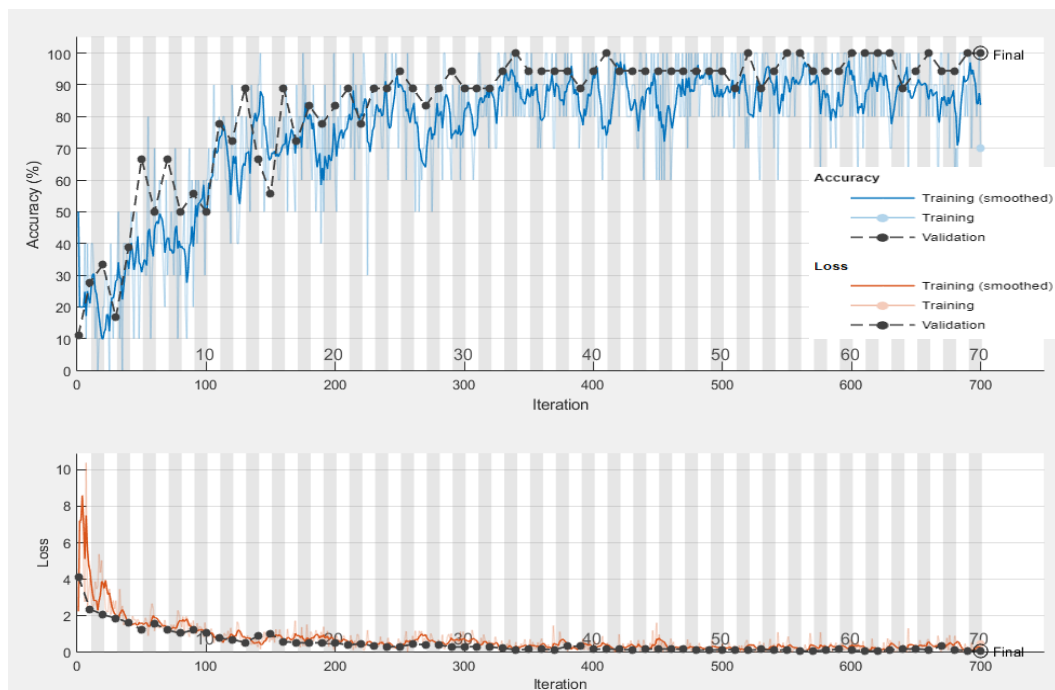


Fig. 5.19 Training process accuracy and loss–multimodal (face + ear + iris)

Fig. 5.18, Fig. 5.19, Fig. 5.20 and Fig. 5.21 illustrate training process accuracy and loss value of MVGG16 for face modality based unimodal, ear modality based unimodal, iris

modality based unimodal, and integrated face, ear, and iris modality based multimodal, respectively.

Table 5.6 Performance analysis of unimodal recognition using MVGG16

Biometric Model	Traits	Accuracy (%)	EER (%)	Recognition Time
Unimodal	Face	100	0	1.98s
	Ear	95.00	0.05	2.22s
	Iris	90.00	0.10	2.10s

Multimodal recognition- In this work, three multimodal systems that combine the modalities of the face, ear, and iris are proposed. This combination is realized at the feature level (feature level fusion technique) in this study. Results of the feature level fusion approach are summarized in Table 5.7. The informative features were extracted using MVGG16, and Softmax was then employed as a classifier after fusion of features. With regard to feature level fusion, MVGG16-Softmax demonstrated the highest accuracy of 100 percent (EER=0 percent), as well as a shorter recognition time of 3.1seconds. The performance of MVGG16 is superior to unimodal recognition systems.

Table 5.7 Performance analysis of multimodal recognition using MVGG16

Biometric Model	Fusion Type	Accuracy (%)	EER (%)	Recognition Time
Multimodal (Face, Ear and Iris)	Feature Level Fusion	100	0	3.1s

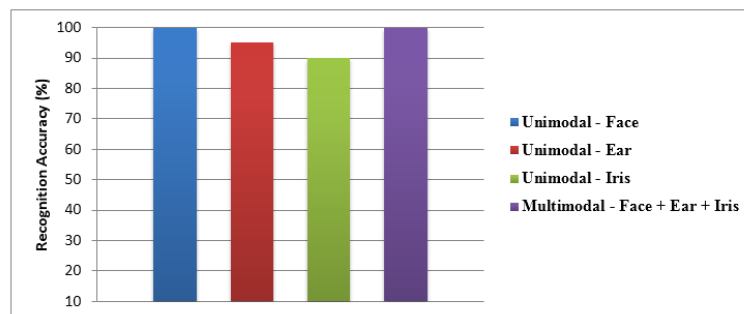


Fig. 5.20 Recognition accuracy for unimodal and multimodal using MVGG16

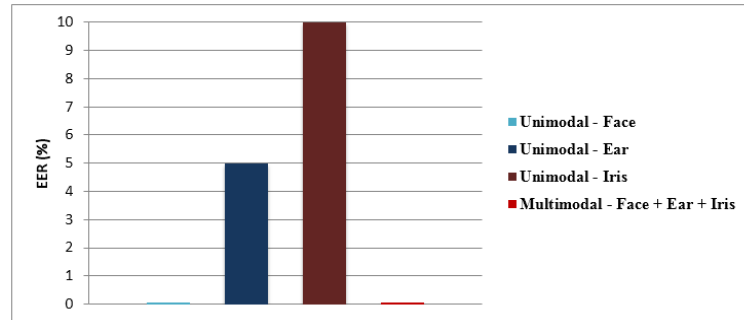


Fig. 5.21 EER analysis for unimodal and multimodal using MVVG16

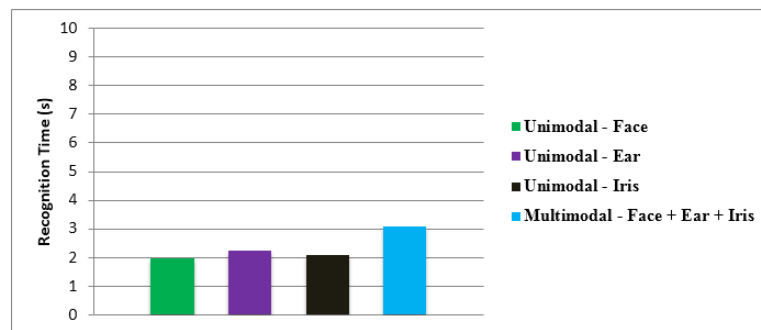


Fig. 5.22 Recognition time for unimodal and multimodal using MVVG16

Performance analyses of unimodal (face, ear, and iris) and multimodal (face + ear + iris) feature-level recognition using MVGG16 are shown in Fig. 5.22, Fig. 5.23, and Fig. 5.24 in terms of accuracy, equal error rate, and calculation of recognition time.

Comparative Study

Our suggested feature level fusion method is contrasted with the work of S. Velachery et al. (Multi-SVM with FFF and Multi-SVM with firefly algorithms) [201], as well as Yang et al. proposed OriCode and C2 Code algorithms [202] and Yang et al. developed Weighted Fusion and Cross-Section Binary Coding methods [203]. Table 5.8 provides a summary of the performance comparison findings for multimodal systems. When compared to earlier efforts, our MVGG16 architecture's performance in feature-level fusion is improved.

Table 5.8 Comparative performance analysis of multimodal biometric recognition

Reference	Methods	Accuracy (%)	EER (%)	Recognition Time (s)
Velachery [9]	Multi-SVM with FFF	95	0.35	5.1
	Multi-SVM with firefly	94	0.7	5.1
Yang et al. [10]	OriCode ^k	-	0.889	-
	C ² Code	-	0.435	-
Yang et al. [11]	Weighted fusion	99.84	0.16	-
	Cross section binary coding	99.67	0.31	-
Purohit Ajmera	Proposed MVGG16	100	0	3.1

5.4 Summary

To summarize, a multibiometric model for user identification was developed in this work. The MVGG16 CNN deep learning algorithm was applied in the proposed system. In order to recognize the user from their face, ear, and iris attributes, feature level fusion was used. This study is the first that we are aware of to look into the application of deep learning methods for a multimodal biometric model with these three traits. Three MVGG16 CNNs were employed in the proposed model to identify each attribute. The SDUMLA-HMT and IIT Delhi dataset were used to assess the model's performance. The experimental findings demonstrated the MVGG16 CNN algorithm. It also demonstrated that using three biometric features rather than just two or one can improve the performance of identification systems.

Instead of utilizing a pretrained model in future work, tailored CNNs from scratch for each attribute. Creating a CNN for iris, for instance, which accepts circle images. The impact of applying deep learning algorithms to a larger range of recognition qualities, such as DNA, signatures, or hand shape, can also be investigated. It will also be exciting to expand the kind of experiments used to evaluate the proposed model using different multimodal datasets and level fusion techniques.

Chapter 6

SEM-ANN based Analysis of Users' Awareness and Acceptability of Multimodal Biometrics

6.1 Introduction

Human Machine Interface is the key research area of contemporary computational technologies [180]. According to Breitinger et al. [180] & Nurse et al. [181], the primary aim of the research advancement is to serve humankind and create an eco-system of continuous support for the betterment of the world. In the words of Anil et al.[183]; Faheem et al.[184]; Khayer et al. [185], interaction with machines for a different purpose is conspicuous. Its success depends upon user acceptability and ease of application. Data privacy and security are also paramount factors in this digital era that is unpredictable in the fast-changing technical development process, especially in financial transactions of all magnitude [186].

Scholarly evidence stated the use of biometrics authentication for financial transaction in offline mode ensures the authenticated user's presence while operating at the point of sale [187],[188],[189]. Moreover, the authentication and verification process during online transactions enhance reliability and improved security from the user and service providers [190].

According to Darganet al. [191]; Durdyev et al.[192],biometric systems are traditionally classified in two key areas i.e., Unimodal and Multimodal biometrics. Unimodal is the system based on a single biometric trait, unlike a multimodal system where more than one physiological or behavioral trait is used to recognize the person. Popular physiological

traits are the face, ear, iris, fingerprint, finger vein, and knuckle, whereas behavioral traits are gait, keystroke, voice, eye movement, facial gradient etc. Over the years, quality datasets of all kinds of traits have been developed for research purposes and are available for experimental testing for academic and commercial uses [193].

Researchers have combined multiple traits for identification problems to improve performance and better data security, known as Multimodal biometric (MMB) fusion. Due to its incredible efficiency, cheap cost, and convenience, biometric technology has become the most extensively used human identification and authentication technology in both the public and private sectors [194].

In the word of biometric systems, also known as Identity Certification (I.C.) systems, are the art of constructing authentication procedures using biometric traits to identify automatically, measure, and validate a living human [195]. The biometric system is based on the concept that everyone is unique in cognitive and behavioral characteristics. Identifiers are permanent, one-of-a-kind, and distinct from one another. The purpose of developing such systems is to improve the digital world's safety and security. Science, security, espionage, identifying information, and commerce all require biometric recognition systems to ensure user authentication and identification. Because it is difficult for a forger to detect and spoof a registered person's biometric modality, biometric technologies have grown popular. High accuracy rates and the difficulty of spoofing are the major criteria that separate old security solutions [195].

In online trade transaction MMB has shown better results in terms of more reliability, enhanced security, and safe transaction [196]. On the other hand, the unimodal process where single traits are utilized for authentication is comparatively more prone to imposter attacks and has higher chances of eye dropping [1].

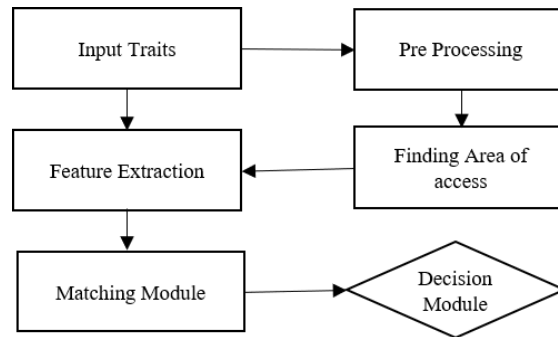


Fig. 6.1 Universal framework of biometric system

The last decade has been a transition period for electronic payment-based application and commercial uses [197],[198],[199]. The rapid development of processors and high-speed internet has provided a foundation for digital payment systems in this era. However, traditional methods like memory-based password gateway are not safe to give such transactions a high level of security and reliability. Instead of fusion of face, ear, iris, keystroke, voice, and fingerprint provides a robust mechanism for authentication [199].

Some Researchers [200],[201], [202], also stated various fusion techniques which can be classified into different categories like rank level, score level, decision level, and feature level, advantages and challenges associated with these methods. For the applications where accurate information requires for decision making, feature level methods are used. The level of fusion and feature extraction position shall decide the name of the methods such as score level fusion [203].

According to Jhaveri, et al. [204], “supervised learning and pattern acknowledgement are critical research areas in information retrieval, data engineering, medical image processing, and intrusion recognition”. This work aims to identify a robust classifier, which can be considered a network-based intrusion detection system.






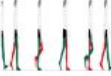




Fingerprint		Iris	
Finger Vein		Retina	
Facial		Gait	
Palm Vein		Hand Geometry	
Handwriting		Speech	

Fig. 6.2 Typical biometric attributes [16]

Another incredible research has been done by Ijaz et al. [205], using the dataset of 858 budding cancer patients to corroborate the performance based on the combinations of iForest with SMOTE and iForest with SMOTETomek. Besides, the study used a mobile application that can gather data on cervical cancer risk factors and delivers results from CCPM for immediate and appropriate accomplishment at the early phase of cervical cancer.

To fulfil the objectives of the research, the following section is conducive to comprehend. Section 6.2 conceptual model has been discussed, Section 6.3 portrays the theoretical constructs, and development of hypotheses. Section 6.4 outlines the research frame and methodology to assess customer awareness and acceptability of biometric transactions. Section 6.5 illustrates data analysis, SEM, testing of hypotheses, followed by a procedure of ANN modeling, sensitivity analysis and interpretation with the help of statistical tools. The last section considers theoretical implications, limitations, future scope, and conclusions.

6.2 Conceptual Model

The model's constructs were evaluated using artifacts found in the existing literature. A literature review of previous works creates the measurement constructs, and the present study seeks to identify and determine the factors influencing cloud computing implementation in Learning Management System (LMS) [206]. The notion of biometrics adoption was evaluated using 21 items (five-point Likert scale) referred from published sources. The constructs viz. User Acceptability, Cognizant Factors towards Biometrics, Technological factors, Perceptual Factors (Fingerprints, Iris, Face Recognition and Voice), and Data Privacy Factor comprise certain items suggested by past studies [191],[1].

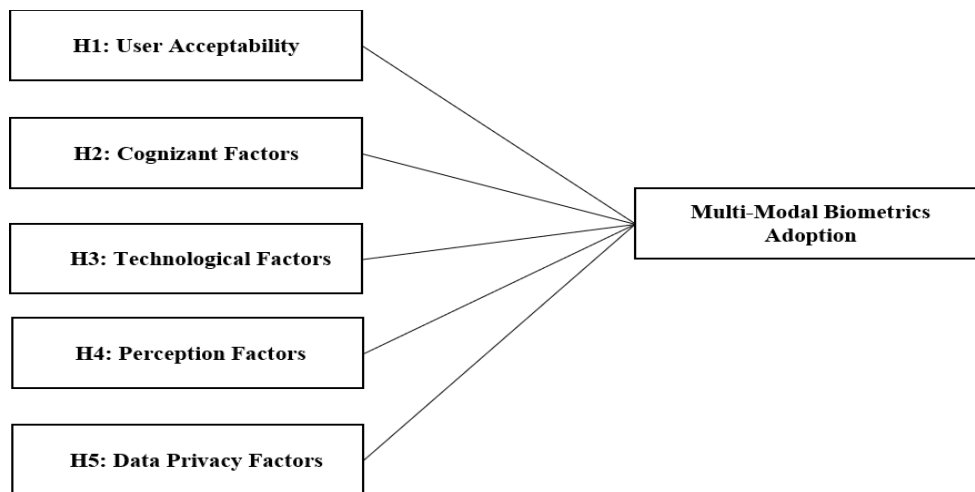


Fig. 6.3 Conceptual framework of the study

This research work contributes a survey related to customer awareness and acceptability of biometric mechanisms while transacting online. As a result, the proposed questionnaire is prepared and receives responses from a variety of customers. Based on feedback received from technical experts/users, hypotheses of the proposed model are tested by framing Structural Equation Modelling (SEM) and followed by artificial neural network (ANN). This is one of the few studies that try to investigate the factors that affect MMB adoption. As a result, the established model makes a significant contribution to the literature in this field.

6.3 Theoretical Constructs and Development of Hypotheses

Diverse factors of biometrics adoption mechanism were examined, and their result was surveyed to find out the degree to which they are embraced. To examine the objectives-user acceptability, cognizant factors, technological factors, perceptual factors, and data privacy were exerted to determine the impact on the adoption performance of the virtual consumers. Distributions of important variables were included in the investigation reported in previous studies. Table.6.1 highlights the constructs and their reported significance in the literature. In different phrases, primary and higher-level data analysis is being done on the selected constructs.

There is stepwise extraction of manifests from past research work, i.e., users' acceptability criterion towards biometrics usage while performing the online transaction, the notion of the Cognizant Factors towards Biometrics usage by the users, Technological factors that evoke the practices of biometrics applications. Subsequently, Perceptual Factors (Fingerprints, Iris, Face Recognition and Voice) influence MMB, followed by data privacy & security while using biometrics submission.

6.3.1 Hypotheses for the Research

The following hypotheses are proposed based on the literature review and conceptual model:

H₁: There is a connotation between user acceptability and MMB while transacting online.

H₂: There is implicit association between technical factors and the adoption of MMB adoption.

H₃: There is evidence of cognizant Factors that influence MMB.

H₄: Perceptual Factors (Fingerprints, Iris, Face Recognition and Voice) strongly associate with MMB.

H₅: There is an association between data privacy factors and the adoptive operation of biometrics operations.

Table 6.1 Multimodel biometrics and execution measures

Constructs	Instruments for survey	Sources where instruments adopted from
User Acceptability	<p>UAC_1: Acquaint about the protection of online privacy.</p> <p>UAC_2: Accuracy and flexibility of biometrics mechanism.</p> <p>UAC_3: Biometrics would help make living wills more eagerly accessible.</p> <p>UAC_4: Control over personal information released online.</p> <p>UAC_5: Acquaint to the extent to which information is shared with other company.</p>	(Gokulkumari 2020; Ioannou, Tussyadiah, and Lu 2020; Jackson 2009; Teh et al. 2016; Vereycken et al. 2019)
Cognizant Factors towards Biometrics	<p>CGF_1: Level of relying on biometrics during an online transaction.</p> <p>CGF_2: Recognition level about Biometrics mechanisms.</p> <p>CGF_3: Using biometrics for safety and privacy.</p> <p>CGF_4: Biometrics is a better tool during Covid.</p>	(Al-Ahmari et al. 2020; Asadi and Abdekhoda 2020; Dinh, Nguyen, and Nguyen 2018; Sinha and Ajmera 2019)
Technological Factors	<p>TCF_1: Capability for gaining a competitive advantage.</p> <p>TCF_2: Biometrics provide a flexible and robust solution during an online transaction.</p> <p>TCF_3: Biometrics is conducive to the existing legacy systems.</p> <p>TCF_4: Accessibility of I.T. support and infrastructure.</p>	(Anabel Gutierrez 2015; Psomas and Jaca 2016; Raut et al. 2018; Singh and Dadhich 2021; Tornatzky, L.G., Fleischer 1990)
Perceptual Factors (Fingerprints, Iris, Face Recognition and Voice)	<p>PRF_1: Facial biometrics need different tactics such as voice, posture.</p> <p>PRF_2: Using MMB is valuable and safe.</p> <p>PRF_3: Biometric authentications ensure virtual security.</p>	(Alsadoon et al. 2016; Liébana-cabanillas et al. 2018; Raut et al. 2018; Sinha and Ajmera 2019)

	PRF_4: We need to incorporate more specific variables in MMB.	
Data Privacy Factors	<p>DPF_1: Aware of the privacy concerns associated with biometric traits.</p> <p>DPF_2: Willing to purchase a high-priced item online biometrics mode.</p> <p>DPF_3: Concerns about your privacy while making an online purchase.</p> <p>DPF_4: Security of online transactions should improve by biometrics</p>	(Carrión-ojeda et al. 2021; Dadhich 2017; Dadhich et al. 2018; Dargan and Kumar 2020; Gokulkumari 2020; Joshi et al. 2020; Teh et al. 2016)

6.4 Research Methodology

The study is a perfect blend of a quantitative-qualitative frame conducive to exploring the determinants of adoption of biometrics mechanism from users' viewpoint. All the elements used to quantify the study variables were adapted from previous research, with slight terminology adjustments to acclimate them to the unique biometrics' context from the users' perspective [207].

6.4.1 Survey Instruments and Data Collection

The survey was conducted by focus group discussions and walk-through evaluations of selected respondents who indulged in biometrics features while conducting online transactions. This confirmed the content validity of the questionnaire used to gauge each unobserved variable stated in Table 6.2. This objective was accomplished using the observational approach of focus groups [208]. In this instance, a few focus groups were formed comprising of individuals with a working knowledge of biometrics applications. The study's subject was assigned to the selected groups to familiarize them with biometrics terminology before filling out the final questionnaire. The questionnaire's survey was developed using these inputs and previous research findings.

Primary data is collected using Google form from plentiful individuals using the online transaction. In addition, a personal interview was conducted to obtain the required answer, i.e. biometrics data. This personal interview cleared the doubts of the respondents about an acquaintance of the questionnaire based on biometrics awareness. Similarly, the sample size of the study is based on the following calculation [209]:

$$\frac{Q_{\alpha} \pm Q_{\beta}}{\mu_1 \pm \mu_2} \quad (6.1)$$

Where: n - number of sample size

σ - Standard deviation as considered at 1.2 based on prior studies.

' Q_{α} - Upper tail in the Normal Distribution (SND)'.

Corresponding to $Q_{\alpha} = 1.96, p - value = 0.05$.

' Q_{β} - Lower tail in the SND'.

Corresponding to $Q_{\beta} = -0.84$ at $\beta = 0.2$.

μ_1 and μ_2 - difference in means

6.4.2 Measurement Used in the Study

To elucidate the degree of correlation among the variables in an array of data sets, Exploratory Factor Analysis (EFA) is a technique to examine a factor structure. It facilitates the performance of certain functions, i.e., exploring the pattern of data configurator, illustrating the relationship among various patterns, and extracting valid data to the next level of analysis[209],[210].

To confirm the observed variables in a set of factor structures, Confirmatory Factor Analysis (CFA) is used to test the core hypothesis and evaluate the association between observed variables underlying their latent constructs. CFA is a significant and vital part of the modeling structure equation. It is advantageous and apt to formulate observed variables for measuring and validating the hypothesized model of latent constructs [211]. The researchers have used the CFA to establish relevance of items representing the proposed theory, empirical or scientific research, postulate the association framework a priori and

confirm hypotheses. Thus, CFA is a unique specified factor conjoint technique widely used to validate the model supported by any previous theory [212].

A priori model, number of factors, item loading on each factor, fit indices, error term, standard regression weights on items are elucidated by SEM. The goodness of fit should be considered to conceptualize the model practically. Moreover, Goodness of Fit (GOF) is inversely interrelated to sample size and the number of variables in the model. Therefore, it is vital and mandatory to view reliability in terms of Cronbach Alfa followed by establishing convergent and discriminant validity of a model [213],[214].

With the assistance of Statistical Package of Social Science (SPSS) and Analysis of Moment Structure (AMOS) tools, the registered data was listed, analyzed, and interpreted. Having considered 530 respondents from various respondents, mostly bankers, technical experts and engineers, the researcher intends to test the convergent and discriminant validity of the group variables. Therefore, the questionnaire consists of two sections. The first part encompasses six questions about the all-purpose profile of the experts and actors using biometrics while doing online transactions. The second part entails five major constructs along with 22 variables. The research took place from January to April of 2021.

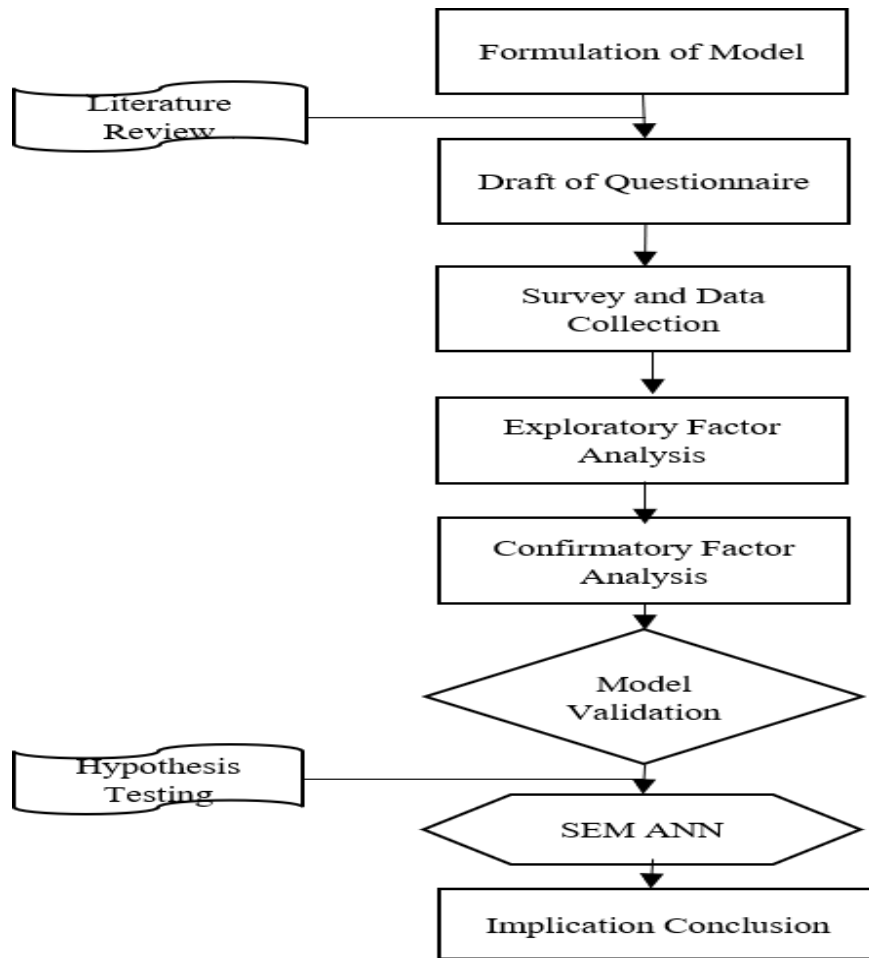


Fig. 6.4 Methodology of research

6.5. Data Analysis and Interpretation

The data analysis is done by examining responses to a set of survey questions intended to measure awareness and acceptability towards adopting MMB mechanism while doing online transactions. Prima facie, it is pertinent to discuss well-versed users' demographic profile and acquainted with the biometric application.

Table 6.2 Demographic profile

Factor	Classification	Frequency	%
Gender	Male	280	52.80
	Female	250	47.20
Age	Below 30	530	100.00
	30-40	253	47.70
	Above 40	231	43.60
Education	U.G.	46	08.70
	P.G.	330	62.30
	Others	105	19.80
Work Experience	Less than 2years	95	17.90
	2-5 years	235	44.30
	Above 5 years	182	34.30
Occupation	Service	113	21.40
	Business	530	100.00
	Professional	227	42.80
Management Level	Junior Management	212	40.00
	Middle Management	91	17.20
	Top Management	530	100.00

Demographic profile depicted in Table 6.2 enunciates that the sample comprised 52.80% male and 47.20% female, principally 47.70% respondents were below the age set of 30 years, and 43.60% were between 30-40 years 08.70% persons were above the age of 40. Besides, most of the technical respondents were U.G., i.e., 62.30%, P.G. professionals 19.80%, and others were only 17.90%. Experts having work experience of fewer than two years 44.30%, 2-5 years 34.30%, above five years 21.40% were the sample population. Among the respondents, major share 42.80% of service class, businesspersons were 40%, share statistics of professional was 17.20%. Eventually, most defendants were from the junior management level, i.e., 36.80%, middle management scored 46.40%, and top management share was 16.80%. Prima facie, the research aims to test the reliability of selected items that have already been discussed; Cronbach's Alpha expressed the score of 0.825 that summarized that the observed variables were reliable enough for confirmatory statistical analysis.

Harman's single variable test

The statistical distribution of the constructs used in the analysis outlines in Table 6.3. To ascertain the symmetry distribution and peakedness of the gathered data, skewness and kurtosis were determined. (i) skewness and kurtosis reflect the closeness of the obtained distribution from normal distribution. Skewness can be right, left, and symmetrical. In positive skewness tail is extended toward right and mean is larger than median, mass distribution is inclined towards left. In negative skewness tail is extended toward left and mean is lesser than median, & mass distribution is inclined towards right. In general, skews are considered within the spectrum of alternatives +1/-1, and the values are usually distorted(Liu et al. 2020; Timans et al. 2016). (ii) Kurtosis is a measure of the tailedness of a distribution. Tailedness is how often outliers occur. Excess kurtosis is the tailedness of a distribution relative to a normal distribution. It highlights the outlier frequency in the distribution. In case of normal distribution, outlier frequency is moderate.

The analysis shows that respondents prioritize user acceptability (20.31), perceptual factors such as fingerprints, iris, facial recognition, speech etc. (15.54), technical factors (15.33), and cognizant factors towards biometrics (12.80) as primary reasons for embracing MMB, followed by data privacy factors (11.84). All in all, there is an asymmetrical distribution of results when it comes to biometric adoption.

6.5.1 Measurement Model

Determinants of biometrics mechanism adoption were used in research to evaluate the users' acceptance and awareness while transacting online. The understanding of the biometric behavioral knowledge and adoption constructs were discovered using an Exploratory Factor Study. Moreover, the Kaiser-Mayo-Olkin score was used to gauge the sample size's adequacy. The measured score of 0.874 intended that the samples were adequate to accomplish the factor analysis. Determinants of MMB were analyzed, and table 4 enunciated the G'F'I (0.936), A'G'F'I (0.812) along with N'F'I (0.866), and R'F'I (0.833)

whereas C'FI (0.920), Tucker-Lewis delineates 0.918 and the value of RMSEA was 0.035 that specifies that the anticipated model is in a decent fit [239].

Table 6.3 Statistical distribution of constructs

Constructs	N	X	Σ	Vari.	Skew.	Kurt.
User Acceptability (Cum. Mean=20.3132)						
UAC_1	530	4.0887	1.00550	1.011	-.817	-.127
UAC_2	530	4.1509	0.97213	0.945	-.889	0.004
UAC_3	530	4.0679	1.05926	1.122	-1.027	0.334
UAC_4	530	4.0453	1.10585	1.223	-1.049	0.245
UAC_5	530	3.9604	1.08797	1.184	-.708	-.423
Technological factors (Cum. Mean=15.3358)						
TCF_1	530	4.2434	0.93817	0.880	-1.357	1.466
TCF_2	530	4.1038	1.02270	1.046	-1.550	2.390
TCF_3	530	3.5849	1.16233	1.351	-.377	-.565
TCF_4	530	3.4038	1.04130	1.084	-.736	-.032
Cognizant Factors towards Biometrics (Cum. Mean=12.8075)						
CGF_1	530	3.0245	1.33350	1.778	-.146	-1.141
CGF_2	530	3.6094	1.41965	2.015	-.605	-.987
CGF_3	530	2.8906	1.42464	2.030	.217	-1.304
CGF_4	530	3.2830	1.33128	1.772	-.351	-1.016
Perceptual Factors (Fingerprints, Iris, Face, and Voice) (Cum. Mean=15.5491)						
PRF_1	530	3.4623	1.24332	1.546	-.420	-.888
PRF_2	530	3.9943	1.19939	1.439	-1.005	-.072
PRF_3	530	4.0491	1.13104	1.279	-1.017	.002
PRF_4	530	4.0434	1.25561	1.577	-1.037	-.207
Data Privacy Factors (Cum. Mean=11.8453)						
DPF_1	530	3.2226	1.21563	1.478	-.357	-.696
DPF_2	530	2.9849	1.25853	1.584	.154	-.945
DPF_3	530	3.0208	1.26070	1.589	.097	-1.008
DPF_4	530	2.6170	1.19486	1.428	.325	-.772

Table 6.4 Goodness-of-Fit for MMB adoption

Particulars	G'F'I	A'G'F'I	N'F'I	R'F'I	C'F'I	T'LI	RMSEA
Ceiling value	>0.900	>0.905	>0.985	>0.906	>0.910	>0.910	>0.01
Achieved value	0.936	0.812	0.866	0.833	0.920	0.918	0.035

The model exposed a good fit measurement i.e., $\lambda^2 = 951.69$, $df = 195$, $CMIN/df = 4.88$, at significant level $p = 0.038$. As typified in table 5, the factor loading, Cronbach alfa, AVE, and C.R. of all factors were highly significant [240].

Table 6.5 Validity and reliability test of standards

Statements	FL.	Cron. α	A-V-E	C-R
UAC_1	0.866	0.902	0.785	0.942
UAC_2	0.802			
UAC_3	0.725			
UAC_4	0.895			
UAC_5	0.816			
TCF_1	0.795	0.887	0.876	0.856
TCF_2	0.815			
TCF_3	0.866			
TCF_4	0.796			
CGF_1	0.812	0.891	0.785	0.839
CGF_2	0.782			
CGF_3	0.769			
CGF_4	0.785			
PRF_1	0.712	0.856	0.806	0.826
PRF_2	0.882			
PRF_3	0.912			
PRF_4	0.885			
DPF_1	0.722	0.887	0.881	0.823
DPF_2	0.882			
DPF_3	0.802			
DPF_4	0.795			

Possible options for normality, linearity, co integration and homoscedasticity were conducted before evaluating the proposed model. The data was then subjected to CFA to see if it matched the proposed theoretical model and confirm the gauged constructs' validity and reliability. The factor correlation among latent items has to be less than the square root of the average variance of each factor. According to Harman's single factor test, just 24% of the variance in all variables is explained by a single factor, demonstrating that CMV is not a problem in this analysis. Another test was run to ensure that no similarities were greater than 0.90, which may mean skewed data [217]. As a result, none of the measured correlations surpasses the suggested threshold, indicating that CMV is not a severe concern in this analysis. With all figures, it has been inferred that the model meets the standards of

reliability, the validity of substance, convergent validity, and discriminant validity. In this way, it was succeeded by the testing of the structural equation model.

Table 6.6 Fornell-Larcker criterion for discriminant validity

Variables	AVE	User Acceptability	Technological factors	Cognizant Factors	Perceptual Factors	Data Privacy Factors
User Acceptability	0.785	0.595				
Technological factors	0.876	0.614	0.672			
Cognizant Factors	0.785	0.526	0.543	0.662		
Perceptual Factors	0.806	0.679	0.748	0.479	0.405	
Data Privacy Factors	0.881	0.426	0.469	0.269	0.559	0.225

The discriminant validity infers the degree to which dormant items are varied from other dormant variables in the selected frame. Also, the construct correlation among dormant items must be less than the square root of AVE of every factor [218]. With these all standards, it has been confirmed that the model estimates the criteria of reliability, the rationality of matter with convergent validity. In this way, it was accomplished by the perusing of the SEM.

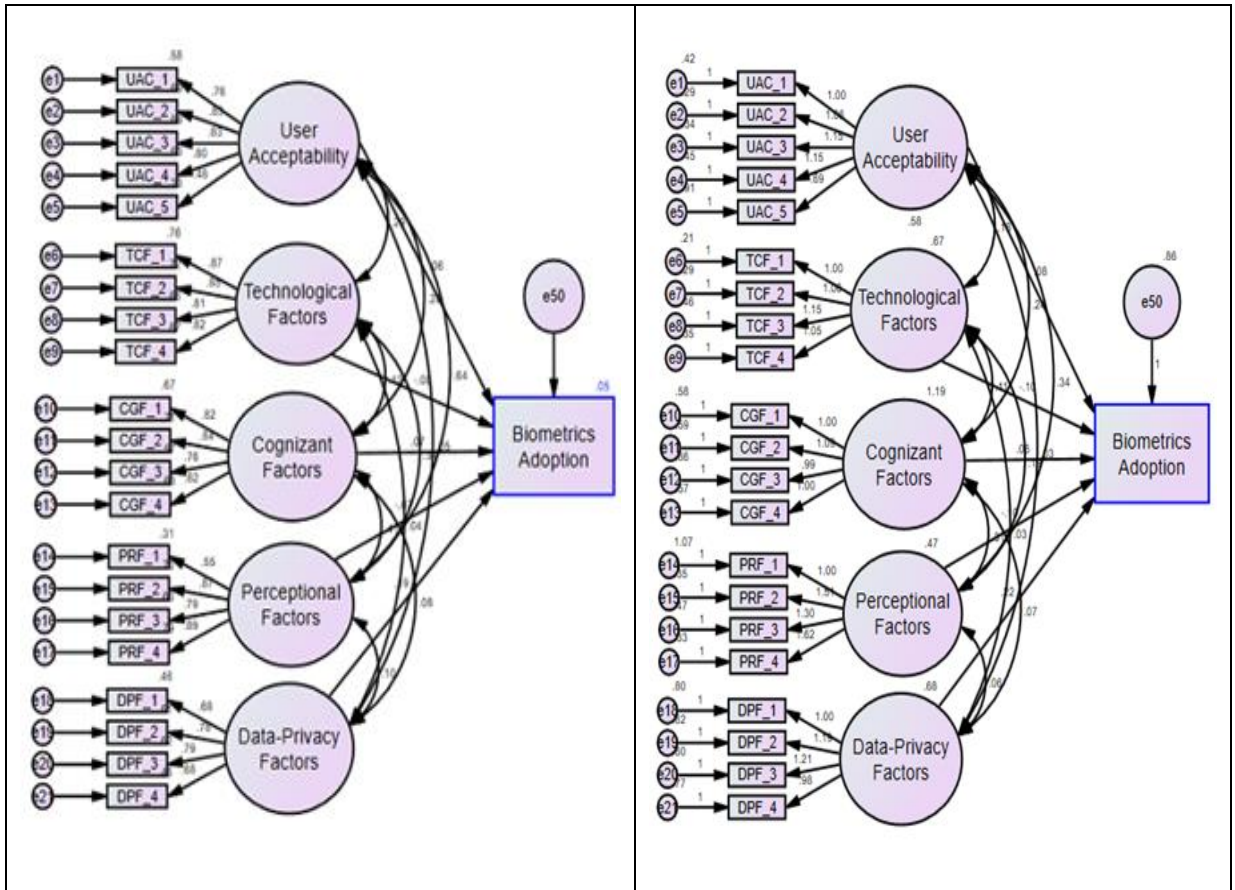


Fig. 6.5 Estimates of CFA model

Fig 6.5 divulged the confirmatory model's standard estimates and eventually examined the proposed hypotheses, structural model fit, and path analysis.

Table 6.7 Summary of standard regression weights of constructs

Items		Direction	Esti.	Std.-Er.	Cri. Ratio	P
UAC_1	<---	Users Acceptability	1.000			
UAC_2	<---	Users Acceptability	1.056	0.055	19.372	***
UAC_3	<---	Users Acceptability	1.152	0.059	19.405	***
UAC_4	<---	Users Acceptability	1.151	0.062	18.509	***
UAC_5	<---	Users Acceptability	0.686	0.064	10.770	***
TCF_1	<---	Technological Factors	1.000			
TCF_2	<---	Technological Factors	1.061	0.043	24.452	***
TCF_3	<---	Technological Factors	1.154	0.051	22.836	***
TCF_4	<---	Technological Factors	1.047	0.045	23.310	***
CGF_1	<---	Cognizant Factors	1.000			
CGF_2	<---	Cognizant Factors	1.090	0.051	21.430	***
CGF_3	<---	Cognizant Factors	0.986	0.052	18.849	***
CGF_4	<---	Cognizant Factors	1.002	0.048	20.934	***
PRF_1	<---	Perceptual Factors	1.000			
PRF_2	<---	Perceptual Factors	1.513	0.112	13.508	***
PRF_3	<---	Perceptual Factors	1.303	0.101	12.914	***
PRF_4	<---	Perceptual Factors	1.619	0.119	13.627	***
DPF_1	<---	Data Privacy Factors	1.000			
DPF_2	<---	Data Privacy Factors	1.190	0.083	14.420	***
DPF_3	<---	Data Privacy Factors	1.205	0.083	14.499	***
DPF_4	<---	Data Privacy Factors	0.982	0.075	13.043	***
BMO_1	<---	Users Acceptability	0.076	0.079	.964	.035
BMO_1	<---	Technological Factors	0.097	0.056	1.723	.045
BMO_1	<---	Cognizant Factors	0.062	0.044	1.407	.040
BMO_1	<---	Perceptual Factors	0.101	0.094	1.066	.020
BMO_1	<---	DataPrivacy Factors	0.217	0.056	3.907	.001

Table 6.8 Estimation of the hypotheses and comparison

S. N	Statements	Consistent with	Inconsistent with	Remarks
H ₁	There is a connotation between the user acceptability and biometrics adoption while transacting online, ($\beta = 0.076$, S. Er = 0.079, Cri. ratio = 0.096)	(Alsadoon et al. 2016; Chandra et al. 2007; Sinha and Ajmera 2019; Stylios et al. 2021)	(Jackson 2009)	Confirmed (p<0.05)
H ₂	There is subsume association between technical factors and adoption of multi-model biometrics adoption, ($\beta = 0.097$, S. Er = 0.056, Cri. ratio = 1.72)	(Cherrat et al. 2020; El-fishawy 2015; Jagadiswary and Saraswady 2016)	(Kalini 2017)	Confirmed (p<0.05)
H ₃	Cognizant Factors towards Biometrics have influenced on BA, ($\beta = 0.062$, S. Er. = 0.044, Cri. ratio = 1.40)	(Buriro et al. 2019; Carrión-ojeda et al. 2021; Nappi et al. 2018)	(Oloyede et al. 2016)	Confirmed (p<0.05)
H ₄	Perceptual Factors (Fingerprints, Iris, Face Recognition and Voice) have a strong association with MMB, ($\beta = 0.101$, S. Er. = 0.094, Cri. ratio = 1.06)	(Buckley and Nurse 2019; Chowdary and Hemanth 2019)	-	Confirmed (p<0.05)
H ₅	There is an association between Data Privacy Factors and adoptive operation of biometrics operations, ($\beta = 0.217$, S. Er. = 0.056, Cri. ratio = 3.90)	(Chong 2013; Choudhury et al. 2021; Dargan and Kumar 2020)	(Rahi et al. 2021)	Confirmed (p<0.05)

The rationalized SEM model exhibits theorized connotation among the latent variables. The assessment of standardized regression loads was applied to fetch an appreciation concerning the proposed disposition, as signified by [219],[220],[221]. In the above table, β , S. Er, Cri. ratios were positive, and eventually, null hypotheses can be rejected. The computed p-values of all five projected hypotheses were less than 0.05, viz. user acceptability and biometrics adoption (0.035), technical factors and adoption of multi-model biometrics adoption (0.045), Cognizant Factors towards Biometrics (0.040), Perceptual Factors (0.020), and Data Privacy (0.001) for the adoptive operation of biometrics. Most of the constructs were consistent with the previous studies. Hence it validates the proposed model good fit. Eventually, expressed hypotheses were supported and accepted.

The researchers used CFA to validate the hypotheses of the research model, as discussed in section 1. Besides, the researcher opted to perform an additional research tool ANN to rate the normalized value of the significant predictors based on the SEM analysis. Due to nonlinear correlations between the independent and outcome variables, the two-stage SEM-ANN methodology entails getting precise classification of performing predictors to biometrics adoption [185], [222]. These two methods are complementing to each other because the SEM is idyllic for hypothesis testing of linear relationships but cannot describe the relationship of nonlinearity, whereas the ANN can detect nonlinear relationships and is not suitable for hypothesis testing [223]

6.5.2 Artificial Neural Network

In the word of [228],[222], ANN is termed as an incredibly analogous scattered processor composed of computational units with a natural propensity for learning from experimental data and making it usable. Moreover, ANN seems to be a more advanced and typically stable method that offers a higher level of precision than traditional tools.

In ANN input Nodes or neurons collect the information that is referred to as synaptic weights. Further, it has already been proved by some prominent studies [225],[185], that the advantage of this methodology is that the neural network model can learn intricate linear and nonlinear relationships between predictors and the adoption decision. At the outset, SEM is applied to evaluate the overall research model and tested the significant hypothesized predictors, which were further used as inputs in the neural network model to assess the relative significance of each predictor item with nonlinear aspect. Input nodes, hidden layers, and output layer make up a neural network. Data are projected into the input layers, and the output information is generated in the output layers. Synaptic weights are assigned to each input and passed to the hidden layers. Using applied weights, a nonlinear activation process uses these values into an output value.

There are numerous other types of neural networks, but the researcher employs one of the most common and well-known instruments, i.e., the feed-forward back-propagation MLP

[250]. A typical neural network has multiple hierarchical levels, including one input, one or more hidden layers, and one output layer. The sigmoid function is used as an activation task in feed-forward networks. These activation function equations are:

Identity (Linear) $f(x) = x$

Hyperbolic Tangent $\tanh(x)f_x = \frac{2}{1+e^{-2x}}$ (6.2)

Sigmoid (Logistic step) $f_x = \frac{2}{1+e^{-x}}$ (6.3)

Where x is the input information. Any linear function can be represented with one hidden layer, but discontinuous functions can be described with two hidden layers. However, only one hidden layer is commonly employed in technology acceptance neural network models [227]. Each layer comprises neurons that connect with neurons in the next layer, and each link is represented by a synaptic weight that can be adjusted. Radial basis, recurrent networks, multi-layer perceptron, and Feed-forward neuronal functional networks are the four significant kinds of ANN. In this work the MLP model is applied to study Users' Awareness and Acceptability towards Adopting a Multimodal Biometrics Mechanism in Online Transactions. The synaptic weights of the relationships will be changed through an iterative training process while using the training samples to train the network[212].

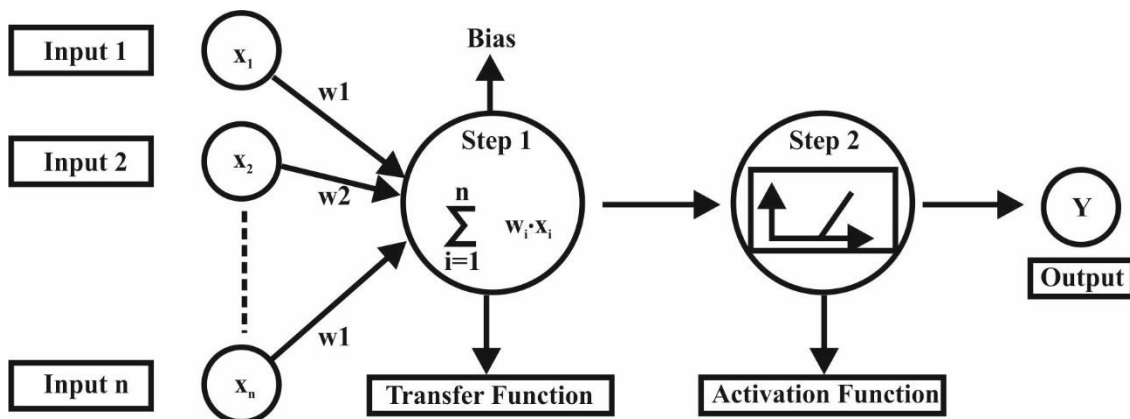


Fig. 6. 6 ANN model used for process

As shown in the Fig. 6.6, Input data is applied through multiple input nodes such as X_1 , X_2 , and X_3 and they are modulated with respect to their weights W_1 , W_2 , and W_3 . The

output Y of the neuron is a function of inputs and weights as per the Equation 6.4 & 6.5. The step 2, to induces nonlinearity in the process, nonlinear activation function (AF) is utilized. The aim of the (AF) is to initiate nonlinearity into the output of a neuron, which is vital since the real-world data are mostly nonlinear.

$$\mu_K = \sum_{i=1}^N W_{ki} X_j \quad (6.4)$$

and

$$Y_K = \varphi(\mu_K + D_K) \quad (6.5)$$

where μ_k is the linear combiner output because of input signal k . W_{kj} ($j = 1, 2, \dots, n$) is the respective weights of neuron k , where Φ is the activation function.

The flow used for in-built ANN algorithm in SPSS:

Input : # Data Preprocessing of the indicator's variables viz. User Acceptability: UAC,
Technological factors: TCF, Cognizant Factors: CGF,
Perceptual Factors: PRF,

Data Privacy Factors: DPF.

- Import the Libraries
- Load the Dataset
- Split Dataset into X and Y
- Encode Categorical Data
- Split the X and Y Dataset into the Training set and Test set
- Build ANN Model
- Initialize the ANN
- Add the input layer and the first hidden layer
- Train and compile the ANN
- Fit the ANN to the Training setting (hyper parameter tuning +loss function)

Output : # Predict the Test Set of Multimodal Biometrics (MMB)

Pre-processed dataset from SEM again used as input variable for ANN and transformed into numerical representation for adoption of Multimodal Biometrics analysis (see table 6.9).

6.5.3 Results of Neural Network Modeling

The neural network approach was examined using the widely used statistical program R-software. The statistically relevant determinants from the SEM analysis were inserted into this model at this point. From the findings of the structural equation, five constructs have been considered vital for further research. As a result, these items were rendered as input variables in the input layers. In this case, biometrics were chosen as the output layer's dependent variable. Furthermore, a cross-validation tool was used to overcome the over-fitting issue of the model[228],[229].

Table 6.9 RMSE value for training and testing data (N-530)

Sample size (Training)	SSE	RMSE	Sample size (Testing)	SSE	RMSE	RMSE	Total Sample
						(Training-Testing)	
466	214.87	0.683	64	25.113	0.652	0.03	530
468	219.101	0.688	62	23.2	0.638	0.05	530
467	229.522	0.705	63	31.686	0.739	0.034	530
474	237.42	0.711	56	13.711	0.519	0.193	530
467	236.439	0.715	63	25.445	0.662	0.053	530
471	222.779	0.691	59	24.461	0.673	0.018	530
475	216.567	0.679	55	18.539	0.609	0.07	530
462	215.448	0.687	68	39.887	0.796	0.109	530
473	221.31	0.688	57	16.193	0.558	0.13	530
465	219.682	0.691	65	29.767	0.704	0.013	530
Mean	223.717	0.694	Mean	24.248	0.084	0.057	-
Σ	8.324	0.012	σ	7.739	0.081	0.057	-

Note: SSE-Sum of error, RMSE-Root-mean-square of errors, N-sample size

Basically, when network memories the noise and become fit too closely to the training set, the model becomes “overfitted”. In such cases, network is unable to generalize with new

data. In the ANN model, Gandhmal et al. [230] proposed that hidden nodes should be in the range of 1–10. During the study process, twenty percent of data were used for experimentation, while eighty percent were used for training. Table 6.9 shows the RMSE values for both training and testing data points, as well as the mean and standard deviation. The outputs show that the mean RMSE values for the training and testing model are 0.694 and 0.057, respectively, whereas the σ of training data are 0.012 and 0.081 for testing data.

$$\text{SSE (Sum squared error)} = \sum_{i=1}^N (\text{Predicted} - \text{Actuals})^2 \quad (6.6)$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^N (\text{Predicted} - \text{Actuals})^2}{N}} \quad (6.7)$$

N is number of samples/inputs.

The RMSE values with comparatively small σ and indicate a higher level of accuracy in the statistical results [231]. We, thus, affirm that the model is a good fit. Using a similar approach as Anabel et al. [232] applied, the researcher computed the R^2 0.87%, and the result reveals that the ANN model predicts multimodal biometrics adoption with an accuracy of 87 percent. The results also show that the ANN mined very secure connections between the significant predictors and the output variables. Furthermore, in evaluating meaningful work, the sensitivity analytics were determined with the average importance of the predictor. The normalized relative significance of each forecast in the model was calculated by dividing each predictor's relative importance by the highest predictor.

Table 6.10 shows each predictor's Normalized and Sensitivity Assessment. The results of the ANN, however, outline that the average explanation of the Data Privacy construct was significant, i.e., DPF_3 (93%), DPF_2 (50%) and DPF_4 (34%). Perceptual construct- PRF_2 (49%) and PRF_3 (33%) was relatively most important predictor of BA, whereas in User Acceptability-UAC_2 (37%), UAC_3& UAC_5 (41%), Technological Factors- TCF_2 (35%), followed by Cognizant factors- CFG_1 (33%) towards biometrics adoption (see table 6.10). Thus, the model confirms the best fit from both CFA-ANN approaches.

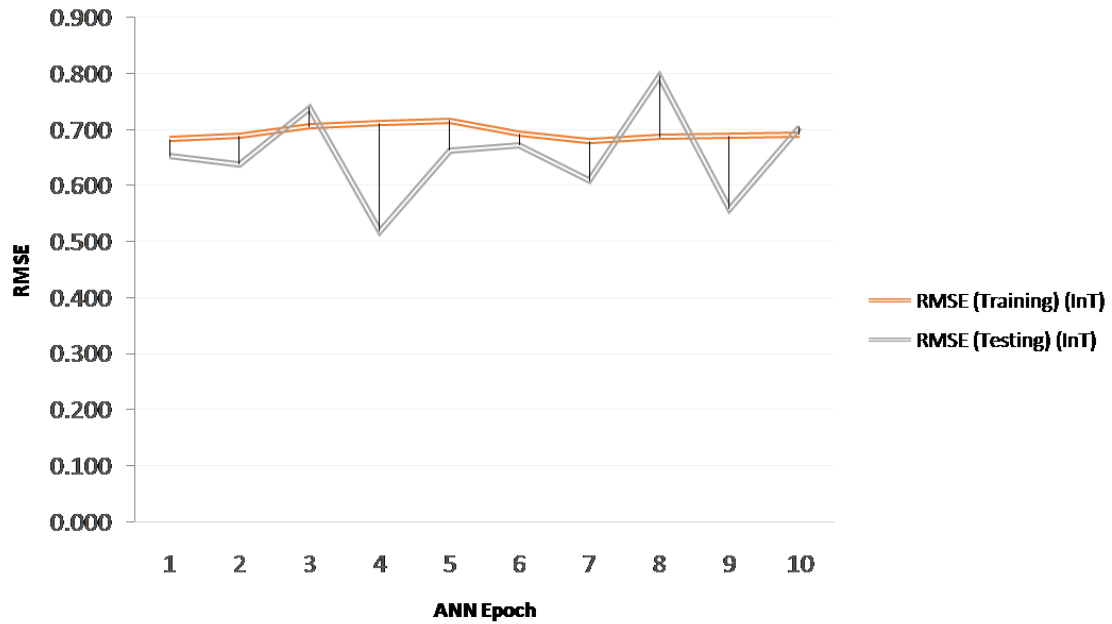


Fig. 6.7 RMSE statistics of training and testing

6.6 Theoretical Implications

Regarding the quantitative outcomes, this study's results contribute to advancing the literature in biometrics and its awareness among various segments of users. Furthermore, this study contributes significantly to our understanding of biometrics adoption at the virtual level in this study, the integration of certain constructs, i.e., Cognizant factor and Perceptual factors, has rendered a new theoretical contribution to the issues influencing biometrics adoption. Although these factors' application was ubiquitous in marketing research, there is a lack of studies that examined its comprehensive effects with other variables. Thus, this is perhaps the first research encompassing five constructs and 21

Table 6.10 Normalized and sensitivity analysis

Neural Network	NI (i)	NI (ii)	NI (iii)	NI (iv)	NI (v)	NI (vi)	NI (vii)	NI (viii)	NI (ix)	NI (x)	Average
UAC_1	30%	21%	28%	45%	31%	25%	44%	22%	18%	77%	34%
UAC_2	19%	73%	37%	31%	43%	20%	62%	8%	47%	26%	37%
UAC_3	52%	88%	31%	47%	37%	34%	26%	14%	42%	43%	41%
UAC_4	19%	50%	49%	24%	58%	18%	18%	22%	14%	74%	34%
UAC_5	16%	50%	48%	34%	59%	28%	28%	23%	24%	75%	41%
TCF_1	23%	38%	37%	42%	28%	22%	19%	34%	6%	40%	29%
TCF_2	58%	63%	18%	54%	22%	13%	50%	9%	39%	24%	35%
TCF_3	31%	42%	19%	13%	21%	51%	18%	24%	15%	41%	27%
TCF_4	17%	43%	14%	25%	25%	28%	31%	13%	33%	21%	25%
CFG_1	39%	50%	16%	28%	57%	23%	34%	24%	9%	55%	33%
CFG_2	16%	22%	6%	30%	57%	34%	21%	34%	19%	67%	31%
CFG_3	15%	62%	31%	16%	22%	19%	12%	24%	10%	70%	28%
CFG_4	21%	39%	20%	13%	44%	40%	26%	35%	17%	24%	28%
PRF_1	10%	31%	20%	34%	66%	18%	24%	24%	35%	33%	30%
PRF_2	100%	51%	58%	28%	26%	38%	99%	13%	23%	54%	49%
PRF_3	28%	45%	16%	24%	25%	10%	27%	50%	21%	85%	33%
PRF_4	65%	27%	21%	35%	22%	23%	73%	32%	10%	10%	32%
DPF_1	27%	65%	20%	51%	27%	16%	21%	9%	11%	33%	28%
DPF_2	55%	13%	46%	31%	42%	43%	84%	79%	11%	100%	50%
DPF_3	73%	100%	100%	100%	75%	100%	100%	100%	91%	91%	93%
DPF_4	41%	57%	26%	37%	21%	29%	55%	23%	45%	7%	34%

variables to explicate biometrics adoption in the present chaotic scenario. Contrary to current linked studies using linear models, we adopted a two-stage SEM-ANN approach consisting of a linear and nonlinear ANN model. It is a new technique since a decrease in one predictor can be neutralized by increasing another predictor in a linear compensatory model [212]. We have effectively resolved the weakness of linear models and have therefore made an innovative theoretical contribution to established literature using a quasi-ANN model.

In summary, the study aims to provide a systematic, in-depth, and consistent understanding of the direct and indirect effect of biometrics on online transactions, which may be helpful

for researchers, administrators, and academicians while developing frameworks to deal with complex technological changes. Existing research has predominantly focused on the drivers of MMB adoption at the virtual platform; however, this study takes it a step further by developing and testing hypotheses about continent-level factors that affect the use and adoption of biometrics techniques for online operation.

6.7 Limitations and Future Research

Even though the findings of this study have significant consequences for researchers and practitioners, it has the following limitations: First, the representative sample for this empirical research is very small. By gathering more data, future research would be able to provide more credible scientific findings. Second, SEM was used in this research to analyze the research model. Even though SEM is a standard statistical technique in information management research, it does not hierarchically evaluate independent variables. Future research should consider computational methods such as AI-ML, which can isolate variables at the personal and organizational levels. Third, this work conducted an empirical investigation using a survey approach. If an in-depth interview or qualitative methods and measurement can be used to get more feasible and meaningful findings in terms of theoretical contributions. Fourth, the research focused exclusively on educational organizations located on a single continent. It is recommended that future cross-country research on MMB can be done to generalize the findings. Finally, the study examined users' knowledge and willingness to adopt the MMB in a limited time frame.

6.8 Summary

The numerous applications of biometrics are gaining traction in business and society. The study focuses on analyzing user awareness of MMB and its acceptability for online transactions in the current dynamic world under pandemic. The study was conducted on the five underlying perspectives named User Acceptability, Cognizant Factors towards Biometrics, Technological factors, Perceptual Factors (Fingerprints, Iris, Face Recognition and Voice) and Data Privacy Factors. The study illustrates the relationship

between previously used constructs and evaluates a multimodal biometrics adoption. A questionnaire was prepared and circulated to the 530 biometrics users; on that basis, the corresponding answers were obtained for analysis. The collected replies were just from the professionals/experts who had performed online tasks of MMB. SEM is first employed to gauge the wholesome research model and tested the prominent hypothesized predictors, which are then used as inputs in the neural network to evaluate the relative significance of each predictor variable. SEM-ANN analysis used to identify the inclusive link among variables, including linear-nonlinear and non-compensatory correlations (Dadhich et al. 2021; Khayer et al. 2020)[250,228]. Additionally, the researcher extended the prior study by incorporating two new constructs, i.e., Cognizant, and Perceptual factors, with already tested constructed elements like user acceptability, technological considerations, and data privacy factors.

Using Back Propagation Algorithm (BPA) of ANN algorithms, we found a significant effect of DPF_3 (93%), DPF_2 (50%) and DPF_4 (34%) on the adoption of MMB. In Perceptual construct, PRF_2 (49%), and PRF_3 (33%) was relatively most important predictor whereas, in User Acceptability, UAC_2 (37%), UAC_5 & UAC_3 (41%) was vital to be considered. Only one item, TCF_2 (35%), from Technological Factors, followed by Cognizant factors, i.e., CFG_1 (33%), confirmed the best fit model to adopt MMB. The study is a unique and comprehensive attempt when compared to past studies, as it considered cognizant and perceptual factors, thereby extending the analytical outlook of MMB literature. The study also explored several new and valuable theoretical implications for adopting multimodal instruments of biometrics adoption.

Chapter 7

Conclusions and Future Scope

7.1 Preamble

The overall conclusion drawn from the thesis covering the detailed simulation study of multimodal biometrics, its implementation challenges, and Convolution Neural Network (CNN) based experiments and user perception analysis about its use in live cases are explored. The reported work highlights the strength and challenges of multimodal biometric implementations. In a later segment, detailed study of user awareness and perception of multimodal biometric applications in online transactions is presented. Furthermore, new research directions which have emerged as a result of thesis work are presented in future recommendations.

7.2 Motivation

Thesis represents the research work undertaken by us. Following are the issues and challenges which motivated the work presented in this thesis.

- Unimodal Biometric system's vulnerability against spoof attacks, imposters and low performance has motivated us to explore Multimodal Biometric fusion-based recognition.
- The potential of feature level fusion for multimodal biometric systems.
- Optimization of features to overcome the dimensionality and compatibility issues in feature-level fusion.
- Need for efficient CNN-based multimodal biometric system for recognition.
- Need for a detailed study of User awareness and perceptions about the adoption of Multimodal Biometric systems in online transactions.

The research work presented is primarily concerned with the various issues and the problems encountered as well as appropriate solutions to address the issues. This thesis reports a step toward the development of the feature level fusion based multimodal biometric systems.

The work started with unimodal and multimodal recognition process with standard feature extraction and classifications. Feature level fusion based multilevel multimodal process has been found to outperform other unimodal work with similar traits. It was validated against standard databases.

Features are the key differentiator of the authentication process and good quality information extraction from modalities ensures high accuracy. Fusing features from different modalities generates a fused feature vector which is used for matching. It has been observed that the level of matching creates a significant difference in performance in different scenarios. Due to the availability of intact raw features, feature-level fusion has been opted as a key fusion scheme in the entire work and the results also validated our decision.

In areas where there were challenges of the curse of dimensionality, optimization methods have been applied to data before making decisions. Optimization reduces complexities and so computation costs. This thesis presents a Modified Grey Wolf Optimization-based multimodal biometric system that outperforms other reported work in similar conditions. Optimization of features before fusion has solved the issue of large-size feature vectors.

The work also presents CNN (VGG16) based multimodal biometric fusion process. Convolution neural network-based architecture like Alexnet, VGG, and Resnet have outperformed the traditional approach of classification and predictions. CNN-based multimodal biometric models are more suitable for implementation on portable platforms like lightweight portable devices. The presented MVGG16 architecture has shown better time efficiency.

In the final part, the thesis presents user perception of multimodal biometric systems and their use in real-time applications. User awareness depends on multiple factors such as cognizant factors, technological factors, perceptual factors, and data privacy factors.

7.3 Conclusions

Following are the thesis work's conclusions

- Feature-level fusion is the core approach of the entire presented work and different methods are used for unimodal and multimodal authentication with feature-level fusion. To enhance the performance joint fusion scheme and multi-level fusion were also adopted. In some experiments, score level, feature level, and combined fusion were performed together. Different classifiers like Random Forest, KNN, SVM, and MSVM were compared and tested on Palm print, Fingerprint, Ear, Iris, and Face modalities.
- Modified Grey Wolf Optimization method with feature level fusion has helped to solve the issue of dimensionality. Palm print, Ear, and Fingerprint combinations were selected for the authentication. Gabor and HMSB operators produced texture features, whereas shape feature was extracted from ear modality. For selecting the optimal feature, the OGWO+LQ algorithm was used. For recognition, Multi kernel Support Vector Machine (MK SVM) algorithm was used. OGWO+ LQ features were classified using MSVM and results have outperformed other contemporary approaches. This process obtained Sensitivity, Specificity, and Accuracy of 0.9166, 0.92, and 0.97 respectively.
- LCNN-based continuous user authentication for E-proctoring using face, and fingerprint keystroke inputs have shown good results. A modified wolf optimization process was used for controlling feature vector size issues. The optimum feature fusion method ensured a good quality of features. For final stage recognition, LCNN and Salp Swarm Optimization (SSO) produced 99.13 % Accuracy and 96.46 % F-1 Score at a False Positive Rate of 4.61%. The proposed approach was tested on different datasets.
- Thesis presents a techno commercial and scientific analysis of user awareness for multimodal biometric systems in the last phase. The study was conducted on five underlying perspectives name User Acceptability, Cognizant Factors towards Biometrics, Technological factors, Perceptual Factors (Fingerprints, Iris, Face Recognition, and Voice), and Data Privacy Factors. The study illustrates the relationship between previously used constructs and evaluates multimodal biometrics adoption. In this study, the integration of certain constructs, i.e., Cognizant factors and Perceptual

factors, has rendered a new theoretical contribution to the issues influencing biometrics adoption. Although these factors' application was ubiquitous in marketing research, there is a lack of studies that examined its comprehensive effects with other variables. SEM-ANN analysis has been used to identify the inclusive link among variables, including linear-nonlinear and non-compensatory correlations. Technological and Cognizant factors have highlighted that multimodal biometric process are safer and more future oriented.

7.4 Future Recommendations

The thesis work has opened up new research avenues. Some of the intriguing aspects that can be addressed as a result of the work done are as follows.

- In the thesis, optimization of feature level fusion process is presented. Its real time implementation can be explored.
- Mobile devices have hardware limitations due to weight, storage capacity, and power consumption constraints. Multimodal biometric authentication for mobile devices with a limited database would be a challenging task. This can be explored and implemented in the future.
- The SEM ANN approach for the study of user perception was executed in an SPSS environment. Its real-time implementation could be performed on high-end technical platforms like MATLAB to get more insights into user behavior.
- Designing a dedicated Convolution Neural Network (CNN) from the scratch for multimodal biometric authentication would be challenging work to explore.

References

1. Gokulkumari, G., “Analytical outlook on customer awareness towards biometrics mechanism of unimodal and multimodal in online transactions,” *Multimedia Tools and Applications*, vol. 79, issue 41-42, pp. 31691–31714, 2020.
2. Purohit, H., & Ajmera, P.K., “Multimodal biometric systems, a brief study,” *International Journal of Innovative Technology and Exploring Engineering*, vol.8, issue 7, pp. 108-111, 2019.
3. Purohit, H., & Ajmera, P.K., “Fusion of Palm Print and Palm Geometry,” *Advances in Intelligent Systems and Computing*, vol. 870. pp. 558-563, 2019.
4. Nigam, A., Tiwari, K., & Gupta, P., “Multiple Texture Information Fusion for Finger-Knuckle-Print Authentication System,” *Neurocomputing*, vol. 4, issue 3, pp. 188-185, 2015.
5. Maltoni, D., Maio, D., Jain, A.K., & Prabhakar, S., “Handbook of Fingerprint Recognition,” Springer professional computing, vol.2, 2003.
6. Frischholz, R. W., & Dieckmann U., “BioID: A multimodal biometric identification system,” *Computer Systems*, vol.33, issue 2, pp. 64-68, 2000.
7. Rattani A, Kisku DR, Bicego M, & TistarelliM. “Feature level fusion of face and fingerprint biometrics,” *BTAS*. vol. 27, pp. 1-6, 2007
8. Jain, A., Nandkumar, K., & Ross, A., “Score Normalization in Multimodal Biometric Systems” *Pattern Recognition*, vol. 38, issue 12, pp. 2270–2285, 2005.
9. Dadhich, Manish, “An Analysis of Factors Affecting on Online Shopping Behavior of Customers,” *ZENITH International Journal of Business Economics & Management Research*, vol.7, issue 1, pp.20–30, 2017.
10. Purohit, H., Ajmera, P.K., “Optimal Feature Level Fusion for Secured Human Authentication in Multimodal Biometric System,” *Machine Vision and Applications*, vol. 32, issue 24, pp.12-34, 2021.
11. Dargan, Shaveta, and Kumar, M., “A Comprehensive Survey on the Biometric Recognition Systems Based on Physiological and Behavioral Modalities,” *Expert Systems With Applications*, vol. 143, pp.113-114, 2020.
12. Bours, P. “Continuous keystroke dynamics: A different perspective towards biometric evaluation,” *Information Security Technical Report*, vol.17, issue-2, pp.36-43, 2012.
13. Peng, Jialiang, Ahmed A. Abd El-Latif, Qiong Li, and Xiamu Niu "Multimodal biometric authentication based on score level fusion of finger biometrics," *Optik-International Journal for Light and Electron Optics*, vol.125, issue 23, pp.6891-6897, 2014.
14. Saevanee, H., Clarke N., Furnell S., and Biscione, V., “Continuous user authentication using multi-modal biometrics,” *Computers & Security*, vol. 53, pp.234-246, 2015.
15. Mondal, S. and Bours, P., “A computational approach to the continuous authentication biometric system,” *Information Sciences*, vol.304, pp.28-53, 2015

16. Gao, Xizhan, Quansen Sun, and Haitao Xu., "Multiple rank supervised canonical correlation analysis for feature extraction, fusion and recognition," *Expert Systems with Applications*, vol. 84, pp. 171-185, 2017.
17. Yang, Jinfeng, and Xu Zhang "Feature-level fusion of fingerprint and finger-vein for personal identification," *Pattern Recognition Letters*, vol. 33, issue 5, pp.623-628, 2012.
18. Veluchamy, S., and Karlmarx, L.R., "System for multimodal biometric recognition based on finger knuckle and finger vein using feature-level fusion and k-support vector machine classifier," *IET Biometrics*, vol.6, issue 3, pp. 232-242, 2016.
19. Nagar, Abhishek, Nandakumar, K., and Jain, A.K., "Multibiometric cryptosystems based on feature-level fusion," *IEEE transactions on information forensics and security*, vol.7, issue 1, pp. 255-268, 2012.
20. Yang, Wencheng, Song Wang, Jiankun Hu, Guanglou Zheng, and Craig Valli "A Fingerprint and Finger-vein Based Cancelable Multi-Biometric System." *Pattern Recognition*, vol.78, issue 12, pp. 242-251, 2018.
21. Slanzi, Gino, Gaspar Pizarro, and Juan D. Velásquez "Biometric information fusion for web user navigation and preferences analysis: An overview," *Information Fusion*, vol.38, pp.12-21, 2017.
22. Miao, Di, Man Zhang, Zhenan Sun, Tieniu Tan, and Zhaofeng He "Bin-based classifier fusion of iris and face biometrics," *Neuro computing*, vol.224, pp. 105-118, 2017.
23. Duan, M., Li, K., Liao, X., Li, K., Tian, Q. "Features-enhanced multi-attribute estimation with convolutional tensor correlation fusion network," *ACM Trans. Multimed. Comput. Commun. Appl. (TOMM)*, vol.15, issue3, pp.1–23, 2019.
24. Duan, M., Li, K., Li, K. "An ensemble cnn2elm for age estimation," *IEEE Trans. Inf. Forens. Secur.*, Vol. 13, issue 3, pp. 758–772, 2017.
25. Duan, M., Li, K., Yang, C., Li, K. "A hybrid deep learning CNN–ELM for age and gender classification," *Neurocomputing*, vol. 275, pp.448–461, 2018.
26. Peng, G., Zhou, G., Nguyen, D.T., Qi, X., Yang, Q. and Wang, S. "Continuous authentication with touch behavioral biometrics and voice on wearable glasses," *IEEE transactions on human-machine systems*, vol. 47, issue 3, pp.404-416, 2016.
27. Ali, M.L., Monaco, J.V., Tappert, C.C. and Qiu, M. "Keystroke biometric systems for user authentication," *Journal of Signal Processing Systems*, vol.86, issue 2, pp.175-190, 2017.
28. Ullah, A., Xiao, H. and Barker, T., "A study into the usability and security implications of text and image-based challenge questions in the context of online examination," *Education and Information Technologies*, vol. 24, issue 1, pp.13-39, 2019.
29. Ullah, A., Xiao, H. and Barker, T., "A dynamic profile questions approach to mitigate impersonation in online examinations," *Journal of Grid Computing*, vol. 17, issue 2, pp.209-223, 2019.
30. Prakash, A., "Continuous user authentication-based score level fusion with hybrid optimization," *Cluster Computing*, vol. 22, issue 5, pp.12959-12969, 2019.

31. Yang, Y., Guo, B., Wang, Z., Li, M., Yu, Z. and Zhou, X., "BehaveSense: Continuous authentication for security-sensitive mobile apps using behavioral biometrics," *Ad Hoc Networks*, vol. 84, pp. 9-18, 2019.
32. Nunes, I.D.O., Eldefrawy, K., and Lepoint, T., "SNUSE: A secure computation approach for large-scale user re-enrollment in biometric authentication systems," *Future Generation Computer Systems*, vol. 98, pp. 259-273, 2019.
33. Chatterjee, K., "Continuous User Authentication System: A Risk Analysis Based Approach," *Wireless Personal Communications*, vol. 108, issue 1, pp. 281-295, 2019.
34. Wang, F., Li, Z. and Han, J., "Continuous user authentication by contactless wireless sensing," *IEEE Internet of Things Journal*, vol.6, issue 5, pp. 8323-8331, 2019.
35. Kiyani, A.T., Lasebae, A., Ali, K., Rehman, M.U., and Haq, B., "Continuous user authentication featuring keystroke dynamics based on robust recurrent confidence model and ensemble learning approach," *Sensors*. vol. 21, pp. 22-42, 2020.
36. Mario P. C., Yu Guan, Aad van Moorsel, "Mobile Based Continuous Authentication Using Deep Features," *EMDL'18: Proceedings of the 2nd International Workshop on Embedded and Mobile Deep Learning*, pp 19–24 , 2018.
37. Shende, P., Dandawate, Y. "Convolutional neural network-based feature extraction using multimodal for high security application," *Evolutionary intelligence*, vol. 14, pp. 1023–1033, 2021.
38. Suleiman, S., Dabouei, A., Kazemi, H., Dawson, J., and Nasrabadi, N.M., "Multi-Level Feature Abstraction from Convolutional Neural Networks for Multimodal Biometric Identification," *24th International Conference on Pattern Recognition (ICPR)*, pp. 3469-3476, 2018.
39. Jialiang Peng, Ahmed A. Abd El-Latif, Qiong Li, Xiamu Niu, "Multimodal biometric authentication based on score level fusion of finger biometrics," *Optik*, vol.125, issue 23, pp. 6891-6897, 2014.
40. Earnest E. Hansley, Maurício Pamplona Segundo, Sarkar, S., "Employing fusion of learned and handcrafted features for unconstrained ear recognition," *IET Biometrics*, vol.7 iss.3, pp.215-223, 2018.
41. Sinha, H., Ajmera, P.K., "Upgrading Security and protection in ear biometrics," *IET Biometrics*, vol. 8, issue 4, pp. 259 – 266, 2019.
42. Sinha, H., Awasthi, V., Ajmera P. K., "Audio classification using braided convolutional neural networks," *IET Signal Processing*, vol. 14, issue 7, pp. 448-454, 2020.
43. Al-Waisy, R. Qahwaji, S. Ipson and S. Al-Fahdawi, "A multimodal biometric system for personal identification based on deep learning approaches," *Seventh International Conference on Emerging Security Technologies (EST)*, pp. 163-168, 2017.
44. Kim Wan, Song JM, Park KR. "Multimodal Biometric Recognition Based on Convolutional Neural Network by the Fusion of Finger-Vein and Finger Shape

- Using Near-Infrared (NIR) Camera Sensor,” *Sensors (Basel)*; vol.18, issue 7, pp. 22-96, 2018.
45. Sun Q., Zhang J., Yang A., Zhang Q. “Palmpoint Recognition with Deep Convolutional Features,” *Advances in Image and Graphics Technologies. IGTA Communications in Computer and Information Science*, vol. 757, 2018.
 46. Omkar M. Parkhi, Andrea Vedaldi and Andrew Zisserman. “Deep Face Recognition,” *Proceedings of the British Machine Vision Conference (BMVC)*, pp. 41.1-41.12, 2015.
 47. Singh, M. Singh, R., Ross, A., “A Comprehensive Overview of Biometric Fusion,” *Information Fusion*, vol. 52, pp. 187-205, 2019
 48. Su Tang, Shan Zhou, Wenxiong Kang, Qiuxia Wu, Feiqi Deng, “Finger vein verification using a siamese CNN,” *IET Biometrics*, vol. 8, issue 5, pp. 306-315, 2019.
 49. Gunasekaran, K., Raja, J., Pitchai, R., “ Deep multimodal biometric recognition using contourlet derivative weighted rank fusion with human face, fingerprint and iris images,” *Journal for Control, Measurement, Electronics, Computing and Communications: Computational Intelligence and Capsule Networks*, vol. 60, issue 3, pp. 253-265, 2019.
 50. Umit Kacar, Murvet Kirci, “ScoreNet: Deep Cascade score level fusion for unconstrained ear recognition,” *Cluster Computing*, vol. 8, issue 2, pp. 109 – 120, 2019.
 51. Dadhich, M. “An Analysis of Factors Affecting on Online Shopping Behavior of Customers,” *ZENITH International Journal of Business Economics & Management Research*, vol.7, issue 1, pp. 20–30, 2017.
 52. Birda, R.K., & Dadhich, M., “Study of ICT and E-Governance Facilities in Tribal District of Rajasthan,” *ZENITH International Journal Of Multidisciplinary Research*, vol. 9, issue 7, pp.39–49, 2019.
 53. Porwik, Piotr, and Rafal Doroz., “Adaptation of the Idea of Concept Drift to Some Behavioral Biometrics : Preliminary Studies,” *Engineering Applications of Artificial Intelligence*, vol. 99, pp. 2021.
 54. Anil, Anu P., and Satish K.P, “Investigating the Relationship Between TQM Practices and Firm’s Performance: A Conceptual Framework for Indian Organizations,” *Procedia Technology*, vol. 24, pp. 554–61, 2016.
 55. Rahi, Aysha, Abrar Shakil, Nayeema Ferdausy, and Moshiur Rahman. “Effect of Eco-Physiological Factors on Biometric Traits of Green Mussel *Perna Viridis* Cultured in the South-East Coast of the Bay of Bengal , Bangladesh,” *Aquaculture Reports*, vol.19, pp.100-562, 2021.
 56. Sarier, Neyire Deniz, “Comments on Biometric-Based Non-Transferable Credentials and Their Application in Blockchain-Based Identity Management,” *Computers & Security*, vol. 105, pp. 102-243, 2021.
 57. Sinha, Harsh, Awasthi, V., and Pawan K. Ajmera, “Audio Classification Using Braided Convolutional Neural Networks,” *IET Research Journals*, vol.14, issue 7, pp. 448-454, 2020.
 58. El-fishawy, Nawal, “Multi-Biometric Systems : A State of the Art Survey and

- Research Directions,” *International Journal of Advanced Computer Science and Applications*, vol.6, issue 6, pp.128–38, 2015.
59. Guan, Yu, Yunlian Sun, Chang-tsun Li, and Massimo Tistarelli, “Human Gait Identification from Extremely Low-Quality Videos: An Enhanced Classifier Ensemble Method,” *IET Biometrics*, vol. 3, pp. 84–93, 2014.
 60. Heracleous, Loizos, and Templeton College. “Biometrics: The next Frontier in Service Excellence, Productivity and Security in the Service Sector,” *Managing Service Quality*, vol. 16, issue 1, pp.12–22, 2016.
 61. Stylios, Ioannis, Spyros Kokolakis, Olga Thanou, and Sotirios Chatzis, “Behavioral Biometrics & Continuous User Authentication on Mobile Devices: A Survey,” *Information Fusion*, vol. 66, pp. 76–99, 2021.
 62. Jagadiswary, D., and Saraswady,D., “Biometric Authentication Using Fused Multimodal Biometric,” *Procedia - Procedia Computer Science* vol. 85, pp. 109–16, 2016.
 63. Rivest, R.; Shamir, A., Adleman, L., “A Method for Obtaining Digital Signatures and Public-Key Cryptosystems,” *Communications of the ACM*, vol. 21, issue 2, pp. 120-126, 1977.
 64. Choudhury, S., Kumar, A., and Laskar, S.H, “Adaptive Management of Multimodal Biometrics --- A Deep Learning and Metaheuristic Approach,” *Applied Soft Computing Journal*, vol. 1, pp. 107-344, 2021.
 65. Srinivasu, Parvathaneni Naga, Jalluri Gnana Sivasai, Muhammad Fazal Ijaz, Bhoi, A.K.,Wonjoon Kim, and Kang., J.J., “Classification of Skin Disease Using Deep Learning Neural Networks with MobileNet V2 and LSTM,” *Sensors, MDPI*, vol.21, pp. 1–27, 2021
 66. Panigrahi, Ranjit, Samarjeet Borah, Akash Kumar Bhoi, Muhammad Fazal Ijaz, Moumita Pramanik, Yogesh Kumar, and Rutvij H. Jhaveri, “A Consolidated Decision Tree-Based Intrusion Detection System for Binary and Multiclass Imbalanced Datasets,” *Mathematics*, vol. 9, issue 7, pp. 2–35, 2021.
 67. Ajeenkya D. Patil, Y., Nagra, G., Gopal, R., “International Journal of Management,” *International Journal of Ma.*, vol.4, issue 2, pp. 1–5, 2013.
 68. Liang, Yaling, and Chang-Tsun Li., “Gait Recognition Based on the Golden Ratio,” *EURASIP Journal on Image and Video Processing*, vol. 2, issue 9, pp. 1-8, 2016.
 69. Nader, J., Alsadoon, A., Prasad, P.W.C., Singh, A.K., and Elchouemi. A., “Designing Touch-Based Hybrid Authentication Method for Smartphones,” *Procedia Computer Science*, vol.70, pp. 198–204, 2015.
 70. Oloyede, Muhtahir O., Student Member, and Hancke, G.P., “Unimodal and Multimodal Biometric Sensing Systems: A Review,” *IEEE access*, vol.4, pp.7532–55, 2016.
 71. Sinha, Harsh, and Ajmera. P.K., “Upgrading Security and Protection in Ear Biometrics,” *IET Biometrics*, vol. 8, issue 4, pp. 259–66, 2019.
 72. Sireesha, V., and Sandhyarani. K., “Overview of Fusion Techniques in Multimodal,” *International Journal of Engineering Research & Technology, NCDMA*, pp. 3–8, 2014.
 73. Stylios, Ioannis, Spyros Kokolakis, Olga Thanou, and Sotirios Chatzis.

- “Behavioral Biometrics & Continuous User Authentication on Mobile Devices : A Survey,” *Information Fusion*, vol. 66, pp. 76–99, 2021.
74. Teh, Pin Shen, Ning Zhang, and Ke Chen, “TDAS : A Touch Dynamics Based Multi-Factor Authentication Solution for Mobile Devices,” *International Journal of Pervasive Computing and Communications*, vol.12, issue 1, pp.127–53, 2016.
 75. Venkatraman, Sitalakshmi.“Biometrics in Banking Security : A Case Study.” *Information Management & Computer Security*, vol. 16, issue 4, pp. 415–30, 2008.
 76. Venkatesh, Viswanath; Morris, Michael G.; Davis, Gordon B.; Davis, Fred D. “User Acceptance of Information Technology: Toward a Unified View,” *MIS Quarterly*, vol. 27, issue 3, pp. 425–478, 2003.
 77. Abomhara, Mohamed, Sule Yildirim Yayilgan, and Livinus Obiora Nweke, “A Comparison of Primary Stakeholders’ Views on the Deployment of Biometric Technologies in Border Management : Case Study of SMart Mobility at the European Land Borders,” *Technology in Society*, vol. 64, pp. 24-36, 2021.
 78. Joshi, Mahesh, Bodhisatwa Mazumdar, and Somnath Dey, “A Comprehensive Security Analysis of Match-in-Database Fingerprint Biometric System R,” *Pattern Recognition Letters*, vol. 138, pp. 247–66, 2020.
 79. Dwivedi, Yogesh K., Walton, P., and Michael D. Williams, “Artificial Intelligence (AI): Multidisciplinary Perspectives on Emerging Challenges , Opportunities , and Agenda for Research, Practice and Policy,” *International Journal of Information Management*, vol. 57, pp. 34-79, 2021.
 80. Alsadoon, Abeer, Pham, L., and Elchouemi. A., “An Enhanced Model of Biometric Authentication,” *International Conference on Advances In Electrical, Electronic And Systems Engineering*, pp. 14–16, 2016.
 81. Cherrat, El, Alaoui, R., and Bouzahir, H., “Convolutional Neural Networks Approach for Multimodal Biometric Identification System Using the Fusion of Fingerprint, Finger-Vein and Face Images,” *Peer Journal Computer Science*, vol. 6, issue 248, pp. 1–15, 2020.
 82. Carrión O., Dustin, Rigoberto F., and Israel P., “Analysis of Factors That Influence the Performance of Biometric Systems Based on EEG Signals,” *Expert Systems with Applications*, vol. 165, pp. 45-52, 2020.
 83. Oleish, Linda, and Awad Allah, “The Effect of Integration of Resource Consumption Accounting and Total Quality on Strategic Cost Management in Sudanese Food Industrial Companies,” *International Conference on Advances in Electrical, Electronic and Systems Engineering*, pp. 138-143, 2018.
 84. Dadhich, M., Pahwa, M.S., and Rao. S.S., “Factor Influencing to Users Acceptance of Digital Payment System,” *International Journal of Computer Sciences and Engineering*, vol. 06, issue 09, pp. 46–50, 2018.
 85. Langevin, J., Reyna, J.L., Ebrahim igharehbaghi, S., Sandberg, N., Fennell, P., “Developing a Common Approach for Classifying Building Stock Energy Models,” *Renewable and Sustainable Energy Reviews*, vol. 133, 2020.
 86. Rogers, Everett, "Continuous User Authentication on Mobile Devices: Recent progress and remaining challenges," in *IEEE Signal Processing Magazine*, vol. 33, issue 4, pp. 49-61, 2016.

87. Tornatzky, L.G., Fleischer, M. "The Processes of Technological Innovation, TOE" *Journal of Technological Transfer*, vol.16, pp.45–46, 1990.
88. Davis, F. D., Bagozzi, R. P., Warshaw, P. R., "User Acceptance of Computer Technology: A Comparison of Two Theoretical Models," *Management Science*, vol. 35, issue 8, pp. 982–1003, 1989.
89. Ajzen, I., "Theory of planned behavior. Organizational behavior and human decision processes," vol. 50, pp. 179-211, 1991
90. Bandura, A. "Social Cognitive Theory: An Agentic Perspective," *Annual Review Psychology*, vol. 52, pp. 1-26., 2011.
91. Faheem, Aslam, Mughal Khurram S., Ali Ashiq, and Tariq., M.Y, "Forecasting Islamic Securities Index Using Artificial Neural Networks: Performance Evaluation of Technical Indicators," *Journal of Economic and Administrative Sciences*, vol. 37, issue 2, pp. 253-271, 2020.
92. Dadhich, M., Pahwa, M.S., Jain, V., Doshi., R., "Predictive Models for Stock Market Index Using Stochastic Time Series ARIMA Modeling in Emerging Economy," *Advances in Mechanical Engineering*, pp. 281–90, 2021.
93. Kumar, A., and Passi, A., "Comparison and combination of iris matchers for reliable personal authentication," *Pattern Recognition.*, vol. 43, no. 3, pp. 1016–1026, 2010.
94. CASIAIris Ageing database.
<http://biometrics.idealtest.org/dbDetailForUser.do?id=14>
95. Multimedia University Iris Database (MMU) - V1 and V2
<https://mmuexpert.mmu.edu.my/ccteo> Accessed: 1-10-2020.
96. CUHK Iris Image Dataset, http://www.mae.cuhk.edu.hk/~cvl/main_database.htm
Accessed: 20-5-2020.
97. University of Bath: University of Bath Iris Image Database,
<http://www.bath.ac.uk/eleceng/research/sipg/irisweb/index.html>.
98. Kumar, A and Chenye Wu, "Automated human identification using ear imaging," *Pattern Recognition*, vol. 45, issue 3, pp. 956-968, 2012.
99. Phillips, P.J, Moon, H., Rizvi, S.A, and Rauss, P.J, "The FERET evaluation methodology for face recognition algorithm," *IEEE Trans. on PAMI*, vol. 22, no. 10, pp.1090-1104, 2000.
100. Belhumeur, P.N, Hespanha, J.P, and Kriegman, D.J, "Eigenfaces vs. Fisher faces: Recognition using class specific linear projection," *IEEE Trans. on PAMI*, vol. 19, no. 7, pp. 711-720, 1997
101. Samaria, F.S., and Harter, A.C, "Parameterization of a stochastic model for human face identification," *Proc. of 2nd IEEE workshop on Applications of Computer Vision*, pp. 138-142, 1994
102. Essex Face Database. [Online]: <http://cswww.essex.ac.uk/mv/allfaces/index.html>.
103. CASIA-Fingerprint Version 5.0 of the CASIA Fingerprint Image Database,
<http://biometrics.idealtest.org/findTotalDbByMode.do?mode=Fingerprint>
104. FVC2000, "<http://bias.csr.unibo.it/fvc2000/>," 2010.
105. FVC2002, "<http://bias.csr.unibo.it/fvc2002/>," 2010.

106. FVC2004, "<http://bias.csr.unibo.it/fvc2004/>," 2010.
107. FVC2006, "<http://bias.csr.unibo.it/fvc2006/>," 2010.
108. J. Fierrez, J. Ortega-Garcia, D. Torre-Toledano, J. Gonzalez-Rodriguez, "BioSec baseline corpus: a multimodal biometric database," *Pattern recognition*, vol. 40, pp. 1389–1392, 2007.
109. Ghoualmi, L., Chikhi, S., Draa, A., "A SIFT based feature level fusion of iris and ear biometrics," Springer International Publishing, Switzerland, LNAI 8869, pp. 102-112, 2015.
110. Zhang, Y.M., Ma, L., Li, B., "face and ear fusion recognition based on multi-agent," In *Proceedings of the Machine Learning and Cybernetics, International Conference*, pp. 46-51, 2008.
111. Ramya, M., Muthukumar, A., Kannan, S., "Multibiometric based authentication using feature level fusion," *International Conference on Advances in Engineering Science and Management*, pp.191-195, 2012.
112. Fernandez, F., Gonzalez,P., Albacete, V., Garcia, J., "Iris recognition based on SIFT features," *IEEE International Conference on Biometrics, Identity and Security*, pp.1-8, 2009.
113. Yang, G., Pang, S., Yin, Y., Li, Y., Li, X., "SIFT based iris recognition with normalisation and enhancement," *Int. J. Mach. Learn.Cybern*, vol. 4, issue 4, pp. 410-407, 2013.
114. Kavitha, S.N., Prasanna, S.C., "Multimodal biometric identification with the aid of advanced transforms and random forest classifier," *International Transaction on Engineering and Science*, vol. 1, issue 3, pp. 1-10, 2019.
115. Meena, S., Doegar, A., "Hybridization of feature level fusion with ant colony optimization in multimodal biometrics," *IJEAT*, vol. 8, issue 6, pp. 56-78, 2019.
116. Jitendra, "The Human Identification System Using Multiple Geometrical Feature Extraction of Ear –An Innovative Approach," *International Journal of Emerging Technology and Advanced Engineering*, vol. 2, issue 4, pp.662-666, 2012.
117. Anupam, S, "Ear Biometrics: A New Approach," *Conference: Advances in Pattern Recognition - The Sixth International Conference*, pp. 46-50, 2007.
118. Asmaa, Kareem "Human ear recognition using geometrical features extraction," published in *proceeding of conference on Communication, Management and informationTechnology (ICCMIT)*, *Procedia Computer Science* vol. 65, pp. 529-537, 2015.
119. Bele Bertrand , Collette Jean-Luc, Lanotto Michel, "Adaptive Prediction of Steel Hardness on Hot Strip Mill Using Neural Networks and K-Means Classifier," *IFAC proceedings volumes*, vol. 37, issue 15, pp. 125-130, 2004.
120. Li Yuan, C. Yu, "Ear verification under uncontrolled conditions with convolutional neural networks," *IET Biometrics*, special issue: Unconstrained Ear Recognition, vol. 7 Issue 3, pp. 185-198, 2018.

121. Geethanjali, N., and K. Thamaraiselvi "Feature Level Fusion of Multimodal Biometrics and Two Tier Security in ATM System," *International Journal of Computer Applications*, vol. 70, issue 14, pp. 17-23, 2013.
122. Bhardwaj, S. K. "An Algorithm for Feature Level Fusion in Multimodal Biometric System," *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)* vol. 3, pp. 3499-3503, 2014.
123. Jacob, Anna Jisma, Nikhila T. Bhuvan, and Sabu M. Thampi "Feature level fusion using multiple fingerprints," *Comput Sci.-New Dimens. Perspect* vol. 4, issue 1, pp. 13-18, 2011.
124. Ahmad, Muhammad Imran, Wai Lok Woo, and Satnam Dlay "Non-stationary feature fusion of face and palmprint multimodal biometrics," *Neuro computing*, vol.177, pp. 49-61, 2016.
125. Mondal, Arindam, and Amanpreet Kaur "Comparative Study of Feature Level and Decision Level Fusion in Multimodal Biometric Recognition of Face, Ear and Iris," vol. 5, issue. 5, pp. 822-842, 2016.
126. Xin, Ma, and Jing Xiaojun "Correlation-based identification approach for multimodal biometric fusion," *The Journal of China Universities of Posts and Telecommunications*, vol. 24, issue 4, pp. 34-50, 2017.
127. Yuan, Yun-Hao, Quan-Sen Sun, Qiang Zhou, and De-Shen Xia. "A novel multi set integrated canonical correlation analysis framework and its application in feature fusion," *Pattern Recognition*, vol. 44, issue 5, pp. 1031-1040, 2011.
128. Sarier, Neyire Deniz "Multimodal biometric Identity Based Encryption," *Future Generation Computer Systems*, vol. 80, pp. 1-17, 2018.
129. Gordon, M. and Pathak, P, "Finding Information on the World Wide Web: The Retrieval Effectiveness of Search Engines," *Information Processing and Management*, vol. 35, pp. 141-180, 1999.
130. Broder, A., "A taxonomy of web search," *SIGIR Forum*, vol. 36, issue 2, pp.3-10, 2002.
131. Salton G. "Automatic Text Processing: The Transformation, Analysis, and Retrieval of Information by Computer," Addison-Wesley Longman Publishing Co., Inc, ISBN 78-0-201-12227-5, 1989.
132. Niinuma, K., Park, U. and Jain, A.K., "Soft biometric traits for continuous user authentication," *IEEE Transactions on information forensics and security*, vol. 5, issue 4, pp.771-780, 2010.
133. Moini, A. and Madni, A.M., "Leveraging biometrics for user authentication in online learning: a systems perspective," *IEEE Systems Journal*, vol. 3, issue 4, pp. 469-476, 2009.
134. Frank, M., Biedert, R., Ma, E., Martinovic, I. and Song, D., "Touch analytics: On the applicability of touchscreen input as a behavioral biometric for continuous authentication," *IEEE transactions on information forensics and security*, vol. 8, issue 1, pp. 136-148, 2012.

135. Tsai, P.W., Khan, M.K., Pan, J.S. and Liao, B.Y., "Interactive artificial bee colony supported passive continuous authentication system," *IEEE Systems Journal*, vol. 8, issue 2, pp. 395-405, 2012.
136. Mondal, S. and Bours, P., "A computational approach to the continuous authentication biometric system," *Information Sciences*, vol. 304, pp. 28-53, 2015.
137. Saevanee, H., Clarke, N., Furnell, S. and Biscione, V., "Continuous user authentication using multi-modal biometrics," *Computers & Security*, vol. 53, pp. 234-246, 2015.
138. Murillo-Escobar, M.A., Cruz-Hernández, C., Abundiz-Pérez, F. and López-Gutiérrez, R.M., "A robust embedded biometric authentication system based on fingerprint and chaotic encryption," *Expert Systems with Applications*, vol.42, issue 2, pp.8198-8211, 2015.
139. Liu, C.L., Tsai, C.J., Chang, T.Y., Tsai, W.J. and Zhong, P.K., "Implementing multiple biometric features for a recall-based graphical keystroke dynamics authentication system on a smartphone," *Journal of Network and Computer Applications*, vol. 53, pp. 128-139, 2015.
140. Rodwell, P.M., Furnell, S.M. and Reynolds, P.L., "A non-intrusive biometric authentication mechanism utilizing physiological characteristics of the human head." *Computers & Security*, vol. 26, issue 7, pp. 468-478, 2007.
141. Azzini, A., Marrara, S., Sassi, R. and Scotti, F., "A fuzzy approach to multi-modal biometric continuous authentication. Fuzzy Optimization and Decision Making," *KES-2007*, pp. 801-080, 2008
142. Bours, P., "Continuous keystroke dynamics: A different perspective towards biometric evaluation," *Information Security Technical Report*, vol. 17, issue1, pp. 36-43, 2012.
143. Srinivasa, K.G. and Gosukonda, S., "Continuous multi-modal user authentication: coupling hard and soft biometrics with support vector machines to attenuate noise," *CSI transactions on ICT*, vol. 2, pp. 129-140, 2014.
144. Wójtowicz, A. and Joachimiak, K. "Model for adaptable context-based biometric authentication for mobile devices," *Personal and Ubiquitous Computing*, vol. 20, issue2, pp.195-207, 2016.
145. Ramu, T. and Arivoli, T., "A framework of secure biometric-based online exam authentication: an alternative to the traditional exam," *Int J scient Res*, vol. 4, issue 11, pp. 52-60, 2013.
146. Fayyoubi, A. and Zarrad, A., "Novel solution based on face recognition to address identity theft and cheating in online examination systems," *Advances in Internet of Things*, vol. 4, issue 2, 2014.
147. Traoré, I., Nakkabi, Y., Saad, S., Sayed, B., Ardigo, J.D. and de Faria Quinan, P.M., "Ensuring online exam integrity through continuous biometric authentication," *In Information Security Practices Springer, Cham*, pp. 73-81, 2017.
148. Sabbah, Y.W., Saroit, I.A. and Kotb, A.M., "A Smart Approach for Bimodal Biometric Authentication in Home-Exams (SABBAH Model)," *Biometrics and Bioinformatics*, vol.4, pp.32-45, 2012.

149. Fluck, A., Pullen, D. and Harper, C. "Case study of a computer-based examination system," *Australasian Journal of Educational Technology*, vol. 25, issue 4, 2009.
150. Adebayo, O. and Abdulhamid, S.M. "E-exams system for Nigerian universities with an emphasis on security and result integrity," *Computer Science*, vol. 23, pp. 345-351, 2014.
151. Ullah, A., Xiao, H., Barker, T., and Lilley, M., "Evaluating security and usability of profile-based challenge questions authentication in online examinations," *Journal of Internet Services and Applications*, vol. 5, issue 1, pp. 2-18, 2014.
152. Peng, G., Zhou, G., Nguyen, D.T., Qi, X., Yang, Q. and Wang, S., "Continuous authentication with touch behavioral biometrics and voice on wearable glasses". *IEEE transactions on human-machine systems*, vol. 47, issue 3, pp. 404-416, 2016.
153. Ali, M.L., Monaco, J.V., Tappert, C.C. and Qiu, M., "Keystroke biometric Systems for user authentication," *Journal of Signal Processing Systems*, vol. 86, issue 2, pp. 175-190, 2017.
154. Ullah, A., Xiao, H. and Barker, T., "A study into the usability and security implications of text and image-based challenge questions in the context of online examination," *Education and Information Technologies*, vol. 24, issue 1, pp.13-39, 2019.
155. Ullah, A., Xiao, H. and Barker, T., "A dynamic profile questions approach to mitigate impersonation in online examinations," *Journal of Grid Computing*, vol. 17, issue 2, pp. 209-223, 2019.
156. Prakash, A., "Continuous user authentication-based score level fusion with hybrid optimization," *Cluster Computing*, vol. 22, issue 5, pp.12959-12969, 2019.
157. Yang, Y., Guo, B., Wang, Z., Li, M., Yu, Z. and Zhou, X., "BehaveSense: Continuous authentication for security-sensitive mobile apps using behavioral biometrics," *Ad Hoc Networks*, vol. 84, pp. 9-18, 2019.
158. Vasanthi, M., Seetharaman, K., "Facial image recognition for biometric authentication systems using a combination of geometrical feature points and low-level visual features," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, issue 7, pp. 4109-4121, 2022.
159. Sunaryono, D., Siswanto, J., Radityo Anggoro, "an android-based course attendance system using face recognition," *Journal of King Saud University - Computer and Information Sciences*, vol. 33, issue 3, pp. 304-312, 2021.
160. Kevin S. Killourhy and Roy A. Maxion. "Comparing Anomaly Detectors for Keystroke Dynamics," in *Proceedings of the 39th Annual International Conference on Dependable Systems and Networks (DSN-2009)*, pp. 125-134, 2009.
161. Loy, C.C. Lai, W.K., and Lim., C.P., "Keystroke Patterns Classification using the ARTMAP-FD Neural Network," *International Conference on Intelligent Information Hiding and Multimedia Signal Processing, IHHMSP*, pp. 61-64, 2007.
162. Loy, C.C., Lim, C.P., and Lai. W.K., "Pressure-based Typing Biometrics User Authentication Using the Fuzzy ARTMAP Neural Network," *International Conference on Neural Information Processing, Taiwan, ICONIP*, 2005.

163. FVC2004 - Third International Fingerprint Verification Competition (unibo.it)
164. The Database of Faces (cam-orl.co.uk)
165. GitHub-massimomauro/FASSEG-repository: Datasets for multi-class and multi-pose face segmentation.
166. Jaafar, Haryati and Dzati Athiar Ramli. "A Review of Multibiometric System with Fusion Strategies and Strategies and weighting factor," *International Journal of Computer Science Engineering (IJCSE)*, vol. 2, no. 4, pp. 158–165, 2013.
167. Domingos, P. "A Few Useful Things to Know About Machine Learning," *Communication, ACM*, vol. 55, pp. 78–87, 2012.
168. Mikołajczyk, A.; Grochowski, "M. Data augmentation for improving deep learning in image classification problem," In *Proceedings of the International Interdisciplinary PhD Workshop, Swinoujscie*, pp. 117–122, 2018.
169. Shorten, C.; Khoshgoftaar, T.M. "A survey on Image Data Augmentation for Deep Learning," *J. Big Data*, vol. 1, pp. 1-48, 2019.
170. Pan, S.J.; Yang, Q. "A Survey on Transfer Learning. *IEEE Transction*," *Knowl. Data Eng.*, vol. 22, pp. 1345–1359, 2010.
171. Hinton, G.; Osindero, S.; Teh, Y.W. "A fast-learning algorithm for deep belief nets," *Neural Comput.*, vol. 18, pp. 1527–1554, 2006.
172. Hubel, D.H.; Wiesel, T.N. "Receptive Fields and Functional Architecture of Monkey Striate Cortex," *J. Physiol.*, vol. 195, pp. 215–243, 1998.
173. Simonyan, K.; Zisserman, A. "Very deep convolutional networks for large-scale image recognition," *CoRR*, 2014.
174. Yin, Y.; Liu, L.; Sun, X. SDUMLA-HMT "A multimodal biometric database," In *Proceedings of the Chinese Conference on Biometric Recognition*, vol. 7098, pp. 260–268, 2011.
175. Veluchamy, S., Karlmarx, L. "System for multimodal biometric recognition based on finger knuckle and finger vein using feature-level fusion and k support vector machine classifier," *IET Biometrics*, vol. 6, issue 3, pp. 232–242, 2017.
176. Yang, W., Huang, X., Zhou, F., et al., "Comparative competitive coding for personal identification by using finger vein and finger dorsal texture fusion," *Inf. Sci.*, 268, pp. 20–32, 2014.
177. Yang, W., Qin, C., Wang, X., et al.: "Cross section binary coding for fusion of finger vein and finger dorsal texture," *IEEE Int. Conf. on Industrial Technology (ICIT)*, pp. 742–745, 2016.
178. Abomhara, Mohamed, Sule Yildirim Yayilgan, and Livinus Obiora Nweke, "A Comparison of Primary Stakeholder's iews on the Deployment of Biometric Technologies in Border Management : Case Study of SMart MobILity at the European Land Borders," *Technology in Society*, vol.64, 2021.
179. Breitinger, Frank, Ryan Tully-doyle, and Courtney Hassenfeldt, "A Survey on Smartphone User 's Security Choices, Awareness and Education," *Computers & Security*, vol.88, pp. 1–14, 2020.
180. Buckley, Oliver, and Jason R. C. "The Language of Biometrics : Analysing Public Perceptions," *Journal of Information Security and Applications*, vol.47, pp. 112–19, 2019.

181. Anil, Anu P., and Satish K.P. “Investigating the Relationship Between TQM Practices and Firm’s Performance: A Conceptual Framework for Indian Organizations,” *Procedia Technology* vol. 24, pp. 554–61, 2016.
182. Faheem, Aslam, Mughal Khurram S., Ali Ashiq, and Mohmand Yasir Tariq. “Forecasting Islamic Securities Index Using Artificial Neural Networks: Performance Evaluation of Technical Indicators,” *Journal of Economic and Administrative Sciences*, vol. 37, issue 2, pages 253-271, September. 2020.
183. Khayer, Abul, Shamim Talukder, Yukun Bao, and Nahin Hossain, “Cloud Computing Adoption and Its Impact on SMEs ’ Performance for Cloud Supported Operations : A Dual-Stage Analytical Approach,” *Technology in Society*, vol. 60, pp. 101-225, September 2020.
184. Buriro, Attaullah, Bruno Crispo, and Mauro Conti., “AnswerAuth : A Bimodal Behavioral Biometric-Based User Authentication Scheme for Smartphones,” *Journal of Information Security and Applications* vol. 44, pp. 89–103, 2019.
185. Carrión-ojeda, Dustin, Rigoberto Fonseca-delgado, and Israel Pineda, “Analysis of Factors That Influence the Performance of Biometric Systems Based on EEG Signals,” *Expert Systems With Applications*, vol.165, 2021.
186. Gupta, S., Choudhary, H., & Agarwal, D. R. “An Empirical Analysis of Market Efficiency and Price Discovery in Indian Commodity Market,” *Global Business Review*, vol. 19, issue 3, pp. 771–789, 2018.
187. Kolli, Sekhar, C., and Tatavarthi. U.D., “Fraud Detection in Bank Transaction with Wrapper Model and Harris Water Optimization-Based Deep Recurrent Neural Network,” *Kybernetes*, Emerald Publishing Limited, vol.50, pp.1731-1750. 2020.
188. Chowdary, Kalpana, M., and Hemanth. D., “Human Emotion Recognition Using Intelligent Approaches : A Review,” *Intelligent Decision Technologies*, vol. 13, pp. 417–33, 2019.
189. Dargan, Shaveta, and Kumar., M., “A Comprehensive Survey on the Biometric Recognition Systems Based on Physiological and Behavioral Modalities,” *Expert Systems With Applications*, vol.143, pp. 113-114, 2020.
190. Durdyev, Serdar, Ali Ihtiyar, Audrius Banaitis, and Derek Thurnell. “The Construction Client Satisfaction Model : A PLS-SEM Approach,” *Journal of Civil Engineering and Management*, vol. 24, issue1, pp. 31–42, 2018.
191. Yogesh K., Tubadji, Paul Walton, and Michael D. Williams, “Artificial Intelligence (AI): Multidisciplinary Perspectives on Emerging Challenges , Opportunities , and Agenda for Research , Practice and Policy,” *International Journal of Information Management* vol. 57, 2021.
192. Joshi, Mahesh, Bodhisatwa Mazumdar, and Somnath Dey, “A Comprehensive Security Analysis of Match-in-Database Fingerprint Biometric System R.” *Pattern Recognition Letters* vol.138, pp.247–66, 2020.
193. Chong, Alain Yee-loong, “A Two-Staged SEM-Neural Network Approach for Understanding and Predicting the Determinants of m-Commerce Adoption.” *Expert Systems With Applications* vol.40, issue4, pp.1240–47, 2013.
194. Choudhury, Surabhi Hom, Kumar, A., and Laskar., S.H. “Adaptive Management

- of Multimodal Biometrics - A Deep Learning and Metaheuristic Approach, ”
Applied Soft Computing Journal, vol.106, pp. 81-88, 2021.
195. Khan, Maleika Heenaye-Mamode, and Khan, N.M., “Analysing Factors Affecting Hand Biometrics during Image Capture,” *Procedia - Computer Science*, vol. 32, pp. 521–28, 2014.
 196. Königstorfer, Florian, and Stefan Thalmann, “Applications of Artificial Intelligence in Commercial Banks – A Research Agenda for Behavioral Finance,” *Journal of Behavioral and Experimental Finance*, vol. 27, pp. 56-63, 2020.
 197. Krishnakumar, P., Rameshkumar, K., and Ramachandran, K.I., “Machine Learning Based Tool Condition Classification Using Acoustic Emission and Vibration Data in High Speed Milling Process Using Wavelet Features,” *Intelligent Decision Technologies*, vol. 12, pp. 265–82, 2018.
 198. Kumar, Sachin, and Zymbler., M., “A Machine Learning Approach to Analyze Customer Satisfaction from Airline Tweets,” *Journal of Big Data*, vol. 6, issue 62, pp. 1–16, 2019.
 199. Leong, Lai-ying, Teck-soon Hew, Keng-boon Ooi, and June Wei, “A Two-Stage Structural Equation Modeling-Artificial Neural Network Approach,” *International Journal of Information Management*, vol. 51, pp. 102-47, 2019.
 200. Liébana-cabanillas, Francisco, Veljko Marinkovic, Iviane Ramos, De Luna, and Zoran Kalinic. “Predicting the Determinants of Mobile Payment Acceptance : A Hybrid SEM-Neural Network Approach,” *Technological Forecasting & Social Change*, vol.129, pp.117–30, 2017.
 201. Nappi, Michele, Stefano Ricciardi, and Massimo Tistarelli., “Context Awareness in Biometric Systems and Methods : State of the Art and Future Scenarios,” *Image and Vision Computing*, vol.76, pp. 27–37, 2018.
 202. Panigrahi, Ranjit, Samarjeet Borah, Akash Kumar Bhoi, Muhammad Fazal Ijaz, Moumita Pramanik, Rutvij H. Jhaveri, and Chiranji Lal Chowdhary, “Performance Assessment of Supervised Classifiers for Designing Intrusion Detection Systems: A Comprehensive Review and Recommendations for Future Research,” *Mathematics*, vol. 9, issue 6, pp. 1–32, 2021.
 203. Ijaz, Muhammad Fazal. “Data-Driven Cervical Cancer Prediction Model with Outlier Detection and Over-Sampling Methods,” *Sensors*, MDPI, vol. 20, pp. 1–22, 2020.
 204. Abbas, Jawad, “Impact of Total Quality Management on Corporate Sustainability through the Mediating Effect of Knowledge Management,” *Journal of Cleaner Production*, vol. 244, pp.118-806, 2020.
 205. Dadhich, M., Pahwa, M.S., and Rao. S.S., “Factor Influencing to Users Acceptance of Digital Payment System,” *International Journal of Computer Sciences and Engineering*, vol. 06, issue 09, pp.46–50, 2020.
 206. Ju, Teresa L., Binshan Lin, Chinho Lin, and Hao Jung Kuo, “TQM Critical Factors and KM Value Chain Activities,” *Total Quality Management and Business Excellence*, vol.17, issue.3, pp. 373–93, 2006.
 207. Arifin, Wan Nor, and Muhamad Saiful Bahri Yusoff, “Confirmatory Factor Analysis of the Universiti Sains Malaysia Emotional Quotient Inventory Among

- Medical Students in Malaysia,” SAGE Open, vol.6, issue 2, 2016.
208. Gupta, Garima, and Nagpal., S., “Green Dimensions, Environment Orientation and Size: Impact Assessment on Operational Performance of Manufacturing Firms,” *Global Business Review*, vol. 1, issue 14, pp.1–14, 2020.
 209. Siyal, Abdul Waheed, and Donghong Ding, “M-Banking Barriers in Pakistan : A Customer Perspective of Adoption and Continuity Intention,” *Data Technologies and Applications* vol. 53, issue 1, pp. 58–84, 2018.
 210. Yacob, Peter, Lai Soon Wong, and Saw Chin Khor, “An Empirical Investigation of Green Initiatives and Environmental Sustainability for Manufacturing SMEs,” *Journal of Manufacturing Technology Management*, vol. 30, issue 1, pp. 2–25, 2019.
 211. Dubey, Rameshwar, Angappa Gunasekaran, and Sadia Samar Ali. “Exploring the Relationship between Leadership, Operational Practices, Institutional Pressures and Environmental Performance: A Framework for Green Supply Chain,” *International Journal of Production Economics*, vol. 160, pp. 120–32, 2015.
 212. Hair, J. F., Jr., Black, W. C., Babin, B. J., & Anderson, R. E.. n.d. *Multivariate Data Analysis (7th Ed.)*. Upper Saddle River, NJ: Pearson Prentice Hall, 2009.
 213. Dadhich, Manish, Himanshu Purohit, and Anand A. Bhasker, “Determinants of Green Initiatives and Operational Performance for Manufacturing SMEs,” *Materials Today: Proceedings* vol. 46, pp. 10870-874, 2021.
 214. Modisane, Pheny, Osden Jokonya, and Pheny Modisane. “ScienceDirect Evaluating the Benefits of Cloud Computing in Small , Medium and Enterprises (SMMEs) Enterprises (SMMEs).” *Procedia Computer Science* vol.181,pp.784–92, 2021.
 215. Singh, Kumar, G., and Dadhich., M, “Impact of Total Quality Management on Operational Performance of Indian Cement Manufacturing Industry- A Structural Equation Remodeling Approach.” *TURCOMAT*, pp.:22–41, 2021.
 216. Mashelkar, R. A. “Exponential Technology , Industry 4 . 0 and Future of Jobs in India.” *Review of Market Integration*, vol.10, issue2, pp.138–57, 2018.
 217. Agrawal, Nishant Mukesh, “Modeling Deming’s Quality Principles to Improve Performance Using Interpretive Structural Modeling and MICMAC Analysis,” *International Journal of Quality and Reliability Management*, vol. 36, issue 7, pp. 1159–80, 2019.
 218. Birda, R.K., & Dadhich.,M., “Study of ICT and E-Governance Facilities in Tribal District of Rajasthan,” *ZENITH International Journal Of Multidisciplinary Research*, vol. 9, issue7, pp.39–49, 2019.
 219. Singhal, Neeraj, “An Empirical Investigation of Industry 4.0 Preparedness in India,” *Vision*, vol. 1 , issue 12, pp.1–12, 2020.
 220. Omar, Normah, Zulaikha Amirah Johari, and Malcolm Smith, “Predicting Fraudulent Financial Reporting Using Artificial Neural Network,” *Journal of Financial Crime*, vol. 24, issue 2, pp. 362–87, 2015.
 221. Bernard, Robert M., Eugene Borokhovski, Richard F. Schmid, Rana M. Tamim, and Philip C. Abrami, “A Meta-Analysis of Blended Learning and Technology

- Use in Higher Education: From the General to the Applied,” *Journal of Computing in Higher Education*, vol. 26, issue 1, pp. 87–122, 2014.
222. Langevin, J., Reyna, J.L., Ebrahimigharehbaghi, S., “Developing a Common Approach for Classifying Building Stock Energy Models” *Renewable and Sustainable Energy Reviews*, vol. 133, 2020.
 223. Gyamfi, Nana Kwame, “Bank Fraud Detection Using Support Vector Machine,” *IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)* pp. 37–41, 2018.
 224. Tornatzky, L.G., Fleischer, M. “The Processes of Technological Innovation,TOE,” *Journal of Technological Transfer*, vol. 16, pp.45–46, 1990.
 225. Sadasivam, Sudha, G., and Maha Lakshme, S., “Corporate Governance Fraud Detection from Annual Reports Using Big Data Analytics Mutyala Subrahmanyam and Dasaraju Himachalam Bhanu Prasad Pinnamaneni,” *Int. J. Big Data Intelligence, Inderscience Enterprises Ltd.* vol. 3, issue 1, pp. 51–60, 2016.
 226. Heracleous, Loizos, and Templeton College. “Biometrics : The next Frontier in Service Excellence , Productivity and Security in the Service Sector,” *Managing Service Quality*, vol.16, issue1, pp.12–22, 2006.
 227. Teh, Pin Shen, Ning Zhang, and Ke Chen, “TDAS : A Touch Dynamics Based Multi-Factor Authentication Solution for Mobile Devices,” *International Journal of Pervasive Computing and Communications*, vol. 12, issue1, pp.127–53, 2016.
 228. Gandhmal, Dattatray P., and Kumar, K., “Wrapper-Enabled Feature Selection and CPLM-Based NARX Model for Stock Market Prediction,” *The Computer Journal*, vol. 8, pp. 341-351, 2020.
 229. Raut, Rakesh D., Priyadarshinee, P., Bhaskar B. Gardas, and Kumar, M., “Analyzing the Factors Influencing Cloud Computing Adoption Using Three Stage Hybrid SEM-ANN-ISM (SEANIS) Approach,” *Technological Forecasting & Social Change*, vol. 134, pp. 98–123, 2018.
 230. Gutierrez, A., Boukrami, E., and Lumsden.,R, “Technological Organisational and Environmental Factors Influencing Managers ’ Decision to Adopt Cloud Computing in the UK,” *Journal of Enterprise Information Management*, vol. 28, issue 6, pp. 788–807, 2015.
 231. Barde, S., Zadgaonkar, A.S., Sinha, G.R., *Multimodal Biometrics using face, ear and iris modalities*, IJCA, NCRAIT, vol. 2, pp. 9-15, 2014.
 232. Jain, A.K., Prabhakar, S., Hong, L et al., “A multichannel approach to fingerprint classification,” *IEEE Trans.PAMI*, vol. 21, issue 4, pp. 348-359, 1999.
 233. GovindaRajan, R. et al. “Feature level fusion using hand and face biometrics,” *Proceedings of SPIE*. vol. 57 , pp. 196-204, 2005.
 234. Sinha, G.R., Shrivastava, A., et al., “Biometric Security Technologies : A case study,” *International Journal of Image Processing and Applications (Interantional Science Progress)*, I(1): pp. 1-5, 2010.
 235. Benitez, A.B., Chang, S.F., “Multimedia knowledge integration, summarization and evaluation in proceeding of Workshops on Multimedia,” *Data Mining* ,vol. 2326, pp. 39-50, 2012.

236. Ayodele, O., Adegbenjo, A., "Current practices in information fusion for multimodal biometrics," American Journal Of Engineering Research(AJER), e-ISSN:2320-0847, p-ISSN:2320-0936, vol.6 , Issue 4 , pp, 148-154, 2017.
237. Srivastava, H., "Personal Identification using Iris recognition system, a review," IJERA, vol-3, issue-3, pp.449-453, 2013.
238. Daugman, J.G., "High confidence visual recognition of persons by a Test of Statistical independence," IEEE Trend Pattern Anal. Mach. Intelligence, vol.15, issue 11, pp.1148-1161, 1993.
239. Ranjan, S., Prabhu, S., Swarnalatha, P., Magesh, G., SundaraRajan R, "Iris recognition system," IRJET, e-ISSN:2395-0056, vol. 4, issue 12, 2017.
240. Choudhary, R., Ghosh, D., Agarwal, P., Bandyopadhyay, S.K., "Ear based biometric authentication system," WJERT, vol-2, issue-5, 2016.
241. Choras, M., "Emerging methods of Biometric human identification," Proceedings of the second International Conference on Innovative Computing, Information and control, pp:365-369, 2007.
242. Kuduk, N., Hinge, A., KshirSagar, K., "Biometric Recognition System," IRJET, vol.04, issue 01, pp. 111-115, 2017.
243. Besbes, F, Trochili, H., Solaiman, B. "Multimodal Biometric System Based on Fingerprint Identification and Iris Recognition," Information and Communication Technologies: From Theory to Applications, ICTTA 2008, 3rd International Conference, pp. 231-233, 2008.
244. Patil, P., Bormane, D.S., "Multilevel Security using score level fusion of face and ear based biometric modalities," IJERT, vol-3, Issue-5, 2014.
245. Ning Wang, Ahmed, A., Qiong Li, Latif, A., Jialang, P., Xiami, N., "Multi biometric complex fusion for visible and thermal face images," IJSP, vol.6, issue 3, pp. 67-71, 2013.
246. Karanwal, S., "Secure and reliable multimodal biometric systems using two level and three level biometric traits," International Journal of Advanced Research in Computer Science and Software Engineering, vol.3, no.7, pp. 34-39, 2013.
247. Metri, P., Ghorpade, J., Butalia, A., "Facial emotion recognition using context based multimodal approach," IJIM & AI, vol.1, no.4, pp.12-15, 2012.
248. Daramola, S.A., Oluwaniuyo, O.D., "Automatic ear recognition system using BPNN," IJVIPNS- IJENS, vol.11, issue 1, 2011.
249. Yang, Z., Ronggay, W., "Multilevel Modified Finite Random Transform Network for image upsampler," IEEE, 2015.
250. Jawale, J.B., Bhalchandra, A.S., "The human identification system using multiple geometrical feature extraction of ear- an innovative approach," IJERT, vol-3, issue-5, pp. 108-113, 2014.
251. Kisku, D.P., Gupta, P., Mehrotra, H., Singh, J., "Multimodal belief fusion for face and ear biometrics," Intelligent Information Management, vol.1, pp. 166-177, 2009.
252. Shubhangi, S., Deshmukh, R.R., "Scale Invariant Feature Transform based Multimodal biometric system with face and finger," IJCSIS, vol. 11, issue 7, pp. 45-51, 2013.

253. Soyel, H., Demirel, H., “ Facial expression recognition based on discriminative scale invariant feature transform,” *Electronic letters* , vol.46, issue 5, pp. 343, 2010.
254. Faisal R., Al –Osaimi, Mohd. Bennamoun, Mian, A., “Spatially optimised data level fusion of texture and shape for face recognition,” *IEEE Transactions on Image Processing*, vol.21, issue 2, pp. 859-22, 2012.
255. Asha, S.,Chellappan, C., “Authentication of E- Learners using multimodal biometric technology,” *IEEE*, pp.1-6, 2008.
256. Lowe, D.G., “Distinctive image features from Scale invariant keypoints,” *International Journal of Computer Vision*, vol.60, issue 2, pp. 91-110, 2006.

List of Publications

Journals

1. Purohit, H., Ajmera, P.K., "Optimal feature level fusion for secured human authentication in multimodal biometric system," *Machine Vision and Applications (Springer)*, vol.32, issue 24, pp.1-24, 2021.
2. Purohit, H., Ajmera, P.K., "Multi-modal biometric fusion based continuous user authentication for E-proctoring using hybrid LCNN-Salp swarm optimization," *Cluster Computing*, (Springer) vol. 25, pp.827–846, 2022.
3. Purohit, H., Ajmera, P.K. "Analytical Study on Users' Awareness and Acceptability towards Adoption of Multi-modal Biometrics (MMB) Mechanism in Online Transactions: A Two-Stage SEM-ANN Approach," *Multimedia tools and applications*, (Springer), vol. 82, pp. 14239-14263, 2022.
4. Purohit, H., Ajmera, P.K., "Multi-modal biometric systems: A brief study," *International Journal of Innovative Technology and Exploring Engineering*, (New India), vol.8, Issue 7, pp. 108–111, 2018.
5. Purohit, H., Ajmera, P.K., "Modified VGG 16 Based Multimodal Biometric Person Authentication," Under review- *Journal of Super Computing*.

Conferences

1. Purohit, H., Ajmera, P.K., "Multimodal Multilevel Fusion of Face Ear Iris with Multiple Classifications", *Modelling simulation and Intelligent Computing* (Lecturer notes in Electrical Engineering) Mosaicom2020, vol. 659, pp. 345-355, 2020.
2. Purohit, H., Ajmera, P.K., "Fusions of Palm Print with Palm-Phalanges Print and Palm Geometry" *International Conference on Advanced Computing Networking and Informatics*, (Advances in Intelligent Systems and Computing) ICANI2018, vol. 870, pp. 553–560, 2018
3. Purohit, H., Ajmera, P.K., "Automated Person Authentication Using Face, Iris and Ear Multimodal Biometric Fusion", *ICT for Competitive strategies*, ICTCS2019, edition 1st, 2019.

Book Chapter

1. Purohit, H., Ajmera, P.K., "Contemporary Biometric System Design", *Handbook of Research on Engineering Innovation and Technology Management in Organization*, ISBN 1799827720, pp 292-324, 2017.

Candidate's Biography

Himanshu Purohit obtained his bachelor's degree in Electronics and Communication Engineering from the University of Rajasthan, India. Then he obtained his master's degree in Communication Engineering and PhD in Signal Processing majoring in Multimodal Biometric authentication, both from BITS Pilani, India. Currently, he is working as a Head of Corporate Interface & Admissions, Executive head of Incubation and Assistant Professor ECE, at Sir Padampat Singhania University, Udaipur, India. His current research interests are Multimodal Biometric Fusion, Artificial Intelligence, Electrical Vehicle Design, and Signal Processing.

Supervisor's Biography

Dr. Pawan K Ajmera: Presently working at EEE BITS Pilani as an Associate Professor. His research interest includes multimodal biometrics, signal processing, speech processing, and artificial intelligence. He has published many papers in journals of international repute and the highest standard. He is having more than a decade of experience in teaching engineering subjects at premier institutes in India.

Annexure

Important Databases and its features

Sr. No.	Dataset Name	Modality	Number of Samples	Distinguishing Factors
1	FERET	Face	14,126	Images taken in controlled conditions with variation in expression, lighting, and pose
2	LFW	Face	13,233	Images taken from the internet with large variation in pose, expression, and lighting
3	CASIA-Iris	Iris	756	Images taken under both visible and near-infrared illumination
4	PolyU Palmprint	Palm	3,258	Images taken under both controlled and uncontrolled conditions with variation in pose, illumination, and occlusion
5	NIST SD 27	Fingerprint	258,000	Images taken under various conditions with varying image quality and resolution
6	Honkong Polytech-Disguise and Makeup	Face	2460 from 410 samples	These images have been acquired under real environment which primarily focuses on make-up and disguises covariates. This database also provides ground truth (eyeglass, goggle, mustache, and beard). Each person in the database has six images, the first image is selected to represent clean impression, i.e., without any makeup and disguises, and is from the frontal pose. The rest of the five face images are faces with different levels of makeup or disguise creating accessories.
7	IIT D Near IR Database	Face	612 images from 102 users	The IIT Delhi near infrared face image database consists of the faces collected from the students and staff at IIT Delhi, New Delhi, India. This database has been acquired in IIT Delhi campus during Feb - Jun 2007 (still in progress) using a webcam with near infrared illumination. All the subjects in the database are in the age group 17-50 years. The database was acquired in two stages and in each stage three images were acquired for every user. The database is organized into two folders (each with 306 images from 102 users). The resolution of these images is 768 × 576 pixels.
8	I-SOCIAL-DB	Iris	3,286 images of 400 individuals	This dataset is composed of 3,286 ocular regions, extracted from 1643 high resolution face images of 400 individuals, collected from public websites. For each ocular region, a human expert extracted the coordinates of the circles approximating the inner and outer iris boundaries and performed a pixelwise segmentation of the iris contours, occlusions, and reflections.
9	HYPUCross Spectral Iris Images	Iris	12,540 images of 209 users	This database of iris images has been acquired under simultaneous bi-spectral imaging, from both right and left eyes, using the acquisition setup. Each of the iris images are of 640 × 480 pixels size and with pixel correspondences in both the near-infrared and visible iris images. Each subject's folder consists of two folders, namely VIS and NIR.
10	IIT Delhi Iris Database	Iris	1120 images of 224 users	The IIT Delhi Iris Database mainly consists of the iris images collected from the students and staff at IIT Delhi, New Delhi, India. This database has been acquired in Biometrics Research Laboratory during Jan - July 2007 (still in progress) using JIRIS, JPC1000, digital CMOS camera.

				All the subjects in the database are in the age group 14-55 years comprising of 176 males and 48 females. The resolution of these images is 320 ´ 240 pixels and all these images were acquired in the indoor environment.
11	UBIRIS.v1	Iris	1877 images collected from 241 persons	Database is composed of 1877 images collected from 241 persons during September, 2004 in two distinct sessions. Its most relevant characteristic is to incorporate images with several noise factors, simulating less constrained image acquisition environments. This enables the evaluation of the robustness of iris recognition methods.
12	SVBPI Sclera Blood Vessels, Periocular and Iris)	Iris	1858 images of 55 users	It is a publicly available dataset designated primarily for research into sclera recognition, but it is also suitable for experiments with iris and periocular recognition techniques. It consists of 1858 high-resolution RGB images of eyes from 55 subjects. Images of the dataset were captured during a single recording session with a digital single-lens reflex camera (DSLR) (Canon EOS 60D). The age group was of 15 to 80 years.
13	MOBIUS (Mobile ocular biometrics in unconstrained settings)	Iris	16717 images of 100 subjects	This dataset consists of 16,717 RGB ocular images collected from 100 subjects. The images are high-resolution and were captured with the cameras of 3 different commercial mobile phones and in 3 different capture environments.
14	IIT D Touchless Palmprint database	Palmprint	1610 Images of 230 subjects	The IIT Delhi palmprint image database consists of the hand images collected from the students and staff at IIT Delhi, New Delhi, India. All the images are collected in the indoor environment and employ circular fluorescent illumination around the camera lens. The currently available database is from 230 users. All the subjects in the database are in the age group 12-57 years. Seven images from each subject, from each of the left and right hand, are acquired in varying hand pose variations. The resolution of these images is 800 ´ 600 pixels.
15	PolyU-IIT D Touchless Palmprint DB	Palmprint	12000 images from 60 subjects	This database has been acquired from the volunteers in India and China. This is first such joint database of its kind and acquired at various locations using a general-purpose handheld camera. This new palmprint database has been acquired over several years from (over) 600 different <i>subjects</i> which is largest to-date until August 2020 literature. Each subject in this database has provided his/her left- and right-hand images. The images in this database therefore also have high scale variations and are acquired from subjects. There are twenty images from each of the subjects whose age ranges from 5 years to 72 years.
16	HKPU Hand Dorsal	Hand Dorsal	2505 images of 501 subjects	The Hong Kong Polytechnic University Contactless Hand Dorsal Images Database is contributed from the male and female volunteers. This database has 2505 hand dorsal images from the right hand of 501 different subjects that illustrate three knuckle patterns in each of the four fingers from the individual subject. This database also provides segmented/normalized major, first minor and second minor knuckle images using completely automated segmentation.
17	IIT D EAR DB	EAR	471 images from 121 subjects	The currently available database is acquired from the 121 different subjects and each subject has at least three ear images. All the subjects in the database are in the age group 14-58 years. The database of 471 images has been sequentially numbered for every user with an integer identification/number. The resolution of these images is 272 ´ 204 pixels and all these images are available in jpeg format. In addition to the original images, this database also

				provides the automatically normalized and cropped ear images of size 50 ´ 180 pixels.
18	Annotated Web Ear DB	EAR	1000 Images of 100 subjects	This dataset was collected using images from the web to ensure large variability based from unconstrained environments. 100 subjects were selected among some of the most famous people, across different ethnicities, genders and ages. For each subject 10 images were selected and images tightly cropped. each image is also annotated and stored in JSON files with the annotations listed below. This dataset was later also incorporated into the UERC datasets.
19	IAPRA JB DB	Face-Videos	138000 images	The IARPA Janus Benchmark face challenge (IJB-B, IJB-C) provides various databases for addressing verification, identification, detection, clustering, and processing of full motion videos. The IJB-C dataset consists of 138000 face images, 11000 face videos, and 10000 non-face images.
20	HKPU-Finger Knuckle DB	Finger Knuckle	1951 images of 228 subjects	The Hong Kong Polytechnic University contactless finger knuckle images database (Version 3.0) is contributed from the male and female volunteers. This database has 1950 finger knuckle images from the index finger of 221 subjects, all the images are in .JPG format.
21	HKPU-Finger Vein DB	Finger Vein	6264 images of 156 subjects	The Hong Kong Polytechnic University finger image database consists of simultaneously acquired finger vein and <i>finger surface texture</i> images from the male and female volunteers. The currently available database has 6264 images from the 156 subjects, all the images are in bitmap (*.bmp) format. In this dataset about 93% of the subjects are younger than 30 years. Each of the subjects provided 6 image samples from index finger middle finger respectively, and each sample consisting of one finger vein image and one finger texture image from left hand.
22	HKPU-Finger Print DB 1.0	Fingerprint CL 2D & CW 2D	3600 images of 300 subjects	A new fingerprints database acquiring from 300 different clients consists of 1800 2D contactless fingerprint images and corresponding 1800 2D contact-based fingerprints is developed and made available publicly. Database was acquired during September 2014 to February 2016.
23	IITD-HKPU Low Resolution DB	Fingerprint	1466 images of 156 subjects	The Hong Kong Polytechnic University low resolution fingerprint database consists of low-resolution finger surface images. This database has 1466 images from the 156 subjects, all the images are in bitmap (*.bmp) format. In this dataset about 93% of the subjects are younger than 30 years.