HUMAN BEHAVIOR ANALYSIS USING SMARTPHONE SENSOR DATA

Ph.D. Thesis

By ANKITA JAIN



DISCIPLINE OF ELECTRICAL ENGINEERING INDIAN INSTITUTE OF TECHNOLOGY INDORE JULY 2018

HUMAN BEHAVIOR ANALYSIS USING SMARTPHONE SENSOR DATA

A THESIS

Submitted in partial fulfillment of the requirements for the award of the degree

of

DOCTOR OF PHILOSOPHY

by ANKITA JAIN



DISCIPLINE OF ELECTRICAL ENGINEERING INDIAN INSTITUTE OF TECHNOLOGY INDORE JULY 2018



INDIAN INSTITUTE OF TECHNOLOGY INDORE

CANDIDATE'S DECLARATION

I hereby certify that the work which is being presented in the thesis entitled "Human Behavior Analysis using Smartphone Sensor Data" in the partial fulfillment of the requirements for the award of the degree of DOCTOR OF PHILOSOPHY and submitted in the DISCIPLINE OF ELECTRICAL ENGINEERING, Indian Institute of Technology Indore, is an authentic record of my own work carried out during the time period from December 2013 to July 2018 under the supervision of Dr. Vivek Kanhangad, Associate Professor, Indian Institute of Technology Indore, India.

The matter presented in this thesis has not been submitted by me for the award of any other degree of this or any other institute.

Signature of the student with date (ANKITA JAIN)

This is to certify that the above statement made by the candidate is correct to the best of my/our knowledge.

Signature of Thesis Supervisor with date

(DR. VIVEK KANHANGAD)

ANKITA JAIN has successfully given his/her Ph.D. Oral Examination held on

Signature of Chairperson (OEB) Date:	Signature of External Examiner Date:	Signature(s) of Thesis Supervisor(s) Date:
Signature of PSPC Member #1 Date:	Signature of PSPC Member #2 Date:	Signature of Convener, DPGC Date:
Signature of Head of Discipline Date:		

ACKNOWLEDGEMENTS

First and foremost, I am grateful to IIT Indore for providing me with this opportunity to pursue Ph.D. in such a prestigious institute.

It gives me immense pleasure to express my sincere thanks to my supervisor **Dr. Vivek Kanhangad**, for his excellent supervision, patience and care during the entire Ph.D. He always provided valuable suggestions and corrected my mistakes patiently. It has been a wonderful learning experience working with him throughout my doctoral work and I look forward to our future collaborations.

I express my sincere thanks to my research progress committee members, **Prof. Ram Bilas Pachori** and **Dr. Somnath Dey**, for their invaluable time, encouragement, and constructive comments, which helped me a lot in improving the quality of my research work.

I am extremely grateful to **Bommakanti K. Chaitanya** (**B.Tech. CS**) for helping me develop the Android application used for collection of touch gesture and gait data. I would also like to thank all the volunteers, who have participated in data collection process, for their valuable time and patience.

I had a memorable time and learning experience with my colleagues from PRIA Lab T. Sunilkumar, Shruti Bhilare, Vijay Anand, Shishir Maheshwari and Mahesh Gour and with my friends Anshu Mishra, Dr. Devendra Gurjar, Dr. Deepika Gupta, Dr. Nagendra Kumar and Dr. Manish Mandloi. I would like to thank all of them for their help, constant encouragement and moral support during the entire duration of my Ph.D. I gratefully acknowledge the staff members of IIT Indore for their help and support.

I owe a lot to my family, who encouraged and helped me at every stage of my personal and academic life. Specially, I would like to thank my father and my uncle for their constant support and encouragement in every decision of mine. I would also like to thank my mother, who taught me the social and family values.

I would also like to mention special thanks to my husband whom I was lucky to have with me as a Ph.D. scholar at IIT Indore. I would like to thank him for always supporting me and encouraging me throughout the good and the bad times.

Dedicated To My Beloved Family

ABSTRACT

Over the past decade, smartphones have become an integral part of our daily life. With the advancements in technology, smartphones have become feature-rich, affordable and very sophisticated containing several built-in sensors such as orientation sensor, accelerometer sensor and gyroscope sensor. Equipping smartphones with intelligence has been a topic of interest to researchers working in diverse fields such as biometrics, healthcare, and financial services.

Readings from dedicated body-worn inertial sensors have been shown to carry information useful for various tasks including human gait recognition, sports activity recognition, fall detection and health monitoring. In general, the major disadvantage of such approaches is the cost of employing dedicated sensors for data acquisition. In addition, these approaches are less user-friendly as they are likely to cause inconvenience to the users. The possibility of overcoming these drawbacks by utilizing the built-in sensors in smartphones motivated us to perform the analysis of smartphone sensor data to extract information regarding human behavioral characteristics. The analysis of human behavior has been proven to be effective in various applications including biometricbased user authentication, smart spaces, human-machine interactions, physical activity recognition and surveillance. The key advantage is that the human behavior can be captured unobtrusively without requiring a conscious effort on the part of the user. Therefore, the human behavior-based intelligence for smartphones is quite promising.

The prime objective of this thesis is to enhance the capabilities of smartphone-based biometric recognition and smartphone-based health monitoring systems through the analysis of human behavioral information acquired from the smartphone's built-in sensors. Specifically, we analyze the behavioral information with the objective to develop efficient approaches for biometric authentication, gender recognition and physical activity recognition. Nowadays, it is quite common to see people performing banking transactions and storing sensitive information on smartphones. Therefore, it is extremely important that these devices are able to perform user authentication. In this thesis, we propose an approach for user authentication in smartphones using behavioral biometrics. This approach analyses the behavioral data, which is collected while the user performs different gestures during his/her interaction with the device. In addition to the touch point locations, the proposed approach utilizes the information from the built-in accelerometer sensor and orientation sensor in the smartphone. The modified Hausdorff distance (MHD) is employed for matching features of the gestures performed by the user.

The biometric authentication performance can be improved by supplementing traditional biometric information with soft biometrics like gender, age, height, weight, and ethnicity. Such soft biometric attributes can also be exploited in various applications including surveillance, human-machine interactions, and smart spaces. In this thesis, we perform gender recognition while a user interacts with the device as well as when a user walks with a smartphone in the trouser pocket. Our first approach performs gender recognition by capturing the behavioral information while the user interacts with the smartphone's touchscreen. The behavioral data comprising readings from the accelerometer sensor, gyroscope sensor and orientation sensor are acquired during the user's interaction with the device. Two-dimensional attribute maps are then formed using the set of attributes. GIST descriptors computed on these images provide the discriminatory information for gender recognition. Our second approach for gender recognition utilizes gait information collected using built-in sensors in the smartphone. Specifically, readings from the accelerometer sensor and gyroscope sensor are captured while the user walks with the smartphone in the trouser pocket. We propose a histogram of gradient-based approach to extract features useful for gender recognition.

Smartphone-based activity recognition has attracted a lot of attention as it provides information about daily physical activities performed by an individual and consequently, helps improve the health monitoring applications. Excessive sitting and lack of adequate levels of physical activity are associated with health problems such as obesity, diabetes, cardiovascular disease, poor metabolic health and depression, leading to the increased risk of mortality. In this thesis, we present an approach that utilizes readings from the built-in sensors in the smartphone to recognize various physical activities performed by the user. Accelerometer and gyroscope sensor signals are analyzed to identify the activity performed by the user. We propose a descriptor-based approach to compute the discriminatory characteristics for activity recognition.

In summary, the results presented in this thesis clearly suggest that the data acquired from the built-in sensors of a smartphone carries information useful for analysis of the user's behavior. Our experimental results show that the approaches proposed in this thesis achieve state-of-the-art performances for gesture-based biometric authentication, gesture-based gender recognition, gait-based gender recognition, and physical activity recognition.

Number of research papers published/ accepted in refereed journals		
Number of research papers published/ accepted in refereed conference proceedings	02	
Number of book chapters		
Number of Patent, etc.	NIL	

List of Publications/Conference Proceedings/Patents/Book Chapters

(A) Following are the list of publications to be included in thesis: Journals: (03)

- A. Jain and V. Kanhangad, "Exploring orientation and accelerometer sensor data for personal authentication in smartphones using touchscreen gestures," *Pattern Recognition Letters*, vol. 68, no. 2, pp. 351-360, December 2015. (Impact Factor: 1.952)
- A. Jain and V. Kanhangad, "Gender classification in smartphones using gait information," *Expert Systems with Applications*, vol. 93, pp. 257-266, 2018. (Impact Factor: 3.768)
- **3.** A. Jain and V. Kanhangad, "Human activity classification in smartphones using accelerometer and gyroscope sensors," *IEEE Sensors Journal*, vol. 18, no. 3, pp. 1169-1177, 2018. (**Impact Factor: 2.512**)

Conference Proceedings: (02)

- A. Jain and V. Kanhangad, "Investigating Gender Recognition in Smartphones using Accelerometer and Gyroscope Sensor Readings," IEEE International Conference on Computational Techniques in Information and Communication Technologies, 2016.
- 2. A. Jain and V. Kanhangad, "Human Activity Classification in Smartphones using Shape Descriptors," 24th National Conference on Communications (NCC), 2018. (Presented)

(B) Other Publications during PhD (NIL)

Signature of Candidate

Signature of Thesis Supervisor

TABLE OF CONTENTS

Li	ist of Figures	xi
Li	ist of Tables	ΧV
Ac	cronymsxv	vii
1.	Introduction	1
	1.1 Smartphone sensors	3
	1.1.1 Accelerometer sensor	4
	1.1.2 Gyroscope sensor	4
	1.1.3 Orientation sensor	5
	1.2 Related work and motivation	7
	1.2.1 Biometric-based user authentication	9
	1.2.2 Gender recognition	13
	1.2.2.1 Gender recognition using touchscreen gestures	15
	1.2.2.2 Gender classification using gait information 1	18
	1.2.3 Human activity recognition	19
	1.3 Performance measure	22
	1.3.1 Receiver operating characteristic	22
	1.3.2 k-fold cross-validation	22
	1.4 Challenges and objectives	23
	1.5 Thesis contributions	24
	1.6 Organization of the thesis	25
	2. Personal Authentication in Smartphones using Touchscree	en
	Gestures	27
	2.1 Proposed approach	28
	2.1.1 Classification of gestures	29
	2.1.2 Feature extraction	30
	2.1.3 Score computation and normalization	31
	2.1.4 Fusion of scores	34

2.2	Expe	erimental results and discussion	34
	2.2.1	Dataset-I	35
	2.2.2	Performance of different score normalization and fusion techniques	35
	2.2.3	Performance of score normalization techniques with sum rule	for
	fus	ion	39
	2.2.4	Performance of fusion rules with tanh-estimator for score normalizat	ion
	•••		40
	2.2.5	Performance of orientation, accelerometer versus the rest	46
	2.2.6	Dataset-II	48
2.3	Sum	mary	48
3.	Gend	er Recognition in Smartphones using Touchscreen Gestu	res
	•••••	••••••	51
3.1	Prop	osed approach	52
	3.1.1	Data acquisition and classification of gestures	53
	3.1.2	Feature extraction	53
	3.1.3	Feature selection	56
	3.1.4	Information fusion	56
	3.1.5	Classifiers	57
3.2	Expe	erimental results and discussion	57
	3.2.1	Data acquisition	57
	3.2.2	Evaluation on Dataset-I	57
	3.2.3	Evaluation on Dataset-II	61
3.3	Sum	mary	64
	~		
4.	Gend	ler Classification in Smartphones using Gait Informati	on (5
	•••••		05
4.1	Prop	osed method	66
	4.1.1	Data acquisition	66
	4.1.2	Preprocessing	67
	4.1	.2.1 Normalization	67
	4.1	.2.2 Gait cycle extraction	67

	4.1.3	Featu	are extraction	69
	4.1	1.3.1	Computation of gradients	69
	4.1	1.3.2	Histogram binning	69
	4.1	1.3.3	Block normalization	
	4.1.4	Featu	are level combination	
	4.1.5	Class	sification	
	4.2 Exp	erimer	ntal results and discussion	72
	4.2.1	Data	acquisition	72
	4.2.2	Expe	eriments with S-II device	72
	4.2.3	Expe	eriments with Note-II device	75
	4.2.4	Effec	ct of number of gait cycles on the performance	
	4.2.5	Com	parison with existing methods	
	4.2.6	Sum	mary of experimental results	79
	4.3 Sum	mary		80
5	II	A	uiter Classification in Smorthband using A	
5.	Human and Gyr	Activ	vity Classification in Smartphones using A ne Sensors	ccelerometer 81
5.	Human and Gyr	Activ cosco	vity Classification in Smartphones using A pe Sensors	ccelerometer
5.	Human and Gyr 5.1 Propos	Active cosco sed ap	vity Classification in Smartphones using A pe Sensors proach	ccelerometer
5.	Human and Gyr 5.1 Propos 5.1.1	Active cosco sed ap Input	vity Classification in Smartphones using A pe Sensors proach t signal	ccelerometer
5.	Human and Gyr 5.1 Propos 5.1.1 5.1.2 5.1.2	Activ cosco sed ap Input Extra	vity Classification in Smartphones using A pe Sensors proach t signal action of additional signals	ccelerometer 81
5.	Human and Gyr 5.1 Propos 5.1.1 5.1.2 5.1.3	Active cosco sed ap Input Extra Featu	vity Classification in Smartphones using A pe Sensors proach t signal action of additional signals ure extraction	ccelerometer 81 82 82 82 83 83 84 84
5.	Human and Gyr 5.1 Propos 5.1.1 5.1.2 5.1.3 5.1	Active cosco sed ap Input Extra Featu 1.3.1 F	vity Classification in Smartphones using A pe Sensors proach t signal action of additional signals ure extraction Feature set-I (FSI)	ccelerometer 81 82 82 83 83 84 84 84
5.	Human and Gyr 5.1 Propos 5.1.1 5.1.2 5.1.3 5.1 5.1 5.1	Active cosco sed ap Input Extra Featu 1.3.1 F 1.3.2 F	vity Classification in Smartphones using A pe Sensors proach t signal action of additional signals ure extraction Feature set-I (FSI) Feature set-II (FSII)	ccelerometer 81 82 82 83 83 84 84 84 84 85 87
5.	Human and Gyr 5.1 Propos 5.1.1 5.1.2 5.1.3 5.1 5.1 5.1.4 5.1.5	Active cosco sed ap Input Extra Featu 1.3.1 F 1.3.2 F Infor	vity Classification in Smartphones using A pe Sensors proach t signal action of additional signals action of additional signals re extraction Feature set-I (FSI) Feature set-II (FSII) feature set-II (FSII)	ccelerometer 81 82 82 83 83 84 84 84 84 85 87 87
5.	Human and Gyr 5.1 Propos 5.1.1 5.1.2 5.1.3 5.1 5.1 5.1.4 5.1.5	Active cosco sed ap Input Extra Featu 1.3.1 F 1.3.2 F Infor Class	vity Classification in Smartphones using A pe Sensors proach	ccelerometer 81 82 82 83 83 84 84 84 84 85 87 88 88 88 88 88 88
5.	Human and Gyr 5.1 Propose 5.1.1 5.1.2 5.1.3 5.1.3 5.1 5.1.4 5.1.5 5.2 Exper-	Active cosco sed ap- Input Extra Featu 1.3.1 F 1.3.2 F Infor Class imenta	vity Classification in Smartphones using A pe Sensors	ccelerometer 81 82 82 83 83 84 84 84 84 85 87 87 88 88 88
5.	Human and Gyr 5.1 Propose 5.1.1 5.1.2 5.1.3 5.1.3 5.1.4 5.1.5 5.2 Exper- 5.2.1 5.2.1	Active cosco sed ap- Input Extra Featu 1.3.1 F I.3.2 F Infor Class imenta Evalu	vity Classification in Smartphones using A pe Sensors proach t signal action of additional signals action of additional signals ure extraction Feature set-I (FSI) Feature set-II (FSII) rmation fusion sifiers al results and discussion uation on Dataset-I	ccelerometer 81 82 82 83 83 84 84 84 84 85 87 87 88 88 88 88 88
5.	Human and Gyr 5.1 Propose 5.1.1 5.1.2 5.1.3 5.1 5.1.4 5.1.5 5.2 Exper- 5.2.1 5.2.2	Active cosco sed ap- Input Extra Featu 1.3.1 F 1.3.2 F Infor Class imenta Evalu	vity Classification in Smartphones using A pe Sensors	ccelerometer 81 82 82 83 83 84 84 84 85 87 87 88 88 88 88 88 88 88 88

6.	Conclusions & Future Research Directions	99
	6.1 Conclusions	99
	6.2 Future research directions 1	.02
Re	eferences1	05

LIST OF FIGURES

1.1	The number of smartphone users worldwide from 2014 to 2020 1
1.2	A typical smartphone board with sensors 3
1.3	x, y and z directions for accelerometer sensor defined relative to the
	smartphone
1.4	x, y and z directions for gyroscope sensor defined relative to the smartphone 5
1.5	Azimuth, pitch and roll for orientation sensor defined relative to the
	smartphone
	-
2.1	Overview of the proposed approach for user authentication
2.2	Orientation along roll of a user at two time instances
2.3	Orientation along roll of another user at two time instances
2.4	ROCs for score level combination of gestures with min rule and different
	normalization techniques
2.5	ROCs for score level combination of gestures with max rule and different
	normalization techniques
2.6	ROCs for score level combination of gestures with product rule and different
	normalization techniques
2.7	ROCs for score level combination of gestures with sum rule and different
	normalization techniques
2.8	ROCs for score level combination of gestures with tanh-estimator for score
	normalization 41
2.9	Distribution of genuine and impostor scores for orientation feature 43
2.10	ROCs of individual features with DTW and the proposed MHD based
	matching 43
2.11	ROCs for score level combination (sum rule) of all gestures 45
2.12	Performance comparison of orientation feature with the sum rule based score
	level combination of rest of the features

2.13	Performance comparison of combination of orientation and accelerometer features
	with combination of rest of the features
3.1	Block diagram of the proposed approach for gender recognition
3.2	Samples of swipes (a) R2L (b) SU performed by 10 male and 10 female users (the
	images are enhanced for better visibility) 55
3.3	Comparison of individual performances of the GIST features computed from the
	attribute maps for all the gestures
3.4	Performance of gender classification for different combinations of gestures 60
3.5	Comparison of individual performances of the GIST features computed from the
	attribute maps for all the gestures
3.6	Performance of gender classification for different combinations of gestures 63
4.1	Block diagram of proposed approach for gender recognition
4.2	Accelerometer sensor signals in x , y and z directions
4.3	Gyroscope sensor signals in x , y and z directions
4.4	(a) Local minima points of accelerometer signal in z-direction and (b) the
	extracted gait cycles
4.5	Overview of the proposed feature extraction method 69
4.6	Generation of HG descriptor
4.7	Scatter diagram of the top three features determined using Fisher score based
	feature selection method 75
4.8	Scatter diagram for the top three features determined using Fisher score based
	feature selection method
4.9	Classification accuracy versus the number of gait cycles employed for gender
	recognition
5.1	Block diagram of the proposed approach for activity recognition
5.2	Acceleration signals in x, y and z directions when a user performs different
	activities
5.3	Overview of computation of the HG based feature

5.4	Comparison of the centroid distance signature based FDs of different
	activities
5.5	Performance comparison of the classifiers on Dataset-I using FSI, FSII and their
	feature level fusion
5.6	Comparison of time and frequency domain signals with combined
	performance
5.7	Performance comparison of the classifiers on Dataset-II using FSI, FSII and their
	feature level fusion
5.8	Comparison of time and frequency domain signals with their combined
	performance

LIST OF TABLES

1.1	A summary and comparison of the existing methods for gender classification in
	mobile phones using behavioral biometrics 16
2.1	EERs (%) obtained with different score normalization and fusion techniques 36
2.2	EERs (%) for gesture matching with sum rule based fusion 40
2.3	EERs (%) for gesture matching with tanh-estimator for score normalization 41
2.4	EER (%) of individual features
2.5	EER (%) of individual gestures and their combination 45
2.6	Time (ms) required for matching individual gestures 46
2.7	EER (%) of individual features for dataset-II
3.1	Comparison of feature level and score level fusion of information from the
	attribute maps for all the gestures
3.2	Comparison of the proposed approach and existing work for gender
	recognition
3.3	Comparison of feature level and score level fusion of information from the
	attribute maps for all the gestures
3.4	Comparison of the proposed approach and existing work for gender
	recognition
4.1	Comparison of individual performance of gait information collected from
	accelerometer and gyroscope sensors with the combined performance for gender
	classification
4.2	Cross-speed gender classification accuracy on gait dataset collected using S-II
	device
4.3	Comparison of individual performance of gait information collected from
	accelerometer and gyroscope sensors with the combined performance for gender
	identification

4.4	Cross-speed gender classification accuracy on gait dataset collected using Note-II
	device
4.5	Classification accuracy of the proposed and existing approaches for different
	walking speeds
5.1	Activity classification accuracy with feature level and score level fusion 89
5.2	Classification accuracy using feature sets FSI, FSII and their combination 91
5.3	Confusion matrix of the approach presented in [148]92
5.4	Confusion matrix of the approach presented in [132] 92
5.5	Confusion matrix of the proposed approach
5.6	Performance measures of the proposed and the existing approaches
5.7	Classification accuracy using feature sets FSI, FSII and their combination 94
5.8	Confusion matrix of the approach presented in [149]96
5.9	Confusion matrix of the proposed approach
5.10	Performance measures of the proposed and the existing approaches

ACRONYMS

1D	one-dimensional
2D	two-dimensional
3D	three-dimensional
ACC	accuracy
CNN	convolutional neural network
CPU	central processing unit
DFT	discrete Fourier transform
DTW	dynamic time warping
EER	equal error rate
FAR	false acceptance rate
FD	Fourier descriptor
FRR	false rejection rate
FSI	feature set-I
FSII	feature set-II
GAR	genuine acceptance rate
GPS	global positioning system
GPU	graphics processing unit
HG	histogram of gradient
HOG	histogram of oriented gradients
IB	instance based learning
k-NN	k-nearest neighbor
L2R	left to right swipe
LTE	long-Term Evolution
LOOM	leave-one-out method
MHD	modified Hausdorff distance
N-II	Samsung Galaxy Note-II N7100
NN	neural network
PSO	particle swarm optimization
R2L	right to left swipe

RAM	random access memory
ROC	receiver operating characteristic
S-II	Samsung Galaxy S-II GT-I9100
SD	scroll down
SEN	sensitivity
SoC	system on chip
SPF	specificity
ST	single tap
SU	scroll up
SVM	support vector machine
WEKA	Waikato environment for knowledge analysis
ZI	zoom in
ZO	zoom out

Chapter 1

Introduction

All over the world, mobile phone use has increased significantly over the past decade, mainly due to the convenience it provides to the users, thereby making their daily lives easier. Earlier, mobile phone usage was largely limited to making or receiving phone calls. However, with the advancements in technology, mobile phones have been transformed into *smartphones* that are feature-rich, affordable and one of the widely used electronic gadgets. The advancements in mobile technology have also led to more number of customers embracing smartphones as compared to personal computers. Additionally, smartphones nowadays have become so user friendly that people who are not tech-savvy can also use it without difficulty. It is reported [1] that the factors such as lower price and increased capabilities have resulted in more number of smartphone users worldwide these days as compared to personal computers. According to a survey [2], the result of which is shown in Fig. 1.1, the number of smartphone users worldwide has been increasing significantly every year. In the year 2016, 62.9% of the world population had a mobile phone. Mobile phone ownership is projected to increase further to 67% by the year 2019.



Figure 1.1: The number of smartphone users worldwide from 2014 to 2020 [2].

Nowadays, the use of smartphones and tablets has increased to such an extent that these devices have become indispensable part of our lives. These devices are quite commonly used for performing banking transactions, storing sensitive data, reading emails and news, analyzing daily activities, etc. These devices also help the user to socialize and share data with other people. In addition, these devices can also be used as remote control, recorder, translator and navigator and more importantly, they can also be used for online shopping, which often prompts the user to store banking information on the smartphone. The increased dependency and extensive use of smartphones call for enhanced data security to safeguard user's data stored in these devices. If a smartphone falls into the wrong hands, the consequences can be as serious as a stranger accessing the user's bank accounts and other confidential information such as passwords, emails, personal details and business documents.

The increasing popularity and use of smartphones have motivated the researchers worldwide to explore the information acquired from the smartphone for various tasks such as biometric authentication, access control, surveillance, navigation and health monitoring. The smartphones have become very sophisticated containing several built-in sensors such as motion sensors, position sensors, camera, proximity sensor, pressure sensor, and microphone, as presented in Fig. 1.2 [3]. Smartphones' capabilities are further enhanced by system on a chip (SoC), which contains the central processing unit (CPU), graphics processing unit (GPU), display processor, video processor, etc. With advances in the nanometer technology, the storage capacity of the memory (e.g. random access memory (RAM)) and the internal storage has also increased. In addition, the low power memory employed in smartphones leads to a reduction in the battery consumption. With the incorporation of long-term evolution (LTE) modems, data sharing using advanced smartphones is expected to get easier. The aforementioned advancements and subsequent generation of huge amount of smartphone data has motivated researchers working in diverse fields to acquire and analyze the data in an attempt to derive useful information.



Figure 1.2: A typical smartphone board with sensors [3].

1.1 Smartphone sensors

In the literature, it has been observed that readings from dedicated body-worn inertial sensors carry information useful for various applications related to biometric recognition and health monitoring systems. However, these dedicated inertial sensors result in additional cost. Additionally, these approaches may cause inconvenience to the user and the data collection cannot be performed inconspicuously. These drawbacks inspired us to utilize the readings from built-in sensors of the smartphone to extract information related to the user's behavior.

Equipping smartphones with intelligence has been a topic of interest to researchers working in diverse fields such as biometrics, healthcare, and financial services. Typically, smartphones are equipped with sensors such as motion sensor, position sensor and environmental sensor, which provide readings with high precision and accuracy [4]. These sensors help determine the movement and position of the device, and also monitor environmental changes near the device [5]. Motion sensors, which include accelerometer sensor, gyroscope sensor, gravity sensor and rotation vector sensor, provide acceleration and rotational forces in three directions. Position sensors, on the other hand, provide information about the physical position of the device with respect to the world's frame of reference. They comprise orientation sensor, proximity sensor and magnetometer sensor. Environmental sensors such as barometer, temperature sensor, pressure sensor and light

sensor measure different environmental properties namely, humidity, ambient temperature, ambient pressure and illuminance [4].

In this thesis, we analyze the data acquired from the built-in sensors in the smartphone for biometric-based user authentication, gender recognition and physical activity recognition. Specifically, readings from motion sensors namely, accelerometer sensor and gyroscope sensor, and position sensor namely, orientation sensor are investigated for the aforementioned problems. An Android application is developed to collect the tri-axial readings from these sensors by using a sensor event, which provides sensor readings and the timestamp for the event when there is a change in sensor value (onSensorChanged) [4]. While acquiring the sensor readings, the delay is set to its least value using the command SENSOR_DELAY_FASTEST. The coordinate system for sensor measurements is defined with respect to the screen of the device when the device is held in the portrait mode, as shown in Fig. 1.3. The following subsections present details of the sensors employed in our work [4].

1.1.1 Accelerometer sensor

The accelerometer sensor measures the amount of linear movement applied by the user on the device. Specifically, it provides the linear acceleration (in m/s^2) in x, y and zdirections at each time instant. Fig. 1.3 shows the x, y and z directions defined relative to the smartphone. Most smartphones nowadays come equipped with the accelerometer sensor, which is used to detect shake, tilt and swing of the device.

1.1.2 Gyroscope sensor

Gyroscope sensor measures the angular speed (in rad/sec) applied on the phone along x, y and z directions. Fig. 1.4 shows coordinate system defined relative to the smartphone. The positive rotation is measured in the counter-clockwise direction. This sensor is commonly used to detect rotation such as spin, turn and to watch 360^{0} panoramic view. As the user moves the phone, gyroscope sensor senses the motion and the panoramic view is changed accordingly.



Figure 1.3: x, y and z directions for the accelerometer sensor defined relative to the smartphone.



Figure 1.4: x, y and z directions for the gyroscope sensor defined relative to the smartphone.

1.1.3 Orientation sensor

Orientation sensor records the position of the device with respect to the world's frame of reference. At each event time, it provides orientation values in the following three directions (as shown in Fig. 1.5):

Azimuth: It indicates the rotation around z-axis. Specifically, it is the angle between magnetic north and device's y-axis. Azimuth is 0^0 when the top of the device faces towards north and 180^0 when its top faces towards south.

Pitch: It indicates the rotation around x-axis. It is the angle between the screen of the device and the ground. If the device is held parallel to the ground with its top edge away from the user, it will be positive when the top edge of the device moves towards the ground and will be negative if it moves away from the ground. The pitch values lie between -180° and 180° .

Roll: It indicates the rotation around *y*-axis. It is the angle between the plane perpendicular to the screen of the device and the plane perpendicular to the ground. If the device is held parallel to the ground with its top edge away from the user, the roll value will be positive if the user rotates the right edge of the device away from the ground and will be negative if he moves the right edge towards the ground. The range of roll values is from -90° to 90° .



Figure 1.5: Azimuth, pitch and roll for the orientation sensor defined relative to the smartphone.

In addition to the inertial and position sensor information, there are various behavioral characteristics that can be collected while a user performs different gestures during his/her interaction with the device. An Android application is developed to acquire the

readings when the user performs a touch event such as press, release, and different touch gestures on the touchscreen of the smartphone (OnTouchListener) [4]. The following readings are acquired for smartphone-based user authentication and gender recognition using touchscreen gestures:

1) Event time: It provides the time instant at which the touch event occurs.

2) x - y coordinates: It indicates the position of finger on touchscreen. Specifically, when user performs a gesture, this provides the x - y coordinate values of finger position at each time event.

3) Finger area: This provides an approximation of area of the finger touched on the screen. It returns a scalar value at a given time event. The actual value of touch area in pixels is normalized for the device's explicit range and scaled to a value between 0 and 1.

4) Pressure: This provides the pressure exerted by the user on the touchscreen of the smartphone. The pressure information generally ranges between 0 and 1.

5) Pointer count: This indicates the information on the number of fingers touching on the screen at each time event. This information is employed to differentiate between zoom and other gestures.

1.2 Related work and motivation

In the past decade, readings from dedicated body-worn inertial sensors have been shown to carry information useful for various tasks including human gait recognition [6], sports activity recognition [7], fall detection [8], and health monitoring [9]. In general, the major disadvantage of such approaches is the cost of employing the dedicated sensors for data acquisition. In addition, these approaches are less user-friendly as they are likely to cause inconvenience to the users. The possibility of overcoming these drawbacks by utilizing the built-in sensors in smartphones motivated us to perform the analysis of smartphone sensor data to extract information regarding human behavioral characteristics.

The analysis of human behavior has been proven to be effective in various applications including biometric-based user authentication [10], smart spaces [11], human-machine interactions [12], and surveillance [13]; therefore, the human behavior-

based intelligence for smartphones is quite promising. It has been shown that the signature verification performance can be improved by integrating the human behavioral characteristics such as total duration, number of pen-ups and pen-downs with the handwritten signature [10]. Chen et al. [11] developed a smart system for home energy management based on energy usage patterns of the users. Mondal et al. [12] utilized keystroke information captured during the user's interaction with the computer. Liu et al. [13] demonstrated that inclusion of gait information leads to considerable improvement in the performance of person re-identification. Further, Lu et al. [14] proposed an approach for user authentication that captures in-air-handwriting using wearable inertial sensors.

Human behavior-based intelligence for smartphones using the built-in sensors has also been explored in the literature. Burda [15] and Jin et al. [16] performed user authentication based on the manner in which a user picks up the smartphone from a table or the trouser pocket by using the readings from the built-in accelerometer sensor. In a similar way, Kunnathu [17] performed user authentication by analyzing the accelerometer sensor readings collected when the user picks up the phone and holds it to the ear. On the other hand, Feng et al. [18] acquired motion trajectories for user authentication by using the built-in accelerometer and gyroscope sensors in the smartphone, while the user picks up the phone. Conti et al. [19] captured accelerometer and orientation sensor information while the user answers or makes a phone call. Findling et al. [20] utilized accelerometer sensor readings to unlock a mobile phone without any screen-based interaction. In their approach, an unlocked wrist-watch and a mobile phone are shaken together, and the mobile phone is unlocked when this event is detected in the accelerometer sensor readings. Zhu et al. [21] performed user authentication using readings acquired from the built-in accelerometer and gyroscope sensors, while the user shakes the device to unlock. Hong et al. [22] acquired the accelerometer sensor data for user authentication, while he/she moves the device to form the specific shape in the air. Ryu et al. [23] proposed an approach for user authentication when the verification code is received on smartphone or smart watch. The authors utilized behavioral pattern and environmental information while the user checks the smartphone or smart watch for verification code. The behavioral information was collected using accelerometer and

gyroscope sensors and the environmental information was collected using global positioning system (GPS), wireless access point, Bluetooth, and device model. Li et al. [24] investigated behavioral patterns collected using built-in accelerometer, gyroscope and magnetometer sensors of the smartphone for continuous user authentication. The authors extracted time and frequency domain features from the sensor readings. User authentication in smartphones using gait information extracted from the accelerometer sensor readings was explored in [25, 26]. Primo et al. [27] collected the gait data in different scenarios, including when the device is kept inside the trouser pockets on both sides as well as when the device is held in each hand. As can be observed from the above review of literature, human behavioral characteristics are useful for various tasks. In addition, the key advantage is that the human behavior can be captured unobtrusively without requiring a conscious effort on the part of the user. This thesis deals with the analysis of human behavior for smartphone-based user authentication, gender recognition, and physical activity recognition. Specifically, we analyze the data acquired from the built-in sensors in the smartphone with the objective to develop efficient approaches for the above-mentioned problems.

1.2.1 Biometric-based user authentication

Nowadays, smartphones are commonly used to perform banking transactions and access confidential information such as corporate data, email and social media accounts. Also, since they are often used as a means for storing important and sensitive information, the security of data stored in smartphones is a major concern. Therefore, it is imperative that these devices are able to perform user authentication, even more so, in the event of lost or stolen phones in order to prevent access to user's data by impostors. Traditionally, personal computers have been secured by password-based authentication, which has the inherent limitation that the passwords can be stolen or forgotten. Behavioral biometrics such as keystroke and mouse dynamics provide a more secure and reliable alternative to password-based authentication in computers [29]. Keystroke dynamics performs authentication based on the way the user operates the keyboard [30], while mouse dynamics [31, 32] is an emerging behavioral biometric trait that performs authentication in computers.

can be achieved in two ways; static and continuous authentication. Static authentication is a one-time user verification performed at the time of unlocking the computer [30]. The major limitation of this approach is that once the system is unlocked, it does not continue to provide any security against unauthorized access. In other words, an unauthorized person can operate the system as it has already been unlocked. On the other hand, continuous authentication [33] overcomes this problem by constantly authenticating the user and thereby, preventing unauthorized access to the system by an impostor.

In recent years, with the rapid increase in the use of smartphones and subsequent increase in the security issues associated with it, researchers have focused their effort on developing methodologies for static as well as continuous user authentication. Static authentication in smartphones is achieved using traditional or biometric authentication techniques. Traditional techniques include swipe or click to unlock, creating a specific pattern on the screen to unlock and password-based authentication, whereas commonly employed biometric techniques for static authentication include face [34-36], voice [37-39], fingerprint [40, 41] and keystroke [42-45] based authentication. In addition to this, researchers have explored other physiological biometric traits such as finger knuckle [46], iris [47] and palmprint [48-50]. Multimodal biometric solutions [51-57] have also been proposed for mobile user authentication. However, it may be noted that these biometrics traits are more suited for providing static authentication in mobile phones. A detailed survey of various biometric approaches proposed for user authentication in mobile phones can be found in [58-61].

Keystroke based authentication in computers inspired researchers to explore similar approaches for touchscreen devices. Authors in [62] performed user authentication using the information acquired from a touchpad, which records finger pressure, finger position, hold time and inter-key time. The approach achieved 1% equal error rate (EER) when only the finger pressure information utilized. However, their approach was evaluated on a dataset of only ten users. Luca et al. [63] presented an approach for behavioral biometric based static authentication when the user performs a password pattern on the touchscreen. As discussed earlier, static user authentication has inherent limitations and therefore, continuous authentication approaches are more desirable for securing smartphones. Most

of the existing approaches for continuous authentication are based on behavioral characteristics acquired during the user's continuous interaction with the device.

Frank et al. [64] proposed an approach for continuous user authentication using 30 behavioral features extracted from the touchscreen input. Authors explored two classifiers namely, k-nearest neighbors (k-NN) and support vector machine (SVM) and reported EER between 0% and 4% for different experimental scenarios. However, the authors considered only up-down and left-right scrolling in their study. Feng et al. [65] developed a glove sensor that captures linear and angular acceleration of the finger movement. By combining the information acquired from the touchscreen and the sensor glove, their approach yielded false acceptance rate (FAR) of 4.66% and false reject rate (FRR) of 0.13%. However, the use of additional sensor to improve the accuracy is a major drawback. In another work, Feng et al. [66] proposed to maintain separate templates for each application. Their approach was evaluated on a dataset of only 23 users. In addition, the increased memory requirements due to separate templates might limit its applicability in mobile phones. Authors in [67] presented an approach that uses 21 features, which are fed into a neural network classifier. They also employed particle swarm optimization (PSO) to optimize the neural network and reported an improved EER of 2.92%. Authors in [68] acquired 22 multi-touch gestures such as drag, swipe, pinch and user-defined gestures on an iPad. They employed dynamic time warping (DTW) for matching and achieved an average EER of 7.88%. Antal et al. [69] gathered data from 71 users for horizontal and vertical scrolling using eight different tablets and mobile phones. The authors explored k-NN and random forest algorithm for classification and reported more than 95% accuracy. However, the authors considered only horizontal and vertical scrolling in their work. Zhao et al. [70, 71] proposed a novel method for user authentication, in which the trace of points (swipe gesture) is converted into an image. They represented the trace movement and pressure as shape and intensity values in a twodimensional (2D) image. On a dataset of 30 users, their approach achieved 2.62% EER when six gestures [70] were combined. Fierrez et al. [72] evaluated a behavioral pattern based approach on four public databases of touch gesture data collected while the users performed horizontal and vertical swipes. Mahbub et al. [73] computed a total of 24 features, including stroke duration, mid-stroke pressure and mean resultant length from

the swipe gesture data. Their method achieved EER ranging from 22% to 38% using different classifiers on a dataset of 48 subjects. Kumar et al. [74] investigated the fusion of swiping gestures, typing patterns and phone movement patterns collected from 28 users. Their method achieved 93.33% authentication accuracy for the feature level fusion of swiping and phone movement patterns, and 89.31% authentication accuracy for the score level fusion of typing and phone movement patterns. In [75], the authors demonstrated that the swipe gesture based-continuous authentication using one-class classifier is possible if sufficient training data of the genuine user is available. Sitov et al. [76] collected behavioral data such as hand movement, orientation of the device, tap events and keystroke features, while the user was sitting and walking during his interaction with the device. The authors achieved the best EER of 10.05% on the sitting dataset and 7.16% on the walking dataset, when all the features were combined at score level. Zhang et al. [77] proposed a sparse representation-based approach, which utilized linear and kernelized dictionaries for evaluating swipe gestures. Serwadda et al. [78] collected behavioral data such as pressure, finger area and time event during the user's interaction with the device. They investigated the performance of ten classifiers and achieved the best mean EER of more than 10%. Li et al. [79] utilized WiFi and accelerometer sensor data for user authentication when he or she interacts with the application installed on the smartphone for 3 seconds. In another work, Li et al. [80] collected accelerometer and gyroscope sensor readings for user authentication during the user's interaction with the application installed on the smartphone. Lee et al. [81] proposed an approach for user authentication, which utilized keystroke data such as time, coordinates and finger size in addition to motion sensor readings such as accelerometer, gyroscope and rotation. Authors in [82] investigated deep learning autoencoder, which utilized accelerometer sensor readings acquired during the user's interaction with the device for continuous user authentication. Authors in [83] performed continuous user authentication using behavioral characteristics captured during the user's interaction with the device. The authors computed deep features using Siamese convolutional neural network (CNN) and performed classification was performed using one-class SVM. A detailed survey on touch dynamics-based user authentication in mobile devices can be found in [84].

1.2.2 Gender recognition

In recent years, an increasing number of researchers have focused their efforts on soft biometrics traits such as age, gender and ethnicity of people to enhance the performance of biometric recognition systems. Such soft biometric attributes can be utilized in various applications such as surveillance, biometric authentication, access control, marketing and human-machine interactions. In addition, the gender information can be used in pervasive computing applications. For example, the information can be communicated to devices embedded in the smart space to adjust the intensity and color of the room lighting based on general preferences of the identified gender. Gender information can also be utilized by access control systems in smart spaces, where only a particular gender (male or female) is allowed to enter. Human beings can easily discriminate between male and female by looking at the face, analyzing the style of walking or listening to the speech. However, automated gender identification by computer is still a challenging task [85].

Researchers have extensively studied gender classification using face [86-88] biometrics. The authors in [86] utilized near-infrared and thermal face images for gender classification. They employed histogram of local binary patterns as the feature vector and evaluated the performance using different classifiers such as SVM, k-NN, and Adaboost. Lu et al. [87] identified gender and ethnicity by consolidating the range and intensity information from facial scans. Danisman et al. [88] proposed a fuzzy inference-based gender classification approach, which utilizes hair volume, mustache, and information from a vision-sensor. Additionally, researchers have investigated other physiological biometric traits such as ear [89], fingerprint [90], hand geometry [91] and iris [92, 93] for gender identification.

In the literature, gender recognition using traditional behavioral biometrics such as signature [84], video-based gait [95-97], voice [98, 99] and keystroke [100, 101] has been studied extensively. Other behavioral characteristics [102, 103] have also been investigated for gender recognition. Inspired by the problem of hidden identities in social-networking sites, Peersman et al. [102] proposed an approach to recognize the age and gender based on the user's behavior. This information, for example, can help protect the minors from pedophiles on the prowl in social media. The authors collected data from a Belgian social-networking site for evaluations. The features extracted from the short chat
messages include words, emoticons, and punctuations. Balen et al. [103] performed mouse dynamics based gender recognition and evaluated their approach on a dataset of 94 users. They collected 256 mouse movements from which temporal, spatial and movement accuracy metrics are computed.

A comprehensive review of biometric-based gender recognition is presented in [85] and [104]. Researchers have also investigated multimodal approaches that involve a combination of multiple biometric traits for gender classification [90, 105-107]. Shafey et al. [105] performed a fusion of visual (face) and acoustic information to improve the performance of gender classification. Li et al. [90] proposed a fusion of face and fingerprint for gender recognition. The authors in [106, 107] performed gender classification using face and gait biometric traits. In biometric systems, the authentication accuracy can be enhanced by supplementing traditional biometric traits with soft biometric traits such as gender [108-110]. Jain et al. [108] demonstrated improved performance when soft biometric traits such as gender, height, and ethnicity are incorporated into user authentication that employs face and fingerprint as primary characteristics. Park et al. [109] achieved improvement in performance when soft biometric traits such as gender, ethnicity and facial marks (scars, moles, and freckles) are incorporated into face recognition. Similarly, Idrus et al. [110] demonstrated performance enhancement when soft biometrics such as gender, age, and handedness are combined with the behavioral biometric characteristic namely, keystroke dynamics.

Over the last decade, the use of smartphones has increased rapidly. Currently, smartphones play a significant role in our everyday life. Therefore, researchers have also focused their efforts on developing several applications, including smartphone-based healthcare systems [111-113]. Ogunduyile et al. [111] utilized the built-in sensors in mobile phones for diagnosis, remote monitoring and to provide advice to patients. Ren et al. [112] also utilized the built-in sensors to detect spoofing in mobile healthcare systems. Minutolo et al. [113] proposed an innovative and efficient decision support system to provide remote health monitoring. They also showed the effectiveness of the system in real-time conditions on mobile devices. Incorporating gender information into mobile healthcare systems can enable gender-specific medical advice to the patients, especially in smartphone-based remote medical advice services [114]. Furthermore, gender

information in smartphones can be used to improve user's interaction with the device [115]. Specifically, gender information can be used to present better search results for shopping, themes, and new applications to the smartphone user. Another potential application is the targeted advertisement, in which advertisements can be recommended based on the user's gender [116, 117]. As reported by a study [118], women are more concerned about their safety and use their phones more for safety purposes as compared to men. The existing mobile applications for women safety can also utilize the gender information to improve its safety features.

Researchers have investigated the problem of gender recognition in mobile phones using physiological biometric traits such as face [119] and ocular [120], as well as behavioral biometric traits [121-123]. Eidinger et al. [119] proposed an approach that performs age and gender recognition by analyzing the face images acquired using smartphones. Rattani et al. [120] explored several texture descriptors for ocular biometric-based gender recognition in mobile phones.

Agneessens et al. [121] analyzed audio signals acquired using a smartphone to identify the number of speakers and their gender. The authors in [122] analyzed the mobile phone usage to derive information about the user's gender. Specifically, they analyzed the data relating to mobile phone usage that includes information on the number of incoming and outgoing calls, duration of calls, number of incoming and outgoing messages and the number of unique phone numbers connected with the user. Choi et al. [123] proposed a method that analyzes text messages for predicting mobile phone users' gender.

Table 1.1 shows a summary of the existing methods for mobile phone-based gender recognition using behavioral biometrics. In this table, ACC stands for the classification accuracy.

1.2.2.1 Gender recognition using touchscreen gestures

In the recent past, researchers have investigated gender recognition in mobile phones using keystroke dynamics [124, 125] and touch gestures [115, 125, 126]. The comparison of these methods is presented in Table 1.1. Buriro et al. [124] extracted time-based keystroke features from 4 or 16 digits secret key entered by the user to determine his/her age, gender, and operating hand information. The approach proposed in [125] utilizes

Study	Modality	Methodology	Number of subjects in the dataset	Best performance	Remarks
Agneessens et al. [121]	Speech	Mean of probability distribution function with predefined threshold	1 male and 1 female	ACC: 90%	 (i) Suited for telephone- based applications (ii) Evaluated on a small dataset
Sarraute et al. [122]	Mobile phone usage	A set of features based on incoming and outgoing calls and SMS	284150 males and 215850 females	ACC: 82.8%	 (i) The data should be analyzed for a longer period of time for better performance (ii) It is difficult to predict the gender if there is no incoming or outgoing activity.
Choi et al. [123]	Text data	Based on similarities between the text data and the word-set of each gender	16 males and 16 females	F-score: 0.87	 (i) Requires less computation (ii) Some of the words cannot be classified into either category, i.e., unclassified cases (iii) In some languages, the texts used by the male and female users are subtle
Buriro et al. [124]	Keystroke	A set of timing based keystroke features with random forests classification	45 males and 105 females	ACC: 82.8%	(i) No risk of privacy of users(ii) User needs to remember password
Antal and	Keystroke	A set of 71 features with random forests classification	18 males and 18 females	ACC: 64.76%	(iii) The performance is not very promising
[125]	Swipe gestures	A set of 9 features with random forests classification	38 males and 38 females	ACC: 57.16%	(i) No need to remember password
Antal et al. [126]	Swipe gestures	A set of 15 features with k- NN, random forests and SVM classifiers	9 males and 9 females	ACC: 88%	(ii) Evaluated on a small dataset.(iii) Only horizontal swipes are considered.
Miguel- Hurtado et al. [115]	Swipe gestures	A set of 14 features with multilinear logistic regression classifier	57 males and 59 females	ACC: 78.2%	 (i) Vertical swipes are also considered. (ii) The performance is not very promising, therefore it does not seem a reliable way to classify gender (iii) query sample is the average of each feature across all the samples from the user (iv) zoom-in and zoom-out gestures are not considered

 Table 1.1: A summary and comparison of the existing methods for gender classification in mobile phones using behavioral biometrics

Weiss and Lockhart [133]	Gait	A set of statistical features with NN, IB3 and J48 classifiers	38 males and 28 females	ACC: 71.2%	(i) (ii) (iii) (iv)	Can be performed unobtrusively Data from only the accelerometer sensor is employed Evaluated on a small dataset Variations in walking speed are not considered
Jain and Kanhangad [135]	Gait	Multi-level local pattern-based features with bagging classifier	25 males and 17 females	ACC: 77.45%	(i) (ii) (iii) (iv)	Can be performed unobtrusively Readings from both accelerometer and gyroscope sensors are explored Evaluated on a small dataset Different walking speeds are considered but no cross-speed experimental results

behavioral information extracted from a 14-digit password as well as from the swipe gestures performed by the user while interacting with touchscreen-based mobile phones. To predict the user's gender using keystroke patterns, the authors employed time-based features, pressure and finger area information. On the other hand, they computed features such as velocity, mean of finger area and mean of accelerometer sensor readings for gender classification using touch swipes. Their approach achieved classification accuracy of 64.76% and 57.16% on the keystroke and swipe gesture datasets, respectively. In their work, only horizontal swipe gesture has been explored for gender recognition. Antal et al. [126] computed statistical features namely, pressure, finger area and x - y coordinate values from horizontal swipes. They evaluated the performance using 3-fold crossvalidation on a dataset of 18 subjects. The main limitation of their work is that the performance was evaluated on a small dataset with user overlap between the training and test set. In addition, the authors considered only horizontal swipe in their study. The authors in [115] acquired data relating to both horizontal and vertical swipes from 57 males and 59 females. The performance of their approach was evaluated using 10-fold cross-validation. Since the average of each feature across all samples (for every swipe) of a subject is computed, the issue of user overlap does not arise. The drawback is that their test feature vectors are also computed by averaging each feature across all the samples of a user.

All of the aforementioned existing works utilize pressure information acquired from the smartphone sensor. This is a major drawback as most of the latest smartphones do not sense the pressure exerted by fingers on the touchscreen. The usefulness of information that can be extracted from motion sensors in smartphones remains largely unexplored, except for the approach proposed by Antal et al. [125] that utilized accelerometer sensor readings for touch gesture-based gender identification. Furthermore, the gestures considered in the existing works are limited to horizontal and vertical swipes, and gestures such as zoom-in and zoom-out have never been explored for touch gesture-based gender recognition. Although the existing studies have validated the potential of features extracted from touch gesture data for gender recognition and reported promising results, there is a pressing need to further improve the performance of smartphone-based gender recognition before it can be incorporated into real-world applications.

1.2.2.2 Gender classification using gait information

Vision-based gait biometric traits have also been employed for gender classification. Hu et al. [95] employed shape descriptor for spatial information and periodic shape variations for temporal information. Li et al. [96] developed an algorithm to obtain gait information from the movements of different parts of the silhouette. Lu et al. [97] investigated arbitrary walking directions to recognize the identity and gender. Igual et al. [127] and Kastaniotis et al. [128] performed gender recognition using gait data captured from a depth camera. Igual et al. [127] proposed a fast feature extraction algorithm that utilizes a three-dimensional (3D) point cloud obtained from gait image sequences. Kastaniotis et al. [128] developed a method in which motion of frames is encoded using an angular representation.

Several approaches have also been proposed for user authentication based on gait biometrics captured using wearable sensors [6, 129-131]. Zhang et al. [6] proposed an approach for user authentication by capturing gait information from wearable accelerometer sensors placed at various locations on the human body. Thang et al. [130] employed only the accelerometer sensor readings, while the authors in [131] utilized gait information captured from both the accelerometer and gyroscope sensors for smartphonebased user authentication. Soft biometric traits such as gender, age and height have been determined in [132] using gait signals acquired from four dedicated body-worn sensors placed at different locations on the body.

The comparison of the approaches available in the literature for gait-based gender recognition is presented in Table 1.1. Weiss and Lockhart [133] identified soft biometric traits such as height, weight, and gender using 43 statistical features extracted from the built-in accelerometer sensor readings. They investigated the performance of different classifiers from the Weka data mining tool [134], particularly, instance-based learning (IB3), J48 decision tree and multilayer neural network. However, their study is limited to gait information obtained from only the accelerometer sensor. In a more recent work [135], the combination of information captured by the accelerometer and gyroscope sensors in a smartphone has been investigated for gender recognition. The results presented in [135] suggest that the combination of gait information captured by the two sensors results in improved performance for gender classification. However, the performance of gait-based gender recognition in smartphones needs to be improved significantly, before it can be incorporated into mobile phone-based healthcare services or other real-world applications.

1.2.3 Human activity recognition

Activity recognition, which is an integral task in several health monitoring applications [136, 137], is performed by capturing the contextual information while a user performs different activities. Excessive sitting and lack of adequate levels of physical activity are associated with health problems such as obesity, diabetes, cardiovascular disease, poor metabolic health and depression [137], leading to the increased risk of mortality. According to a study [138], life expectancy can be increased by two years, if individuals in the USA reduce their sedentary time to less than 3 hours per day. Activity recognition can be used for continuous analysis of the daily activities performed by the user. Such an analysis is useful in understanding the behavior and thereby, making it possible to provide automated suggestions for reducing the risk factor for various non-communicable diseases. In addition to healthcare applications, activity recognition is also useful in applications such as smart homes, security and transportation mode detection [139].

The contextual information for activity recognition can be collected by placing sensors such as camera in the environment. This approach, however, is only suited for applications operating in controlled environments [140]. Another approach to activity recognition is using body-worn sensors, which are appropriate for uncontrolled indoor and outdoor environments. In this approach, the information is usually collected from a set of dedicated body-worn motion sensors, which are placed at different body locations such as wrist, chest and ankle. In comparison, this approach is less user-friendly as the body-worn sensors may cause inconvenience. Also, the data collection in this approach cannot be performed inconspicuously. However, the aforementioned drawbacks can be overcome by utilizing built-in inertial sensors of smartphones for activity recognition. Smartphones, over the last decade, have become an indispensable part of our daily lives. Typically, these devices have built-in sensors such as accelerometer, gyroscope, global positioning system (GPS), magnetometer and microphones. For activity recognition, the data from smartphone sensors can be collected by keeping the phone in the user's trouser pocket, while he/she performs daily activities such as sitting, standing, walking. The key advantage of this approach is that the data collection, which can be performed unobtrusively [140], does not require any additional hardware. Therefore, this approach to activity recognition is well suited for continuous analysis of the user's daily activities.

In the literature, activity recognition using sensors placed in the environment and the dedicated body-worn sensors have been extensively studied. A detailed survey on activity recognition using wearable sensors is presented in [141-143]. The authors in [144, 145] explored motion sensors for activity recognition. Specifically, their approaches involve placing wearable sensors at multiple body locations. In [144], the authors collected bi-axial accelerometer data from sensors placed at four different limb locations and a sensor at the right hip. The approach [145] uses two sets of tri-axial accelerometer sensors attached to the left and the right sides of a waist belt. Tapia et al. [146] developed a real-time activity recognition system using five tri-axial accelerometer sensors and a heart rate monitor. In general, the approaches using multiple wearable sensors achieve high accuracy for activity recognition. However, as mentioned previously, these sensors may become to wear, especially for continuous activity recognition in which users may have to wear the sensors for extended periods of time.

Activity recognition using inertial sensors in smartphones has also been investigated in the literature. Bieber et al. [147] developed a mobile phone application that detects daily physical activities using the built-in accelerometer. Kwapisz et al. [148] performed activity recognition using the accelerometer data collected from a smartphone. The subjects kept the phone in the front pocket of their trousers while performing the daily activities such as walking, jogging, ascending stairs, descending stairs, sitting, and standing. Lv et al. [149] performed activity recognition using the accelerometer sensor readings collected from a smartwatch and a mobile phone. Dernbach et al. [150] investigated the usefulness of both accelerometer and gyroscope sensor readings for classification of simple as well as complex activities. The performance of their approach for the classification of complex activities is not very promising. Kwon et al. [151] presented an unsupervised machine learning based method for activity recognition. Recently, Chen et al. [152] employed readings from the accelerometer and gyroscope sensors in a smartphone for classification of five different activities. The authors explored time, frequency and wavelet domain features and Kolmogorov-Smirnov test based dimensionality reduction for activity recognition. Shoaib et al. [153] performed smartphone based activity recognition using hand and leg movement data collected from the pocket and the wrist positions. In addition to the routine physical activities, the authors considered activities such as typing, smoking and eating. A comprehensive survey of inertial sensors based activity recognition in the smartphone can be found in [154, 155].

Anguita et al. [156] considered six activities namely standing, sitting, laying, walking, downstairs and upstairs. A set of 27 additional signals was derived from the tri-axial accelerometer and gyroscope sensor readings. Their approach, which employs statistical features and SVM classifier, achieved 96.33% average classification accuracy on a dataset of 7352 training and 2947 test samples. Wu et al. [157] investigated a set of time and frequency domain features and k-NN classifier for smartphone based activity recognition. Performance evaluation using 10-fold cross-validation yielded 90.2% average classification accuracy on a dataset that contains 2807 sensor readings corresponding to nine different activities. Since the aforementioned works demonstrated the feasibility of recognizing activities using the built-in sensors in smartphones or a

similar device with promising performance on large datasets, there is a pressing need to develop approaches that provide highly accurate and reliable performance. Moreover, enhancing the performance is imperative for the smartphone-based solution to be a superior alternative to the dedicated body-worn sensor based methods in real-world applications.

1.3 Performance measure

The performance measures employed in this thesis for biometrics-based user authentication and classification approaches are discussed below:

1.3.1 Receiver operating characteristic

The evaluation of user authentication method presented in this thesis is performed using receiver operating characteristic (ROC) curve, which is a plot of FAR versus genuine acceptance rate (GAR) [29]. Additionally, the user authentication method is compared using the scalar performance measure namely, equal error rate (EER). EER corresponds to the crossover point of FAR and FRR curves. While FAR indicates the probability of an imposter being mistakenly accepted as a genuine user, FRR indicates the probability of a genuine user falsely rejected as imposter user. FRR can also be computed as 1-GAR. If the query sample belongs to the same user as the template, the user is referred to as genuine user and the score generated by comparing the two samples is denoted as the user is referred to as imposter user and the score generated by comparing the two samples is denoted as imposter user and the score generated by comparing the two samples is denoted as imposter user and the score generated by comparing the two samples is denoted as imposter user and the score generated by comparing the two samples is denoted as imposter user and the score generated by comparing the two samples is denoted as imposter user and the score generated by comparing the two samples is denoted as imposter user and the score generated by comparing the two samples is denoted as imposter user and the score generated by comparing the two samples is denoted as imposter user and the score generated by comparing the two samples is denoted as imposter user and the score generated by comparing the two samples is denoted as imposter user and the score generated by comparing the two samples is denoted as imposter user and the score generated by comparing the two samples is denoted as imposter score.

1.3.2 k-fold cross-validation

The classification approaches employed in this thesis are evaluated using k-fold crossvalidation. In this scheme, the data samples are divided in to k-partitions [158]. Out of them, k-1 partitions are used for training purpose and the remaining one partition is used for testing. This process is repeated k times by considering each of the k-partitions as a testing partition. Finally, value of the performance metric obtained in each of the repetitions is averaged.

1.4 Challenges and objectives

Lately, there has been much research focus on the analysis of information acquired from smartphones for various applications. In this thesis, we focus on three problems namely, biometric authentication, gender recognition and physical activity recognition using smartphone sensor data. In the preceding sections, we have discussed various drawbacks of the solutions, which exist for the aforementioned problems. In addition, most of the existing approaches have been evaluated on small datasets, and the experiments to ascertain their performance on datasets collected using multiple devices have largely been ignored.

In general, information acquired from the built-in sensors in smartphones has not been fully exploited for gesture-based user authentication and gender recognition. The existing methods for gait-based gender recognition and activity recognition utilize dedicated body-worn sensors for signal acquisition. The dedicated body-worn sensors are generally expensive and not suited for various indoor and outdoor scenarios. Since the user needs to wear these sensors while performing the activity, the data cannot be collected unobtrusively. Additionally, it may be too cumbersome to wear these sensors, especially for continuous monitoring. In addition to the above-mentioned specific drawbacks, obtaining highly accurate and reliable performance for approaches that utilize smartphone sensor data remains a major challenge, which needs to be overcome in order for these solutions to be useful for real world applications.

The prime objective of this thesis is to enhance the capabilities of smartphone-based biometric recognition and smartphone-based health monitoring systems through the analysis of human behavioral information acquired from the smartphone's built-in sensors. The specific objectives of this thesis are as follows:

- (1) To investigate the suitability of the information acquired from the built-in sensors of the smartphone for biometric authentication, gender recognition, and physical activity recognition.
- (2) To devise novel approaches that provide highly accurate and reliable performance and are suitable for real-world deployments.
- (3) To evaluate the performances of the proposed approaches on large datasets.

(4) To ascertain the performances of the proposed approaches on datasets collected using different devices.

1.5 Thesis contributions

In this thesis, we aim to develop efficient approaches for smartphone-based biometric authentication, gender recognition, and physical activity recognition based on human behavioral characteristics. Specifically, we analyze the data acquired from the built-in sensors in a smartphone to extract user's behavioral information useful for the abovementioned problems. We also focus on overcoming some of the drawbacks of the existing approaches. The performances of the approaches proposed in this thesis are evaluated on relatively large datasets collected using two different devices. Additionally, we perform performance comparisons with the existing approaches. Major contributions of this thesis can be summarized as follows:

- (i) We propose an approach for user authentication in smartphones using behavioral biometrics. The approach involves analyzing behavioral traits, while the user performs different gestures during his interaction with the device. In addition to the commonly employed features such as x y coordinate information and finger area, the proposed approach investigates the usefulness of information acquired from accelerometer and orientation sensors. The feature set is further enriched with features such as point curvature and curvature of the swipe. The feature matching is performed using the modified Hausdorff distance (MHD).
- (ii) We present an approach for gender recognition in smartphones using touchscreen gestures performed by the user. The proposed work investigates the information extracted from the built-in sensors in a smartphone to identify the user's gender. Specifically, behavioral characteristics reflected in accelerometer sensor readings, gyroscope sensor readings, orientation sensor readings and finger area are captured during the user's interaction with the touchscreen device. These characteristics are further enriched by deriving a secondary set of attributes such as velocity of swipe, point curvature, and length of swipe. We have considered touch gestures such as horizontal swipes, vertical swipes, zoom in and zoom out for gender classification. Additionally, different combinations of these gestures have been explored to improve

the performance of gender recognition. Our approach involves forming twodimensional attribute maps for each of the six gestures using a set of attributes. GIST descriptors are then computed on the attribute maps to obtain a holistic image representation in the form of a low-dimensional feature vector.

- (iii) We propose a novel approach for classification of the gender of a smartphone user using gait signals. The proposed work explores the combination of gait information collected from accelerometer and gyroscope sensors for gender recognition in smartphones. Histogram of gradient (HG) method is proposed to extract features from the gait data. We have considered three variations in speed namely, slow, normal and fast to determine the gender of the user. In addition, we have performed cross-speed experiments to analyze the impact of speed on the performance of gender recognition.
- (iv) We propose a descriptor-based approach for activity classification using built-in sensors in smartphones. In addition to the acquired accelerometer and gyroscope sensor readings, a set of time and frequency domain signals are derived to identify the activities performed by the user. In the proposed approach, two descriptors namely, HG and centroid signature based Fourier descriptor (FD), are employed to extract feature sets from these signals. We have investigated the performance of the feature and score level fusion of these descriptors. Furthermore, we have demonstrated the usefulness of the additional time and frequency domain signals employed in this work.

1.6 Organization of the thesis

The subsequent chapters of this thesis are organized as follows:

Chapter 2 presents the proposed approach for user authentication in smartphones using touch gestures captured during the user's interaction with the device. This chapter also details the Android application developed to collect the touch gesture data employed to achieve the goal. In addition, this chapter discusses different score normalization and score fusion approaches.

Chapter 3 describes our approach for gender recognition in smartphones using behavioral information collected during his or her interaction with the device. This

chapter also discusses the 2D attribute map formed by employing the primary and secondary sets of attributes. It also provides a brief overview of GIST descriptor.

Chapter 4 presents details of the proposed gender classification approach that utilizes the gait information obtained using the built-in accelerometer and gyroscope sensors in smartphones. This chapter also describes the Android application developed to acquire gait signals for gender recognition. Further, it provides a detailed description of the proposed HG descriptor.

Chapter 5 details the proposed descriptor-based approach for activity classification using the accelerometer and gyroscope sensors in smartphones. This chapter discusses the centroid signature based FD in detail and presents descriptions of the datasets employed for performance evaluation.

Chapter 6 presents the conclusions and directions for future research.

Chapter 2

Personal Authentication in Smartphones using Touchscreen Gestures

Nowadays, people often use their mobile phones for performing banking transactions, accessing and storing confidential information such as corporate data, emails and social media accounts. They also store important and sensitive information on their smartphones. Therefore, it is extremely important that these devices are able to perform user authentication. Password-based authentication can be employed to secure the mobile phones, but passwords can be easily stolen or forgotten. Behavioral biometrics can provide a more secure and reliable alternative to password-based authentication in smartphones. In this chapter, we present a touch gesture based user authentication approach in smartphones. The proposed approach utilizes the built-in sensors of the smartphone along with the commonly used features such as x - y coordinate locations and finger area. The major contributions of this work can be summarized as follows:

- We propose a new approach for user authentication in smartphones using modified Hausdorff distance (MHD) for matching.
- (2) This study investigates the usefulness of information acquired from accelerometer sensor and orientation sensor for user authentication.
- (3) Performance evaluation of the proposed approach is performed on a relatively large dataset of 104 users.
- (4) Performance evaluation on a second dataset of 30 users acquired using another smartphone to study the impact of device on the verification performance.

The rest of the chapter is organized as follows: The proposed approach is detailed in Section 2.1, in which firstly, an overview of the approach is provided followed by detailed descriptions of techniques employed for feature extraction, score computation, score normalization and fusion of scores. Section 2.2 presents experimental results and discussion. This section also presents a description of the datasets used for performance evaluation. Finally, Section 2.3 summarizes this chapter.

2.1 Proposed approach

Fig. 2.1 shows a complete overview of the proposed approach. In this approach, an Android application is employed to capture the behavioral data during user's interaction with the device. Specifically, a set of behavioral data is captured at every touch point when the user performs a specific gesture on touchscreen. In this study, we consider the following seven gestures: left to right swipe (L2R), right to left swipe (R2L), scroll up (SU), scroll down (SD), zoom in (ZI), zoom out (ZO) and single tap (ST). The set of data captured by our application for each of the above gestures include x - y coordinates, accelerometer sensor readings, orientation sensor readings and the area covered by the finger on the screen. In the next step, the acquired data is processed to identify the category of the gesture. This is done in order to facilitate the matching of feature templates of the corresponding gestures in the matching stage.



Figure 2.1: Overview of the proposed approach for user authentication.

In addition to the behavioral data, which we also refer to as raw data, captured by our application, two additional features are computed by using the x - y coordinate information. One of the features is the curvature at each touch point of swipe and the other one is the curvature of swipe. The computation of these features is detailed in Section 2.1.2. During the enrolment phase, the behavioral data, along with the two extracted features are stored as feature templates in the database. In this way, the database contains feature templates corresponding to each of the seven gestures performed by every user. During the authentication stage, the Android application running on the

smartphone captures user's behavioral data. In the first step of processing, the type of gesture performed by the user is identified and the two additional features are extracted to form a query feature set. This is followed by matching of the query feature set with the corresponding feature templates from the database, generating multiple matching scores. These matching scores are then normalized and combined using fusion of scores technique to obtain the final score. Finally, the decision stage utilizes the final score to determine whether the user is genuine or impostor.

2.1.1 Classification of gestures

In this work, we considered the following seven most commonly used gestures in touchscreen devices: L2R, R2L, SU, SD, ZI, ZO and ST. L2R and R2L are commonly used for performing tasks such as unlocking the phone, browsing photos and switching between the home pages. SD and SU gestures are often performed while reading a document and browsing internet. ZI and ZO gestures are used for viewing specific content in images and documents, while ST is used to write a message or mail and more generally, to select an option [70]. As discussed in the previous section, the identification of the gesture performed by the user is an important task in the proposed approach for user authentication. For this purpose, we developed a simple and efficient heuristic method based on the x-y coordinate information captured by the application. In this method, L2R, R2L, SU and SD gestures are identified by computing differences between x and y coordinates of the first and last points of the gesture. If the magnitude of the difference of x-coordinates is greater than that of y-coordinates, the gesture is either L2R or R2L. Further, based on the sign of the difference of x-coordinates, the gesture can be classified as L2R or R2L. In a similar way, SU and SD are identified based on the difference of y-coordinates. ZI and ZO gestures are first separated from other gestures using the finger count on touchscreen. Further, the distance between the start point of one finger and the start point of second finger is calculated. Similarly, the distance between the two end points is also calculated. If the distance between the start points is greater than that of the end points, the gesture is ZO, otherwise the gesture is classified as ZI. The identification of ST is trivial as it contains only a single point.

2.1.2 Feature extraction

The feature set in the proposed approach comprises a set of raw behavioral data acquired by the Android application and the two additional features extracted from the x - ycoordinate information. The raw behavioral data comprises of keystroke data namely xy coordinates and finger area in addition to motion sensor readings such as accelerometer sensor readings and orientation sensor readings. The detailed description of these constituents of our feature set is provided in Section 1.1:

Figs. 2.2 and 2.3 show patterns of orientation along the roll for two users acquired at two time instances. It may be observed from these figures that the two patterns belonging to a user are very similar (high intra-class similarity). It may also be observed that there is hardly any inter-class similarity in this case. This has motivated us to explore the effectiveness of information from the built-in orientation sensor for user authentication.

In addition to the set of behavioral data described above, we computed the following two additional features for every gesture, except for the ST gesture. Since the ST gesture contains only a single point for x-y coordinates, the following features cannot be computed.

• Point curvature: Curvature at each point specifies the slope formed by the user at each successive point. At time t_i , curvature can be computed by [67]:



$$P_{i} = \tan^{-1} \left(\frac{y_{i} - y_{i-1}}{x_{i} - x_{i-1}} \right)$$
(2.

Figure 2.2: Orientation along roll of a user at two time instances.



Figure 2.3: Orientation along roll of another user at two time instances.

where x_i , y_i are the *x* and *y* coordinates of the sample point at event time t_i and x_{i-1} , y_{i-1} are the *x* and *y* coordinates of the sample point at the previous time event (t_{i-1}) . Slope at the first point is considered to be zero.

• Swipe curvature: Curvature of swipe specifies the slope formed by the user while performing a particular gesture. Curvature of swipe can be calculated by [66]:

$$S_i = \tan^{-1} \left(\frac{y_{end} - y_{start}}{x_{end} - x_{start}} \right)$$
(2.2)

where x_{start} and y_{start} are the x and y coordinates of the start point of gesture, x_{end} and y_{end} are the x and y coordinates of the end point of gesture.

2.1.3 Score computation and normalization

During enrolment, feature templates that contain the above mentioned features are created for each of the gestures. In a similar way, a query template is formed when the user performs a gesture on the touchscreen device. During verification, matching of the corresponding feature sets is performed using the modified Hausdorff distance (MHD) [159]. The reason for choosing MHD is that it compares each index of query gesture with every index of training gesture and vice versa, whereas other distance measures the distance between corresponding indices. Additionally, it has also been observed through experiments that in Dubuisson and Jain [160] the modified Hausdorff distance (MHD) is more robust to outliers and consistently outperforms other possible Hausdorff based

distance measures, when used to measure similarity or dissimilarity between two sets of points. Therefore, in this work, we employed MHD for computation of matching scores. The modified Hausdorff distance between two sets $A = \{a_1, a_2, ..., a_p\}$ and $B = \{b_1, b_2, ..., b_q\}$ is defined as

$$H(A,B) = \max(h(A,B),h(B,A))$$
(2.3)

where,

$$h(A,B) = \frac{1}{p} \sum_{a \in A} \left(\min_{b \in B} \left\| a - b \right\| \right)$$
(2.4)

$$h(B,A) = \frac{1}{q} \sum_{b \in B} \left(\min_{a \in A} \left\| a - b \right\| \right)$$
(2.5)

In the above equations, $\|\cdot\|$ represents the L_2 norm between the two points of sets A and B. Also, h(A,B) is the forward Hausdorff distance and h(B,A) is the reverse Hausdorff distance. Essentially, the average of minimum distances from every point of one set with the other set is computed in both forward and reverse directions. The MHD is then computed by finding the maximum of the forward and the reverse Hausdorff distance.

The proposed approach for matching features using MHD generates matching scores for each of the constituents of our feature set. In the case of a scalar feature, the swipe curvature in our case, the feature matching using the MHD effectively reduces to matching the corresponding features using the Euclidean distance. As the information acquired from the accelerometer and orientation sensors is measured in three directions, the matching process generates a total of six scores. The matching of rest of the features that include x - y coordinates, finger area, point curvature and swipe curvature generates a score each. Therefore, a comparison or a match between a query and reference template in the database results in a total of 10 matching scores for each of the gestures, except for ST. As discussed earlier, since point curvature and swipe curvature cannot be computed for ST, the number of scores generated for a comparison between corresponding ST features is only 8. The next step in our approach is to normalize multiple matching scores generated to a common domain. There are various techniques available in the literature for score normalization [160, 161]. In this work, we have investigated the following techniques for score normalization.

1. min-max: min-max normalization technique linearly transforms matching scores into the range of 0 to 1. Normalized matching scores are computed as follows:

$$s_k = \frac{s_k - \min}{\max - \min}$$
(2.6)

where min and max are the minimum and the maximum values of matching scores, respectively.

2. z-score: In this technique, normalized matching scores are computed as follows:

$$s'_{k} = \frac{s_{k} - \mu}{\sigma} \tag{2.7}$$

where μ and σ are the mean and standard deviation of matching scores, respectively.

3. tanh-estimator: The mathematical expression for score normalization using tanhestimator is given as follows:

$$s'_{k} = \frac{1}{2} \left\{ \tanh\left(0.01 \times \left(\frac{s_{k} - \mu_{G}}{\sigma_{G}}\right)\right) + 1 \right\}$$
(2.8)

where μ_G and σ_G are the mean and standard deviation of genuine scores. In Equations (2.6), (2.7), (2.8), s_k and s'_k are the matching score and the normalized matching score, respectively.

4. w-score: The w-score scheme is originally proposed for score normalization in the recognition framework [160], and the approach uses distribution of scores generated by matching a probe to all gallery samples. Since we performed experiments in the verification scenario, which is more appropriate for user authentication in smartphones, overall distribution of genuine and impostor scores is used for score normalization. This is consistent with other score normalization techniques considered in our work.

2.1.4 Fusion of scores

The score normalization technique transforms the scores into a common domain, so that the scores can be combined using a fusion method [162]. There are numerous ways by which scores can be combined. Some of them are sum of scores, maximum score, minimum score, weighted sum and product of scores [163]. In this work, we have explored the following rules for combination of matching scores.

1. Min Rule: In this rule, the combined score is the minimum of the set of scores being combined

$$S = \min(s_1, s_2, \dots, s_n)$$
(2.9)

2. Max Rule: The combined score is the maximum of the set of scores being combined

$$S = \max(s_1, s_2, \dots, s_n)$$
(2.10)

3. Product Rule: Product rule is mathematically expressed as follows:

$$S = \prod_{k=1}^{n} s_k \tag{2.11}$$

4. Sum Rule: Mathematically, the combined score is computed as follows:

$$S = \sum_{k=1}^{n} s_k \tag{2.12}$$

2.2 Experimental results and discussion

Ideally, the proposed approach should be effective on any touchscreen device. To ascertain this aspect of the performance of the proposed approach, two different devices were selected as they were popular smartphones differing in certain aspects. Specifically, they differ in screen size, weight and precision of the finger area measurement.

2.2.1 Dataset-I

Since there is no publicly available database that contains all the behavioral data that we have explored in this work, we developed a database of 104 users (Dataset-I). For the purpose of data acquisition, we developed an Android application on IntelliJ IDEA platform and ran it on Samsung Galaxy S-II GT-I9100 Android phone. Out of 104 users, 82 users were having prior experience of operating touchscreen phones. Participants in the data collection process conducted at our institute primarily included students aged between 19 and 36 years. Specifically, the dataset comprises 9 users between the age 31-36 years, 40 users between 26-30 years and 55 users between 19-25 years. Approximately 80% of the participants were students. Among the student participants, the majority were studying IT and related subjects, while others were pursing mechanical engineering. A few of the student participants were studying the science subjects such as bioscience, mathematics and physics. These participants were asked to perform the following gestures on the touchscreen: L2R, R2L, SU, SD, ZI, ZO and ST. Each of these gestures was performed three times by every user. As described in Section 2.1, the raw data captured by our application consists of x - y coordinates, finger area, accelerometer and orientation sensor readings from the touchscreen.

The following sections present results from a set of experiments carried out to evaluate the performance of the proposed approach. In our initial experiments, we evaluated various combinations of score normalization and fusion techniques. Results from these experiments are presented in Section 2.2.2. Based on these results, the best techniques for score normalization and fusion are identified and these techniques are investigated further.

2.2.2 Performance of different score normalization and fusion techniques

The objective of this set of experiments is to evaluate the performance of different score normalization and fusion techniques for matching gestures in our dataset. Table 2.1 summarizes EERs obtained for score level fusion of all gestures using various score normalization and fusion techniques. As can be seen in this table, using min-max and z-score normalization techniques results in unacceptably high EERs. This may be because

the parameters associated with min-max and z-score normalization are sensitive to outliers, and therefore, these approaches are less robust. The robustness of a score normalization technique, as defined in [160], is its insensitivity towards outliers. If an approach performs better than the other in the presence of outliers, the former approach is more insensitive to outliers than the latter and hence, more robust. As can be seen in this table, tanh-estimator normalization approach consistently provides the best matching performance, except for the case in which fusion is performed with min rule. According to the observations in [160], tanh-estimator is highly efficient and robust as it is less sensitive to outliers in the matching scores. It can also be noted that EER of w-score normalization with max rule is quite high as compared to its performance with other fusion rules. The major problem with the max rule is that if any of the individual scores being combined is 1 (upper limit of normalized scores), then the fused score gets confined to 1. In our experiments, we observed that w-score normalization is more likely (than tanh-estimator) to yield a score with value 1 and this leads to high verification error as majority of the fused scores have a value of 1. This explains the poor performance of the max rule based fusion with w-score normalization. It can also be seen in the table that min, max and product score fusion approaches perform poorly compared to sum rule based fusion. This may be because min, max and product rule are highly sensitive to outliers. A single outlier will affect the fused score in the case of min, max and product rule. On the other hand, fusion using sum rule depends on the scores generated by each constituent of the feature set.

Score normalization	Score level fusion technique				
technique	Min	Max	Product	Sum	
min-max	12.84	47.14	50.14	24.12	
z-score	64.29	56.09	50.02	50.64	
w-score	11.76	50	5.43	1.92	
tanh-estimator	27.80	3.40	0.32	0.31	

Table 2.1: EERs (%) obtained with different score normalization and fusion techniques

In addition, the combination of tanh-estimator for score normalization and sum rule for fusion provides the best matching performance among different combinations considered. Fig. 2.4, 2.5, 2.6 and 2.7 show receiver operating characteristics (ROC) of



Figure 2.4: ROCs for score level combination of gestures with min rule and different normalization techniques.



Figure 2.5: ROCs for score level combination of gestures with max rule and different normalization techniques.



Figure 2.6: ROCs for score level combination of gestures with product rule and different normalization techniques.



Figure 2.7: ROCs for score level combination of gestures with sum rule and different normalization techniques.

score level fusion techniques with different score normalization schemes. It may be noted that for the product and max rule based fusion schemes in Figs. 2.5 and 2.6, ROC curves corresponding to the w-score normalization are not plotted. This is due to the nature of distribution of fused matching scores. Specifically, range of values of genuine scores is quite high with majority of scores having low values and a few of them having very high values. This necessitates a very small increment in threshold to plot a smooth ROC curve. However, a small step size (threshold increment) results in high computational complexity and leads to memory error in MATLAB.

2.2.3 Performance of score normalization techniques with sum rule for fusion

In this section, sum rule is employed for fusion of scores and performance (in terms of EERs) of different score normalization techniques are investigated for matching individual gestures. Here, matching of gestures is performed by considering the entire feature set. A set of genuine and impostor scores generated using the MHD based matching of corresponding features are normalized using min-max, z-score, w-score and tanh-estimator techniques. These normalized scores are then combined using sum rule. Experimental results are presented in Table 2.2, which shows EERs of individual gestures and their combination. It may also be noted from Tables 2.1 and 2.2 that w-score normalization scheme performs significantly better than z-score. A similar trend has been observed in [161] for recognition. However, more important observation is that tanhestimator scheme for score normalization clearly outperforms min-max, z-score and wscore techniques. This may be due to the fact that min-max technique uses minimum and maximum value, while the z-score method employs mean and standard deviation of matching scores; hence both these techniques are quite sensitive to outliers in the matching scores and cause performance degradation. This observation is also consistent with the results presented in [160] for sum rule based fusion. Superior performance of tanh-estimator on our dataset as compared to other schemes considered in this study is probably due to its robustness to noisy data [160, 161].

Gesture	min-max	z-score	w-score	tanh-estimator
L2R	36.53	48.98	12.03	3.42
R2L	20.96	48.37	9.44	3.69
SD	32.03	47.73	13.44	4.68
SU	33.27	48.39	11.83	3.06
ST	25.86	46.14	9.96	7.47
ZI	23.43	45.46	10.83	3.10
ZO	24.26	45.83	7.02	1.61
Combined	24.12	50.64	1.92	0.31

Table 2.2: EERs (%) for gesture matching with sum rule based fusion

2.2.4 Performance of fusion rules with tanh-estimator for score normalization

In this set of experiments, matching scores are normalized using tanh-estimator technique and the performance of four commonly used score level fusion rules is evaluated. Table 2.3 presents EERs achieved while matching individual gestures with this experimental setting. This table also presents EERs for score level fusion of all gestures. The corresponding ROCs for fusion of all gestures are depicted in Fig. 2.8. It can be observed from the table that sum and product rules clearly outperform (with significant reduction in EERs) min and max rules for score level fusion. Interestingly, EERs achieved with the product rule are quite comparable to that of the sum rule. However, a closer observation of the ROCs (in Fig. 2.8) of two of these score level fusion techniques reveals that sum rule based fusion yields consistently higher genuine acceptance rates (GAR) for the same range of false acceptance rates (FAR). This experimental observation shows that sum rule performs better than other fusion schemes is consistent with the observation in [163], in which authors reported experimental results from extensive evaluation of different combination schemes.

For the rest of the experiments in this section, tanh-estimator and sum rule are employed for normalization and fusion of matching scores, respectively. This is based on our observation from the above set of experiments that these techniques for normalization and fusion of matching scores achieve better performance as compared to other techniques.

Gesture	Min	Max	Product	Sum
L2R	28.19	5.03	3.48	3.42
R2L	29.32	6.31	3.64	3.69
SD	37.74	7.51	4.68	4.68
SU	33.51	3.95	3.05	3.06
ST	27.99	8.40	7.54	7.47
ZI	45.34	4.71	3.23	3.10
ZO	24.39	2.32	1.60	1.61
Combined	27.80	3.40	0.32	0.31

Table 2.3: EERs (%) for gesture matching with tanh-estimator for score normalization



Figure 2.8: ROCs for score level combination of gestures with tanh-estimator for score normalization.

As a comparative study, we have also implemented dynamic time warping (DTW) based matching, as it has been found very effective [63, 66, 68] for user authentication based on touchscreen gestures. In order to compare the performance of the proposed approach with DTW, we performed two sets of experiments in the verification mode. The objective of the first set of experiments was to evaluate the effectiveness of individual features, while in the second, we ascertained the matching performance of individual gestures and their combination.

In the first set of experiments, we considered one feature at a time for each of the gestures. In order to compute EER for each of the features, a set of genuine scores are generated by comparing a gesture with the rest of the corresponding gesture samples of the same user. This is repeated for all the gestures considered in this work. This experimental setting assumes that users perform all gestures during authentication. Multiple matching scores generated from matching of these gestures are normalized using tanh-estimator technique and then combined using sum rule to obtain a consolidated score. Since we have collected three samples of each gesture from 104 users, the total number of genuine scores generated is 312. Similarly, a set of impostor scores are generated by matching gestures of a user with the corresponding gestures of all other users in the dataset, resulting in 48204 impostor scores.

The results from this set of experiments are presented in Table 2.4. As it can be observed from this table, the orientation feature achieves EER of 0.56% and outperforms all other features considered in this study. Fig. 2.9 shows distribution of genuine and impostor matching scores for orientation feature. It can be noticed from the figure, there is only a little overlap between the genuine and impostor scores, which explains the promising results obtained for gesture matching using orientation feature.

It is also important to note from Table 2.4 that the proposed approach outperforms the DTW based matching for three of the features namely, orientation sensor, x - y coordinates and the point curvature, with significant improvement in performance. For the rest of the features, the EERs of the two methods are quite comparable, with DTW approach achieving marginal improvement over our approach. Fig. 2.10 shows performance of the features individually in terms of the ROC for DTW and MHD based matching. It can be observed in the figure that accelerometer and orientation sensors perform significantly better than the rest of the features employed in this work. Additionally, at a FAR of 0.01%, GAR of MHD based matching for orientation sensor is 100%, while that of DTW based matching is 96.79%. More importantly, it can be observed from the figure that MHD based matching works better than DTW for all the features except finger area at low FAR. This implies that our approach can be used in the applications, where security is a major concern.

Features	DTW	MHD
Accelerometer sensor	2.91	3.35
Orientation sensor	1.67	0.56
<i>x</i> - <i>y</i> coordinates	17.03	11.24
Finger area	13.06	14.12
Point curvature	27.36	19.70
Swipe curvature	21.83	22.70

Table 2.4: EER (%) of individual features



Figure 2.9: Distribution of genuine and impostor scores for orientation feature.



Figure 2.10: ROCs of individual features with DTW and the proposed MHD based matching.

In second set of experiments, firstly, we investigated the performance of the proposed approach when the user performs only one of the gestures considered in this study. Secondly, we also investigated the performance for score level combination of all gestures. In order to evaluate and compare the performance of the gestures, we considered one gesture at a time.

A set of genuine scores are then generated by matching a query gesture with the rest of the corresponding gesture samples of the same user. As described in Section 2.1.3, matching a pair of gestures involves matching of the corresponding features in the feature set. Multiple matching scores generated in the process are then combined using sum rule to get a consolidated score. In the same way, the impostor scores are obtained by comparing a query gesture with corresponding gestures of other users. The process is repeated for every gesture and corresponding EERs are calculated. Table 2.5 presents EERs obtained for gesture matching using the proposed and the DTW based method. It can be seen that ZO gesture provides the best authentication performance in our dataset. Also, it can be observed that performance of ST gesture is relatively poor. This is quite expected as the ST gesture, being a single point gesture, does not contain much information. More importantly, it may also be observed that our approach outperforms DTW based matching consistently for all gestures. To further investigate the combined performance of all gestures, the scores generated from matching of individual gestures are combined using the sum rule. Table 2.5 also shows EERs obtained for this experiment. The proposed approach achieves EER of 0.31%, a significant improvement (of 80%) over the DTW based method. In other words, the total number of falsely accepted impostor and falsely rejected genuine samples has reduced from 5 to 1. Fig. 2.11 shows comparison of ROCs for the proposed and the DTW based matching for score level fusion of all gestures. It can be seen in the figure that MHD based approach achieves higher GARs for all FARs. Specifically, at a FAR of 0.01%, GAR of MHD based matching for combination of all the gestures is 99.68%, while that of DTW is 97.76%.

Table 2.6 shows average time taken to calculate matching scores for individual gestures using DTW and MHD based matching. It can be observed from this table that the proposed approach, as compared to DTW, requires significantly less time to perform

Gesture	DTW	MHD
L2R	6.22	3.42
R2L	4.87	3.69
SD	8.98	4.68
SU	3.86	3.06
ST	8.04	7.47
ZI	13.20	3.10
ZO	4.88	1.61
Combined	1.55	0.31

Table 2.5: EER (%) of individual gestures and their combination



Figure 2.11: ROCs for score level combination (sum rule) of all gestures.

matching of gestures. It can also be noted that average time taken to match ST gestures is very less as compared to other gestures. It is because the ST gesture contains single point for matching. In addition, two of the features namely, the point curvature and swipe curvature features are not employed for matching ST gestures. Matching of ST gestures using MHD achieves highest performance improvement of 80% over DTW. For score level combination of all gestures, our approach achieves an improvement of 18.84% over DTW based matching. From these experimental results, it is evident that the proposed approach outperforms DTW based matching in terms of both error rate and the computational time.

It may be noted that, in our experiments, we have not considered weighted combination (sum) of the matching scores as we do not have adequate data to train the system on an independent training subset to obtain optimal values for weights. We believe that if optimal weights are employed for combination of matching scores, it may be possible to further improve the performance of the system.

Gesture	DTW	MHD	Performance improvement (%)
L2R	8.66	5.96	31.18
R2L	9.24	6.52	29.44
SD	11.34	8.74	22.93
SU	11.50	8.79	23.56
ST	1.30	0.26	80
ZI	20.76	19.98	3.76
ZO	14.58	12.55	13.92
Combined	77.38	62.8	18.84

Table 2.6: Time (ms) required for matching individual gestures

2.2.5 Performance of orientation, accelerometer versus the rest

In this section, the individual performance of orientation feature is compared with rest of the features. The ROC curves corresponding to the orientation feature and score level combination of rest of the features using sum rule are shown in Fig. 2.12. In terms of EERs, user authentication based solely on orientation feature yielded 0.56% EER, while the sum rule combination of rest of the features yielded an EER of 1.2%.



Figure 2.12: Performance comparison of orientation feature with the sum rule based score level combination of rest of the features.

Furthermore, we have performed experiments to compare performance of the proposed user authentication approach using combination of orientation and accelerometer features with the combination of rest of the features. Fig. 2.13 shows ROC curves from this set of experiments. It may be noted from this figure that the combination of accelerometer and orientation features yielded EER of 0.64%, while the combination of rest of the features yielded EER of 6.26%, suggesting that the combination of accelerometer and orientation features clearly outperforms the combination of rest of the features. The above experimental results indicate that the orientation and accelerometer sensor readings carry significant discriminatory information for user authentication in smartphones. It may be noted that the data acquired from built-in orientation sensor not only provides information on how the user holds the device, but also provides measurement of the continuous changes (on all three axes) in the orientation of the device while the user performs a gesture. The accelerometer sensor also operates in a similar manner providing linear acceleration of the device along the three axes. On the other hand, gesture features such as x - y coordinates, point curvature and swipe curvature rely mainly on the pattern (shape) of the gesture performed on the touchscreen. Intra-class variability of these gesture patterns causes erroneous matches resulting in higher error rates. This probably explains why these features carry limited discriminatory information as compared to orientation and accelerometer sensor readings.



Figure 2.13: Performance comparison of combination of orientation and accelerometer features with combination of rest of the features.

2.2.6 Dataset-II

To investigate if the device used for data acquisition has any impact on the performance of the proposed authentication approach, we evaluated performance on a new dataset, which we refer to as Dataset-II. This dataset contains data from 30 subjects acquired using another smartphone-Samsung Galaxy Note-II N7100. Out of the 30, only two users were not having prior experience of operating touchscreen phones. Similar to Dataset-I, most of the participants in collection of Dataset-II were also students. Experimental setting for data acquisition remained the same as that of Dataset-I. The performance of the proposed authentication algorithm for different features on Dataset-II is summarized in Table 2.7. The key observation in Table 2.7 is that, although the individual EERs corresponding to different features have changed, their performance trend remains the same as in dataset-I. More importantly, it can be observed that the orientation sensor readings offer the best discrimination, followed by the accelerometer. It may be noted that a one-to-one comparison of EERs reported in Tables 2.4 and 2.7 cannot be made as these performance statistics are obtained on two different datasets, with Dataset-II having considerable number (about 50%) of new users. The above results suggest that the device has no considerable impact on the performance of the proposed approach for mobile user authentication.

Features	EER
Accelerometer sensor	4.55
Orientation sensor	0.99
<i>x</i> - <i>y</i> coordinates	7.63
Finger area	19.45
Point curvature	23.07
Swipe curvature	23.17
Combined	0.03

Table 2.7: EER (%) of individual features for Dataset-II

2.3 Summary

In this chapter, we presented an approach for user authentication in smartphones based on behavioral biometrics. The Android application that we have developed in this work runs on the smartphones and acquires a set of behavioral data when the user interacts with the device. The modified Hausdorff distance (MHD) is used for matching of features of the corresponding gestures, after identifying the category of the gesture performed by the user. Performance evaluation on a relatively large dataset (Dataset-I) of 104 users show that the proposed MHD based matching achieves better performance (in terms of both error rate and computational time) than the approach based on DTW. Our approach achieves the lowest EER of 0.31% when all the gestures performed by the users are combined at the score level for user authentication. Our experimental results also show that the information acquired from the built-in orientation and accelerometer sensors carry high discriminatory information for user authentication. This also indicates that the way user performs gestures on mobile phones is unique to some extent and the information can be exploited for user authentication. The performance of the proposed algorithm is also ascertained on a dataset (Dataset-II) of 30 subjects captured using another smartphone. The experimental results show a performance trend similar to the one on the Dataset-I. It can be inferred that changing the device does not affect the performance of the proposed algorithm. The framework for user authentication developed in this work is very well suited for continuous authentication in smartphones.
Chapter 3

Gender Recognition in Smartphones using Touchscreen Gestures

In Chapter 2, we have presented an approach for user authentication using touchscreen gestures. It has been observed in the literature that the performance of biometric authentication can be improved by combining traditional biometric information with soft biometrics like gender, age, height, weight, and ethnicity. Gender classification in smartphones has a lot of potential applications apart from biometric authentication. The gender recognition approaches can also be utilized to validate the gender information provided by the user. For instance, in order to gain illegitimate access to the smartphone an imposter may not provide the correct gender information. In addition, the interaction between human and machine can be enhanced based on the gender information. Moreover, the gender information can be used in pervasive computing applications. For example, the information can be communicated to devices embedded in the smart space to adjust the intensity and color of the room lighting based on general preferences of the identified gender. This chapter presents an approach for gender recognition in smartphones using touchscreen gestures performed by the user. The behavioral characteristics such as accelerometer sensor readings, gyroscope sensor readings and finger area are captured during the user's interaction with the touchscreen device. This set of behavioral characteristics is referred to as primary set of attributes. These characteristics are further enriched by deriving a secondary set of attributes such as velocity of swipe, point curvature, and length of swipe. The two-dimensional attribute maps are then formed for the gestures considered in this work by utilizing the primary and secondary sets of attributes. We have employed GIST descriptors to determine the holistic image representation into a low-dimensional feature vector. The contribution of this work can be summarized as follows: a novel approach is presented for gender recognition in smartphones using touchscreen gestures. The information extracted from the built-in sensors in smartphones is investigated for gender recognition. The suitability of an expanded set of touch gestures that includes horizontal swipes, vertical swipes,

zoom-in and zoom-out is investigated for gender recognition. Additionally, different combinations of these gestures are explored for performance improvement.

The rest of the chapter is organized as follows: The proposed approach for gender recognition is described in Section 3.1. The description of the datasets employed as well as the experimental results and discussion are reported in Section 3.2. Section 3.3 presents summary of this work.

3.1 Proposed approach

Fig. 3.1 shows the block diagram of the proposed approach for gender recognition using touchscreen gestures. The Android application collects behavioral data during a user's interaction with the smartphone. These behavioral characteristics are collected at every touch point when a user performs gestures on the touchscreen. The primary set of behavioral attributes acquired using the Android application consists of accelerometer sensor readings, gyroscope sensor readings, orientation sensor readings and the area covered by the finger on the screen. A total of six gestures are considered in this work namely, L2R, R2L, SU, SD, ZI and ZO. These gestures are classified based on the starting and end points of the collected x - y coordinates. In addition to the primary attributes, a secondary set of attributes are derived from x - y coordinates. Feature selection is performed on a set of image-based features extracted from the primary and secondary gesture attributes to identify the best set of features for gender classification. A classifier then generates matching scores by comparing the query sample with the training templates. The scores thus generated are combined using a score level fusion method. Finally, a decision (based on the consolidated score) is taken as to whether the query sample belongs to a male or a female user.



Figure 3.1: Block diagram of the proposed approach for gender recognition.

3.1.1 Data acquisition and classification of gestures

We have developed an Android application to collect behavioral data during the user's touchscreen-based interaction with the smartphone. The behavioral data collected using the Android application includes x - y coordinates, accelerometer sensor readings, gyroscope sensor readings, orientation sensor readings and the finger area. The x and y coordinates provide locations of the finger at every touch point on the touchscreen when the user performs a gesture. Accelerometer sensor readings signify the amount of linear acceleration in x, y and z directions. Gyroscope sensor records the angular speed in x, y and z directions. Orientation sensor is a position sensor, which determines the position of the device with respect to the world's frame of reference in the three directions [4]. The finger area measurement gives the area covered by the finger on the touchscreen. The acquired data is processed to identify the user's gesture using the approach described in Section 2.1.1, which relies on the starting and end positions of the swipe.

3.1.2 Feature extraction

As mentioned previously, our approach uses a set of secondary gesture attributes derived from the collected x - y coordinate data. This set of attributes includes velocity of swipe, point curvature, and length of swipe. Velocity at each touch point is a measure of distance traveled from previous touch point to current touch point over the time taken. Velocity at t_i is computed as follows:

$$V_{i} = \frac{\sqrt{(x_{i} - x_{i-1})^{2} + (y_{i} - y_{i-1})^{2}}}{t_{i} - t_{i-1}}$$
(3.1)

Point curvature represents the slope formed at each touch point while performing the gesture. The curvature at point i can be determined by [164]:

$$P_{i} = \tan^{-1} \left(\frac{y_{i} - y_{i-1}}{x_{i} - x_{i-1}} \right)$$
(3.2)

Length of swipe is a measure of distance covered by the finger while performing the gesture. At finger location i, length of swipe can be calculated by:

$$L_{i} = L_{i-1} + \sqrt{(x_{i} - x_{i-1})^{2} + (y_{i} - y_{i-1})^{2}}$$
(3.3)

For each gesture, these secondary attributes are computed at every touch point. The next step in our approach involves extraction of features from the primary and secondary gesture attributes. For this purpose, a two-dimensional (2D) attribute map is obtained for every attribute by forming an image, the size of which is the same as that of the touchscreen. This idea of 2D representation of attributes is partly inspired by the approach presented by Zhao et al. [71], in which images are formed by using x - ycoordinates and the corresponding pressure values. In our approach, the pixel value at a particular pixel in each of the 2D attribute maps is the corresponding attribute value at that location on the touchscreen. Since accelerometer sensor, gyroscope sensor, and orientation sensor provide the tri-axial readings at each touch point, nine attribute maps are formed using these attributes. In addition, four attribute maps are formed by using finger area, velocity of swipe, point curvature and length of swipe, as they provide only one value at each touch point. In this way, a total of 13 attribute maps are formed for each swipe gesture. In the case of multi-touch gestures such as ZI and ZO, locations of both the fingers are used to form the attribute maps. Fig. 3.2 shows velocity attribute maps of R2L and SU swipes performed by 10 male and 10 female users. In general, it appears that female users perform longer horizontal swipes and more curved vertical swipes as compared to male users.

To extract discriminatory features for gender classification, GIST descriptor is computed on the attribute maps. GIST descriptor was originally proposed for scene recognition by Oliva et al. [165]. We have chosen this descriptor as it captures holistic spatial properties of an image in the form of a low-dimensional vector. As can be observed in Fig. 3.2, the value of most of the pixels in the attribute maps is zero. Therefore, local descriptors do not seem to be appropriate. In the literature, GIST descriptor has been successfully employed in various applications involving image processing, including recognizing human actions [166], human-human interaction recognition [167], scene completion, [168], object/place recognition [169, 170] and copy detection [170].

Prior to computing GIST descriptor, attribute maps are enhanced by contrast stretching. Thereafter, to compute the GIST descriptors, attribute maps are first convolved with 32 Gabor filters (at 4 scales and 8 orientations) [165]. This process generates 32 output maps of the same size as the input attribute maps. Subsequently, each of the output maps is divided into non-overlapping 4^{\times} 4 grids, and the average of each grid is computed. This results in a vector having 16 elements (average values) for each output map. Vectors thus generated from 32 output maps are concatenated to form the final representation. In this way, our feature extraction using GIST descriptor generates a 512-dimensional feature vector, which represents the 'gist' of the attribute maps at different scales and orientations.



Figure 3.2: Samples of swipes (a) R2L (b) SU performed by 10 male and 10 female users (the images are enhanced for better visibility).

3.1.3 Feature selection

Feature selection is employed to identify the best features for gender recognition from the set of features discussed in Section 3.1.2. In the proposed approach, feature selection is performed using wrapper subset evaluator with the BestFirst attribute selection approach [171]. The default settings available in WEKA data mining toolbox [134] are used for both wrapper subset evaluator and the BestFirst attribute selection. Wrapper approach evaluates the subset of attributes by using induced classifier and finds the best possible feature set. BestFirst searches for attribute subsets in one of the three ways namely, forward, backward and bi-directional. The forward direction search starts with an empty set of attributes set of attributes and deletes the non-improving features. The bi-directional search starts with any random point and adds or deletes features from both directions. In this work, we have employed the forward direction attribute selection approach to identify the best subset of features.

3.1.4 Information fusion

In order to achieve the best performance for gender recognition, we have explored the information fusion at feature level and score level [29]. Feature level fusion is performed by concatenating the feature sets, as shown below:

In order to utilize multiple attributes of a touch gesture and multiple gestures for predicting the user's gender, we have explored feature level and score level [29] information fusion techniques. Specifically, concatenating the feature sets in the following manner performs feature level fusion:

$$F = [f_1 f_2 f_3 \dots f_N]$$
(3.4)

and linear combination of multiple matching scores performs score level fusion, which is mathematically shown as follows:

$$S = w_1 s_1 + w_2 s_2 + \dots + w_N s_N \tag{3.5}$$

where $s_1, s_2, ..., s_N$ are matching scores generated by the classifiers and $w_1, w_2, ..., w_N$ are weights, which are selected in such a way that $w_1 + w_2 + ... + w_N = 1$. In our experiments, we have employed equal weights to combine the scores.

3.1.5 Classifiers

In the proposed approach, k-NN classifier is employed to predict the user's gender based on a set of features computed from his/her touchscreen gestures. This classifier [172] predicts the class label for every test sample based on a set of distances computed between the feature vector corresponding to the test sample and those corresponding to the training samples. Typically, the category of the test sample is determined based on majority voting by considering k smallest distances. In our experiments, Euclidean distance is used as a measure of distance between feature vectors and the parameter k is set to 1.

3.2 Experimental results and discussion

The performance of the proposed approach for gender recognition has been evaluated on two datasets. We have collected these datasets during users' interaction with two different devices. Brief descriptions of the datasets and our experiments are presented in the following sub-sections.

3.2.1 Data acquisition

We have developed an Android application using IntelliJ IDEA platform to acquire behavioral data while a user performs gestures on the touchscreen. As discussed in Section 2.2, behavioral data was collected using two touchscreen devices namely, Samsung Galaxy Note-II N7100 (Note-II) and S-II GT-I9100 (S-II). A total of 126 users from our institute participated in data collection process. Every subject performed six gestures namely, L2R, R2L, SU, SD, ZI and ZO three times. The application collected x - y coordinates, accelerometer sensor readings, gyroscope sensor readings, orientation sensor readings and finger area.

3.2.2 Evaluation on Dataset-I

Dataset-I, which was created using Note-II device, contains touch gesture data from 39 female and 45 male subjects. All our performance evaluations are based on 5-fold cross validation. While partitioning the dataset into 5 folds, we have ensured that there is no overlap of subjects between the training and test set. We have divided the dataset in such a way that all the samples from a subject are grouped into a single fold. Specifically, all the samples of 45 male subjects are distributed equally into 5 folds, whereas all the

samples of 35 female participants are divided equally into 5 folds and the remaining samples from 4 female subjects are distributed equally into four of the five folds. In this way, four of the five folds contain all the samples of 9 male and 8 female subjects and the remaining fold contains all the samples of 9 male and 7 female subjects.

In the first set of experiments, the performance of features computed from seven gesture attributes (discussed in Section 3.1.2) are evaluated individually. The classification accuracies for each of the gestures are shown in Fig. 3.3. Although no single gesture attribute provides the best performance for all the gestures, each attribute performs well for at least one of the gestures.



Figure 3.3: Comparison of individual performances of the GIST features computed from the attribute maps for all the gestures.

In Fig. 3.3, it can be observed that all the attributes considered in this study contain information useful for predicting the user's gender. Therefore, the combination of GIST features extracted from the seven attributes is explored using feature level and score level techniques for each of the gestures. The results of this set of experiments are presented in Table 3.1. As can be observed, the score level fusion performs consistently better than the feature level fusion. The performance of score level fusion can perhaps be improved further by optimizing the weights involved in the fusion. In the remaining sets of experiments, only the score level fusion is employed.

	1	
Gesture	Feature level fusion	Score level fusion
L2R	73.46%	78.23%
R2L	74.26%	77.50%
SD	71.47%	74.27%
SU	73.06%	80.10%
ZI	77.03%	77.40%
ZO	73.07%	76.23%

Table 3.1: Comparison of feature level and score level fusion of information from the attribute maps on Dataset-I

To investigate whether the combination of gestures leads to improved gender recognition, different combinations are evaluated. The bar graph in Fig. 3.4 shows the classification accuracies obtained when gestures are combined at the score level. Initially, all possible combinations of two gestures are considered. The second bar in Fig. 3.4 indicates the resulting minimum, maximum and average classification accuracies. Similarly, other combinations involving 3, 4, and 5 gestures are evaluated and the results displayed in Fig. 3.4. As can be observed, the gender classification accuracy increases with the number of gestures combined. However, the maximum classification accuracy achieved when five gestures are combined is marginally better than the accuracy achieved when all the gestures are combined. Specifically, our approach achieves 94.46% accuracy when all the gestures except R2L are combined, as against 93.65% when all the gestures are combined.



Figure 3.4: Performance of gender classification for different combinations of gestures.

In the next set of experiments, the proposed approach is compared with the current state-of-the-art for gender prediction from swipe gestures [115]. For this purpose, we have implemented their approach and evaluated its performance on our datasets. Since Note-II device does not provide pressure information, it is not utilized while evaluating the existing approach on Dataset-I. Additionally, since Miguel-Hurtado et al. [115] have not considered multi-touch gestures such as ZI and ZO in their work, only four gestures namely L2R, R2L, SD, and SU are considered for a fair comparison of the two approaches. Since their approach achieved the best classification accuracy with decision level fusion, a comparison of results with plurality voting based fusion of decisions is also reported. The results of this set of experiments are presented in Table 3.2. As can be observed, the proposed approach achieves considerably higher classification accuracy compared to the current state-of-the-art approach. Also, our approach provides consistent improvement, including when the gestures are combined at the score and decision levels.

Gestures	Miguel-Hurtado et al. [115]	Proposed approach		
L2R	64.27%	78.23%		
R2L	59.63%	77.50%		
SD	55.88%	74.27%		
SU	53.75%	80.10%		
Score level fusion	60.74%	92.45%		
Decision level fusion	66.77%	84.98%		

Table 3.2: Performance comparison with the existing approach for gender recognition on Dataset-I

3.2.3 Evaluation on Dataset-II

To ascertain the performance of the proposed approach, we have performed experiments similar to the ones explained in Section 3.2.2 on another dataset collected using a different device. Specifically, touch gesture data in Dataset-II is acquired using S-II device. This dataset contains gesture data collected from 17 female and 25 male subjects. As in the previous experiments, we have ensured that there is no overlap of users between the training and test datasets used for 5-fold cross validation. Specifically, all the samples of 25 male subjects are distributed equally into 5 folds, whereas all the samples of 15 female participants are divided equally into 5 folds and the samples of the remaining 2 female subjects are distributed equally into two of the five folds. In this way, two of the five folds contain all the samples of 5 male and 4 female subjects and the remaining folds contain all the samples of 5 male and 3 female subjects.

The results of our first set of experiments are presented in Fig. 3.5, which shows the gender classification accuracies achieved using individual attributes of each of the touch gestures. In general, the performance trend is similar to what has been observed on Dataset-I with no clear winner among the attributes.

In the next set of experiments, we have evaluated the performance when GIST features computed from the seven attributes are combined at the feature and score levels. The classification accuracies achieved using each of the gestures are presented in Table 3.3. As can be observed, score level fusion yields higher classification accuracy than feature level fusion in a majority of cases. Therefore, the rest of our experimental evaluations are based only on the score level fusion of gesture attributes.



Figure 3.5: Comparison of individual performances of the GIST features computed from the attribute maps for all the gestures.

Table 3.3 Comparison of featu	re level and score leve	el fusion of infor	rmation from t	he attribute
	mans on Datas	at II		

Gesture	Feature level fusion	Score level fusion
L2R	77.13%	82.59%
R2L	77.78%	76.39%
SD	70.09%	72.22%
SU	77.10%	77.41%
ZI	81.48%	82.22%
ZO	83.70%	77.40%

Furthermore, we have performed experiments to evaluate different combinations of gestures and the results are presented in Fig. 3.6. The experimental protocol and performance metrics remain the same as the ones used for evaluations on Dataset-I. Our approach achieves the highest gender classification accuracy of 92.96% when all the gestures are combined at the score level.



Figure 3.6: Performance of gender classification for different combinations of gestures.

In the last set of experiments, the proposed approach is compared with the existing approach presented by Miguel-Hurtado et al. [115]. Since S-II device records pressure information, it is considered while evaluating their approach on Dataset-II. A performance comparison between the two approaches is presented in Table 3.4. Again, the proposed approach achieves higher classification accuracy for gender classification using individual gestures as well as their combination. Overall, the results of our evaluations on Dataset-II and Dataset-II demonstrate cross-device capability of the proposed approach.

Dataset-II				
Gestures	Miguel-Hurtado et al. [115]	Proposed approach		
L2R	69.17%	82.59%		
R2L	47.50%	76.39%		
SD	56.11%	72.22%		
SU	57.50%	77.41%		
Score level fusion	62.32%	88.80%		
Decision level fusion	57.50%	78.70%		

Table 3.4: Performance comparison with the existing approach for gender recognition on Dataset-II

3.3 Summary

In this chapter, we have presented an approach for touchscreen gesture-based gender recognition in smartphones. A smartphone application is developed to acquire a set of behavioral data during the user's interaction with the touchscreen-based smartphone. A 2D attribute map is generated for each of the seven gesture attributes, which include accelerometer sensor readings, gyroscope sensor readings, orientation sensor readings, finger area, swipe velocity, point curvature and length of swipe. Our approach uses GIST descriptor to extract holistic spatial features of the attribute maps and a feature selection approach to identify the best set of features for gender recognition. The performance of the proposed approach has been evaluated on two datasets collected using different devices. Our results suggest that the attributes considered in this study provide information useful for touch gesture-based gender recognition. In addition, the combination of gestures using score level fusion technique results in enhanced gender classification. Most importantly, the proposed approach achieves state-of-the-art performance on both the datasets and handles cross-device scenarios well. In summary, the results of this study suggest that the behavioral data acquired during the user's interaction with the device can be utilized to recognize user's gender reliably.

Chapter 4

Gender Classification in Smartphones using Gait Information

In the previous chapter, we have presented an approach for gender recognition in smartphones using the touchscreen gestures collected during the user's interaction with the device. However, for continuous classification of gender, the information from the smartphone should be collected continuously even if the user is not interacting with the device. Therefore, we perform gender recognition while the user walks with a smartphone in the trouser pocket. In this chapter, we present an approach for gender classification using users' gait information acquired from the built-in sensors of a smartphone. Histogram of gradient (HG) method is proposed to extract features from the gait data, which includes a set of signals collected from accelerometer and gyroscope sensors of a smartphone. The motivation behind choosing gait is that it can be captured unobtrusively without any conscious effort from the user. The key contributions of this work are as follows:

- 1) A novel approach for classification of gender of a smartphone user using the proposed histogram of gradient (HG) features.
- 2) The proposed work explores the combination of gait information collected from accelerometer and gyroscope sensors for gender recognition in smartphones.
- 3) A total of 654 gait data is acquired from 109 subjects using smartphone sensors with variations in walking speed. The gait data is collected using two different devices to ascertain the gender recognition accuracy of the proposed approach.

The rest of the chapter is organized as follows: A detailed description of the proposed approach for gender recognition is presented in Section 4.1. Section 4.2 presents details about the gait dataset collected using smartphones. This section also presents experimental results and discussion. Finally, summary of this chapter is presented in Section 4.3.

4.1 Proposed method

The block diagram of the proposed approach for gender classification is shown in Fig. 4.1. The Android application collects the gait biometrics of the user. Specifically, readings from built-in accelerometer and gyroscope sensors of smartphone are recorded with the help of this application. The accelerometer and gyroscope sensor readings in x, y and z directions constitute the gait data, which are preprocessed for resampling and low pass filtering in the time domain. In the next stage, features are extracted from the gait data. Finally, the discriminatory features extracted from accelerometer and gyroscope sensor readings are combined to obtain the final feature representation, which is fed to a binary classifier for classification of the user's gender.



Figure 4.1: Block diagram of proposed approach for gender recognition.

4.1.1 Data acquisition

The Android application captures the accelerometer and gyroscope sensor readings of the user's gait in x, y and z directions of the smartphone. Accelerometer sensor provides the amount of linear acceleration on the phone in x, y and z directions. Fig. 4.2 shows typical accelerometer readings of the user's gait captured by the Android application.



Figure 4.2. Accelerometer sensor signals in x, y and z directions.

On the other hand, gyroscope sensor measures angular speed exerted on the phone by the user along x, y and z directions. Fig. 4.3 shows typical gyroscope readings of the user's gait captured by the Android application.



Figure 4.3. Gyroscope sensor signals in x, y and z directions.

4.1.2 Preprocessing

4.1.2.1 Normalization

The accelerometer and gyroscope sensor readings are normalized to have zero mean and unit standard deviation [131]. Since the sampling rate of the signals captured by the application from accelerometer and gyroscope sensors is not fixed, it is necessary to resample the signal at a fixed sampling rate before further processing. In this work, these signals are resampled at a fixed sampling rate of 100 Hz by using cubic spline interpolation. This is followed by moving average filtering of the signals to reduce noise.

4.1.2.2 Gait cycle extraction

As can be seen in Fig. 4.4, human gait signals generally exhibit periodically repeating patterns with multiple gait cycles. However, the gait cycles in the recorded gait data may vary with time because of many reasons including variations in walking speed of the user and irregular walking behavior. Majority of the existing techniques for gait cycle extraction utilize accelerometer sensor reading in only one direction. However, our preliminary experiments indicated that such an approach might not provide accurate gait

cycles. Therefore, in this work, we have developed a heuristic technique to extract gait cycles using sensor readings in all the three directions. Our approach primarily relies on accelerometer signal in the z-direction for identifying local minima points, which help us in extracting gait cycles. Specifically, a point is considered to be local minimum if the value of the current sample is less than that of its neighboring samples and their differences are higher than a threshold. These local minima points are refined using accelerometer signals in x and y directions. Specifically, only those points that have corresponding local minima points (within a small neighborhood of 0.5 sec) in accelerometer signals in x and y directions are considered for extracting gait cycles. Fig. 4.4(a) shows the local minima points computed for linear acceleration signal in zdirection, in which the set of sample points between two consecutive minima constitute one gait cycle. In our experiments, a total of seven such gait cycles are extracted for further processing. However, we have also performed a set of experiments to study the effect of the number of gait cycles on the gender recognition performance, results of which are presented in Section 4.2.4. Fig. 4.4(b) shows gait cycles extracted from the accelerometer sensor signal in z-direction. Similarly, gait cycles from the rest of the signals - accelerometer sensor signals in x and y directions and gyroscope sensor signals in x, y and z directions are extracted by considering the sample points that correspond to those constituting the gait cycles extracted from accelerometer sensor signal in the zdirection.



Figure 4.4: (a) Local minima points of accelerometer signal in z-direction and (b) the extracted gait cycles.

4.1.3 Feature extraction

The proposed approach for gender classification employs a novel feature extraction technique, which is based on the histogram of oriented gradients (HOG) proposed by Dalal and Triggs [173]. The HOG was originally proposed for human detection in images. Due to its excellent performance, HOG and its variants have been successfully employed for numerous computer vision applications including face recognition [174], character recognition [175], traffic sign recognition [176] and video surveillance system [177]. Besides these applications, HOG descriptors have also been utilized to recognize gender from still images of body parts [178]. In this work, we develop a one-dimensional (1D) version of the HOG descriptor. Since gradients are oriented in only one direction in a 1D signal, we refer to the modified descriptor as the histogram of gradients (HG). Fig. 4.5 shows the computational stages involved in the proposed feature extraction technique. Details of each of these stages are presented in the following sub-sections.



Figure 4.5: Overview of the proposed feature extraction method.

4.1.3.1 Computation of gradients

The first step in the proposed feature extraction process is to compute the gradient and the angle of gradient. For discrete signals, gradients can be estimated using different masks. In this work, we have explored the following masks [173] for computation of gradients: centered [-1,0,1], uncentered [-1,1] and cubic-corrected [1,-8,0,8,-1].

4.1.3.2 Histogram binning

The key processing stage in the proposed feature extraction technique is the histogram binning, which generates a gradient-based histogram feature. In this stage, the gait cycle is divided into six non-overlapping cells and a histogram for each cell is computed using the gradient and the angle of gradient. The range of gradient angle $(0^{\circ} - 180^{\circ})$ is divided uniformly to generate six histogram bins [173]. Fig. 4.6 shows an example of a cell histogram computed using the gradient and angle of gradient. A bin for each element of the cell is identified using its gradient angle and it is voted with the gradient value of the

element. This process is repeated for every element of the cell to generate a cell histogram.



Angle

Figure 4.6: Generation of HG descriptor.

4.1.3.3 Block normalization

The cell histograms computed in the previous stage are normalized by considering a group of cells or a block. In this work, overlapping blocks consisting of two cells are considered. This generates five overlapping blocks for a gait cycle with six cells. The cell histograms of a block can be normalized using various methods and we have explored L_1 -norm, L_1 -sqrt and L_2 -norm based normalization techniques, which are defined as follows [173]:

(1) L₁-norm:

$$H_N = \frac{H}{(\|H\|_1 + \epsilon)} \tag{4.1}$$

(2) L_1 -sqrt:

$$H_N = \sqrt{\frac{H}{(\|H\|_1 + \epsilon)}} \tag{4.2}$$

(3) L₂-norm:

$$H_N = \frac{H}{\sqrt{(\|H\|_2^2 + \epsilon^2)}}$$
(4.3)

where *H* is the histogram generated by concatenating two cell histograms in a block, H_N is the corresponding normalized histogram and \in is a very small constant. The block

normalization process generates five normalized histograms for each gait cycle, which are concatenated to form the HG descriptor of length 60. Finally, the normalized histograms corresponding to seven gait cycles are concatenated to form a feature vector of length 420 for a gait signal.

4.1.4 Feature level combination

As discussed in the previous section, the proposed approach generates a histogram feature of length 420 for each of the gait signals acquired from accelerometer and gyroscope sensors in x, y and z directions. Therefore, there is a need for efficient combination of this information. Broadly, information fusion can be performed in two ways namely, pre-classification fusion and post-classification fusion. Jain et al. [160] observed that pre-classification fusion is expected to perform better than post-classification fusion. Therefore, in the proposed work, information fusion is performed at feature level (pre-classification) by concatenating the normalized histograms to obtain the final representation. In our case, the pre-classification information fusion can be represented as:

$$H_{Ncat} = [H_{N1} H_{N2} H_{N3} \dots H_{Nn}]$$
(4.4)

where, $H_{N1}, H_{N2}, H_{N3}, \dots, H_{Nn}$ are the histogram features and $H_{N_{cat}}$ is the concatenated histogram. In eqn. (4.4), individual histogram features (of length 420) that are computed from accelerometer and gyroscope sensor readings in *x*, *y* and *z* directions are concatenated to form the final feature representation of length 2520.

4.1.5 Classification

Bootstrap aggregating [179], also known as bagging, is designed to provide stable and improved performance for machine learning algorithms used for classification and regression. It is an ensemble method in which multiple predictors are aggregated. It trains an ensemble of decision trees either for classification or regression using bootstrap method. Specifically, it generates an aggregated predictor by using multiple versions of a predictor, which are created by training the base predictor on bootstrap replicates of the training set. Finally, bagging generates an aggregated model by combining the outputs of the individual models using plurality voting in the case of classification.

4.2 Experimental results and discussion

4.2.1 Data acquisition

We have developed an Android application on IntelliJ platform to acquire users' gait information. The application captures accelerometer and gyroscope sensor readings in x, y and z directions of the device, which we collectively refer to as gait data. We have employed two devices namely, Samsung Galaxy S-II GT-I9100 (S-II) and Note-II N7100 (Note-II) for collection of gait data. A total of 109 subjects participated in the data collection process. These subjects were instructed to keep the device in the front pocket of their trouser and to walk in a straight path. To collect gait data corresponding to different walking speeds, subjects were instructed to walk at (what they think is) normal, fast and slow walking speeds. This process was repeated twice, collecting six gait data from each subject. Therefore, our dataset consists of a total of 654 gait data acquired from 109 subjects.

4.2.2. Experiments with S-II device

The S-II device was used to collect gait data from 46 subjects, out of which 25 were male and 21 were female subjects in the age range of 19 and 36 years. This subset of our dataset consists of 276 gait data. Initially, we have performed experiments to evaluate individual performances of accelerometer and gyroscope sensor readings for gender identification. In this set of experiments, we have used the 5-fold cross-validation methodology and ensured that no overlap of subjects existed between our training and testing datasets. This is ensured by partitioning the dataset in such a way that all gait data from a subject is grouped into a single partition. Specifically, gait data belonging to 25 male subjects are equally divided to generate 5 partitions, with each partition containing 5 male subjects' gait data. Similarly, gait data belonging to 20 female subjects are distributed equally into five partitions and the remaining gait data belonging to a female user is considered to be part of one of the five partitions. In this way, four partitions for the 5-fold cross-validation contain gait data belonging to 5 male and 4 female subjects and the fifth partition contains gait data belonging to 5 male and 5 female subjects.

Our preliminary experiments indicated that the centered mask with L_1 -norm based normalization provides better performance than other combinations of gradient

computation and histogram normalization techniques for gender classification. Therefore, we have employed the above techniques for further analysis. The average 5-fold crossvalidation accuracy of the proposed approach on this dataset is presented in Table 4.1. It can be observed from the table that the performance of gender classification using gait data collected from gyroscope sensor is comparable with that of the approach based only on the accelerometer sensor readings, except for the case when subjects walked at slow pace. This observation motivated us to explore the combination of information from gyroscope and accelerometer sensor readings to further improve the gender classification accuracy. We have performed pre-classification fusion by concatenating the two HG descriptors derived from accelerometer and gyroscope sensors readings. The accuracy of the gender classification approach which combines information from accelerometer and gyroscope sensors is presented in the last column of Table 4.1. It can be seen that the gender classification approach based on the combination of information from accelerometer and gyroscope sensors clearly outperforms the one based on either accelerometer or gyroscope sensors. Therefore, it can be concluded from this set of experiments that in addition to accelerometer sensor readings, gyroscope sensor readings also provide the discriminatory information for gender recognition and their feature level combination leads to significant improvement in classification accuracy.

	Gait Information				
Walking Speed	Accelerometer	Gyroscope	Accelerometer + Gyroscope		
Normal	78.56%	81.67%	91.78%		
Fast	81.11%	80%	94.44%		
Slow	83.33%	70.83%	88.89%		

Table 4.1: Comparison of individual performance of gait information collected from accelerometer and gyroscope sensors with the combined performance for gender classification

We have also performed a set of experiments, which considers more realistic scenarios. It is very unlikely that a user walks at the same speed all the time. Therefore, ideally, the performance of gender classification approach should not be affected by variations in the user's walking speeds. To evaluate the performance of the proposed approach under such scenarios, we have performed experiments with training and testing sets consisting of gait data corresponding to different walking speeds, which we refer to as cross-speed gender classification. The experimental results from this set of experiments are presented in Table 4.2. In all these experiments, we have partitioned the dataset into training and testing sets, which contain all users' gait data corresponding to a specific category. However, for normal versus normal, fast versus fast and slow versus slow cases, gait data belonging to 50% of the users is used for training, while the rest of the gait data is used for testing, which ensures that there is no overlap of users between training and testing sets. It can be observed from Table 4.2 that the performance deterioration for cross-speed cases is only marginal, except for the cases such as slow versus fast and fast versus slow. This is possibly due to significant change in the walking speeds between the training and testing sets for those cases.

Train Test Normal		Fast	Slow	
Normal 93.18%		92.22%	92.04%	
Fast	92.39%	95.45%	88.64%	
Slow 89.13%		85.56%	92.86%	

Table 4.2: Cross-speed gender classification accuracy on gait dataset collected using S-II device

To analyze why the HG features are effective for gender classification, we have identified the top 3 features in the proposed HG based feature vector and plotted a scatter diagram. The complete set of gait data corresponding to normal walking speed collected using S-II device is used for this purpose. We have employed the Fisher score based feature selection technique [180], which assigns a rank to each feature based on its discriminative capability, to identify the top 3 features. Fig. 4.7 shows the scatter diagram, which helps us visualize the discriminative capability of these features. As can be seen in this figure, the majority of data points corresponding to male and female subjects are well separated in the three-dimensional feature space. In the original feature space, the separation between the two clusters of data points belonging to male and female subjects is expected to improve further due to the inclusion of additional features. This possibly explains why the proposed features are effective for gender classification using gait data.



Figure 4.7: Scatter diagram of the top three features determined using Fisher score based feature selection method.

4.2.3 Experiments with Note-II device

In this section, we present results from a set of experiments carried out to evaluate the performance of the proposed approach on a dataset collected using a different device. The key objective of this experiment is to ascertain the performance of the proposed gender classification approach and to investigate if the device used for data collection has any impact on its performance. As described in Section 4.2.1, we have employed Note-II device to collect gait data from 63 subjects. Out of 63 subjects, 33 were male and 30 were female subjects aged between 20 and 33 years. The two sets of subjects (who participated in experiments with S-II and Note-II devices) are disjoint. Six gait samples with variations in walking speed (slow, normal and fast) are acquired from each subject. Therefore, this dataset contains a total of 378 gait data.

In the first set of experiments, effectiveness of features extracted from accelerometer and gyroscope sensor readings are investigated separately. In addition, we have also explored the feature level combination of accelerometer and gyroscope sensor readings. In these experiments, we have used 5-fold cross-validation methodology and partitioned the dataset in such a way that no overlap of subjects existed between the five subsets. Specifically, we have created five subsets, with each subset containing gait data of 6 male and 6 female subjects. The gait data belonging to remaining 3 male subjects are distributed into three different subsets. The average 5-fold classification accuracies from this set of experiments are presented in Table 4.3. It can be observed that the classification accuracies presented in Tables 4.1 and 4.3 are comparable and therefore the performance is consistent across the two gait datasets. The results presented in Table 4.3 further demonstrate that gyroscope sensor readings are useful for gait based gender recognition in mobile phones. As expected, the combination of information from accelerometer and gyroscope sensor readings results in significant improvement in classification accuracy.

	Gait Information				
Walking Speed	king Speed Accelerometer Gyroscope		Accelerometer + Gyroscope		
Normal	85.59%	76.08%	90.48%		
Fast	85.13%	76.68%	91.07%		
Slow	78.53%	71.39%	88.46%		

Table 4.3: Comparison of individual performance of gait information collected from accelerometer and gyroscope sensors with the combined performance for gender identification

In the second set of experiments, we have considered cross-speed gender classification scenarios. We have followed the same experimental protocol as the one adopted for cross-speed experiments described in Section 4.2.2. The experimental results are summarized in Table 4.4. The key observation in this table is that the performance trend remains the same as in the first dataset. Specifically, there is slight deterioration in the classification accuracy for majority of the cross-speed cases presented in Table 4.4. However, the performance deterioration is significant for fast versus slow and slow versus fast scenarios, indicating that the proposed approach may not be very effective when there is significant change in walking speeds between training and testing data.

Table 4.4: Cross-speed gender classification accuracy on gait dataset collected using Note-II device

Train Test	Normal	Fast	Slow
Normal	93.33%	88.33%	87.72%
Fast	90.98%	91.67%	80.70%
Slow	83.61%	70.83%	87.50%

In order to visualize the effectiveness of the proposed HG features on this dataset, the top 3 features in the HG based feature vector have been determined using Fisher score based feature selection technique (please refer to Section 4.2.2). The complete set of gait data corresponding to normal walking speed collected using Note-II device has been used for this experiment. Fig. 4.8 shows the scatter diagram of the top 3 features of the proposed HG feature vector. As expected, most of the data points belonging to male and female subjects are well separated in the three-dimensional feature space. Also, it is noteworthy that the distribution of these features belonging to male and female users is quite similar to the one shown in Fig. 4.7.



Figure 4.8: Scatter diagram for the top three features determined using Fisher score based feature selection method.

4.2.4 Effect of number of gait cycles on the performance

To study the effect of the number of gait cycles on the performance of the proposed approach, a set of experiments has been performed by varying the number of cycles employed for feature extraction. Fig. 4.9 shows the classification accuracy versus the number of gait cycles for datasets captured using S-II and Note-II devices. As can be observed, the gender recognition accuracy increases with the number of gait cycles. These results indicate that employing more number of gait cycles for feature extraction leads to better characterization of the user's gender at the expense of the computational performance. Since the maximum number of gait cycles available for feature extraction is 7 for some of the users in the dataset, we have set this parameter to 7 in our experiments.



Figure 4.9: Classification accuracy versus the number of gait cycles employed for gender recognition.

4.2.5 Comparison with existing methods

The performance of the proposed approach has been compared with the existing approaches for gender classification [133, 135]. For a fair comparison, we have implemented the approaches reported in Weiss and Lockhart [133] and Jain and Kanhangad [135] and evaluated their performances on our datasets. As has been done in Weiss and Lockhart [133], the classification is performed using three classifiers namely, instance-based learning (IB3), J48 decision tree (J48) and multilayer neural network (NN), which are available in the Weka data mining tool [134]. In this set of experiments, we have adopted the leave-one-out method (LOOM), which was used for performance evaluation in Weiss and Lockhart [133]. Results from this set of experiments, performed on the datasets collected using S-II and Note-II, are presented in Table 4.5. It can be observed from this table that the proposed approach achieves higher classification accuracy than the existing approaches. It is also noteworthy that our approach provides considerable performance improvement consistently for different walking speeds considered in this study.

		Dataset					
Reference	Classifier	S-II		Note-II			
		Normal	Fast	Slow	Normal	Fast	Slow
Weiss and Lockhart [133]	NN	71.74%	79.76%	63.64%	73.77%	74.60%	75.16%
	IB3	72.83%	82.14%	60.23%	68.03%	73.81%	70.63%
	J48	82.61%	71.43%	54.55%	55.74%	68.25%	74.60%
Jain and Kanh	angad [135]	75%	67.50%	73.86%	74.75%	78.33%	65.79%
This work		89.13%	92.22%	85.23%	88.52%	90%	84.21%

Table 4.5: Classification accuracy of the proposed and existing approaches for different walking speeds

The key strength of our approach is the discriminating power of the proposed HG features that provide superior performance. Apart from the information that these features carry, we believe that the way in which they are computed in our approach contributes to their higher discriminating power. Since the walking speed affects the duration of the gait cycles, we have extracted features from individual gait cycles. This approach makes the extracted features invariant to momentary changes in the user's walking speed. On the other hand, Weiss and Lockhart [133] computed a set of statistical features using windows of fixed duration. Similarly, Jain and Kanhangad [135] computed their local curvature-based features from segments of the gait signals that contain 8 cycles. Due to the fixed-length segments involved in their feature extraction process, the approaches mentioned above are likely to be adversely affected by momentary variations in user's walking speed. This is supported by our experimental results presented in Table 4.5. A disadvantage of the proposed approach is the higher dimensionality of our feature vector. In addition, our approach involves more parameters compared with the existing methods [133, 135].

4.2.6 Summary of experimental results

This study considered more realistic conditions (compared to the existing studies) as we have carried out performance evaluation on two datasets collected using different devices. The experimental results suggest that the fusion of information extracted from linear acceleration and angular velocity gait signals leads to improved gender recognition. In addition, the results of our cross-speed experiments show that the proposed approach performs well when there is limited variation in the walking speed during the training and

test sessions. In other words, the approach is likely to fail if this variation is significant. This is because HG features capture local gradient information by dividing the signal into a number of cells and the use of fixed-size cells in our approach resulted in intra-class variations in the extracted features. A straightforward way to overcome this problem is to adapt the cell size to the speed of walking. This can be achieved through coarse-grained classification of user's gait speed into categories such as slow, normal and fast and assigning appropriate cell sizes to each of these walking speeds.

4.3 Summary

In this chapter, we have presented an approach for gender recognition in smartphones using gait biometrics. An Android application is developed to acquire gait data using built-in accelerometer and gyroscope sensors of smartphones and a novel approach based on histogram of gradients is proposed for feature extraction. The key observations from our experiments can be summarized as follows:

- Features extracted from gyroscope sensor readings also carry information useful for gender classification in smartphones. More importantly, the feature level combination of accelerometer and gyroscope sensor readings results in significant improvement of performance of the proposed approach.
- 2) It is observed from the cross-speed experiments that variations in the user's walking speed have a minimal impact on the performance of the proposed approach, except for cases where there is significant change in walking speeds between training and testing sets.
- 3) Our experimental results also show that the performance of the proposed approach is consistent across two gait datasets collected using two different devices and it achieves higher classification accuracy than the existing works.

Chapter 5

Human Activity Classification in Smartphones using Accelerometer and Gyroscope Sensors

In the previous chapter, it is observed that the performance of gender recognition can be improved by classifying the walking speed of the user into categories namely slow, normal and fast. Further, this information can be utilized to compare the query sample only with the samples belonging to the same category of the walking speed in the training set. Activity classification in smartphones helps us monitor and analyze the physical activities of the user in daily life and has potential applications in healthcare systems. Moreover, excessive sitting and lack of adequate levels of physical activity are associated with health problems such as obesity, diabetes, cardiovascular disease, poor metabolic health and depression, leading to the increased risk of mortality. Activity recognition can be used for continuous analysis of the daily activities performed by the user. Such an analysis is useful in understanding the behavior and thereby, making it possible to provide automated suggestions for reducing the risk factor for various non-communicable diseases. Additionally, signals corresponding to the other activities performed by the user can also be collected using the built-in accelerometer and gyroscope sensors. A lot of research on activity recognition in the literature is focused on using the dedicated bodyworn sensors. These sensors may become too cumbersome to wear, especially for continuous activity recognition in which users may have to wear the sensors for extended periods of time. Therefore, in this chapter, we propose a descriptor-based approach for activity classification using built-in sensors of smartphones. Accelerometer and gyroscope sensor signals are acquired to identify the activities performed by the user.

Since the existing works demonstrated the feasibility of recognizing activities using built-in sensors of smartphones or a similar device with promising performance on large datasets, there is a pressing need to develop approaches that provide highly accurate and reliable performance. Moreover, enhancing the performance is imperative for the smartphone-based solution to be a superior alternative to the dedicated body-worn sensor based methods in real-world applications.

The rest of the chapter is organized as follows: A detailed explanation of the proposed approach for activity classification is presented in Section 5.1. Section 5.2 presents the description of datasets employed in this thesis along with experimental results and discussion. Finally, Section 5.3 summarizes the chapter.

5.1 Proposed approach

Fig. 5.1 shows the block diagram of the proposed approach for activity classification. The input signals are acquired from the built-in accelerometer and gyroscope sensors of the smartphone. The signals captured from accelerometer and gyroscope sensors provide tri-axial linear acceleration and angular velocity information, respectively. Additionally, a set of signals is derived from the input signals. The feature extraction is performed using the proposed descriptors and the resultant feature sets are combined at the feature level. Subsequently, the combined feature vector is fed to the classifier to determine the activity performed by the user.



Figure 5.1: Block diagram of the proposed approach for activity recognition.

5.1.1 Input signal

The set of input signals consists of linear acceleration and angular velocity signals collected using the built-in accelerometer and gyroscope sensors of the smartphone, while the user performs different activities. The tri-axial linear acceleration and tri-axial angular velocity signals are denoted as a_x , a_y , a_z and g_x , g_y , g_z , respectively. Fig. 5.2 shows typical patterns in sample accelerometer sensor signals in x, y and z directions for different activities. It may be seen that patterns corresponding to most of the activities are distinct. For example, acceleration signals in x and z directions show distinct patterns for

walking upstairs and downstairs. Since features play a significant role in classification of activities, a careful observation of the signals in Fig. 5.2 indicates that the feature descriptors need to be designed to effectively characterize patterns of evolution of these signals.



Figure 5.2: Acceleration signals in x, y and z directions when a user performs different activities.

5.1.2 Extraction of additional signals

The proposed approach also utilizes a set of additional signals that are derived from the input signals. This approach is similar to the one adopted by Anguita et al. [156]. Firstly, two magnitude (Mag) signals corresponding to the acquired tri-axial linear acceleration and angular velocity signals are computed as follows:

$$Mag = \sqrt{x^2 + y^2 + z^2}$$
(5.1)

This results in a total of eight time-domain signals. The set of frequency domain signals in our approach includes the magnitude signals computed from the Fourier transforms of the time-domain signals. In this way, a total of eight time-domain and eight frequency-domain signals are generated for feature extraction.

5.1.3 Feature extraction

In the proposed approach for activity recognition, feature extraction is performed using two descriptors namely, histogram of gradient (HG) and the Fourier descriptor (FD). Feature set-I is computed using HG descriptor that provides a locally normalized histogram of gradients. This feature is computed from all the sixteen signals. Feature set-II is computed by forming a closed curve, which we refer to as *activity curve* in this chapter. The activity curve is formed by considering two signals x(t) and y(t) (from the set of 16 signals) and generating a set of 2D data points (x(t), y(t)), which in the 2D space forms a curve. Specifically, a closed curve is obtained by appending the first samples to the respective signals. Feature set-II comprises the shape features extracted from the activity curves. The reason for selecting only two signals at a time is to simplify the extraction of features from these curves.

5.1.3.1 Feature set-I (FSI)

This feature set is computed using a one-dimensional (1D) version of the HOG descriptor, which was originally proposed for human detection in images [173]. HOG and its variants have been shown to perform well for numerous applications including face recognition [174], fungus detection [181], etc. Since the gradients are oriented in only one direction in a 1D signal, we refer to the modified descriptor as the histogram of gradients (HG).

Fig. 5.3 shows the computational stages involved in the HG technique. The first step is to compute the gradient and its angle. For discrete signals, gradients are estimated using the centered masks [-1,0,1]. The key processing stage in this technique is the histogram binning, which generates a gradient-based histogram feature. In this stage, input signal is divided into four non-overlapping cells and a histogram for each cell is computed using the gradient and the angle of gradient. The range of gradient angle (0° - 180°) is divided uniformly to generate seven histogram bins [173]. A bin for each sample of the cell is identified based on its gradient angle and the corresponding gradient value is voted to that particular bin. This process is repeated for every sample of the cell to generate a cell histogram. Cell histograms computed from a signal are normalized by considering a group of cells, which form a block. Specifically, cell histograms belonging to a block are normalized using the L1-norm, as shown below:

$$H_N = \frac{H}{(\|H\|_1 + \epsilon)} \tag{5.2}$$

where *H* is the histogram generated by concatenating cell histograms belonging to a block, H_N is the corresponding normalized histogram and \in is a very small constant. In this work, overlapping blocks with each block consisting of two cells are considered. This generates three blocks (with an overlap of a cell between them) for a signal. Since each block yields a normalized histogram feature of length 14, the dimensionality of the HG based feature vector is 42.



Figure 5.3: Overview of computation of the HG based feature.

5.1.3.2 Feature set-II (FSII)

Feature set-II (FSII) consists of features that characterize the shape of the activity curves. This idea is loosely inspired by the approach presented by Wang et al. [182] for silhouette-based gait recognition, in which the trajectories formed by the gait sequences in the eigenspace are analyzed for gait recognition. In our approach, the shape of an activity curve is represented using the Fourier descriptors.

FD based shape representation is not only very effective but also is translation, rotation and scale invariant [183]. Essentially, FD is the discrete Fourier transform (DFT) of a shape signature, which describes the shape of an object or in general, a closed contour. In the literature, FD based methods have been employed in various pattern recognition applications such as shape classification and shape retrieval [184, 185] and recognition of precipitation type in meteorology [186]. Commonly used signatures for shape representation using FDs are complex coordinates, centroid distance, curvature signature and cumulative angular function [185]. In the proposed approach, the centroid distance signature based FD is employed. Centroid distance of the activity curve with coordinates (x(t), y(t)) is determined as:

$$d(t) = \sqrt{(x(t) - x_c)^2 + (y(t) - y_c)^2}$$
(5.3)
where, d(t) is the centroid distance at time t, which varies from 1 to N and N is the number of coordinates in the curve. The centroid of the curve, (x_c, y_c) is computed as:

$$x_{c} = \frac{1}{N} \sum_{t=1}^{N} x(t) , \quad y_{c} = \frac{1}{N} \sum_{t=1}^{N} y(t)$$
 (5.4)

The discrete Fourier transform of the centroid distance based shape representation is then determined as follows:

$$FD_{n} = \frac{1}{N} \sum_{t=1}^{N} d(t) \exp(\frac{-j2\pi nt}{N})$$
(5.5)

The Fourier coefficients, FD_n , n = 0,1,..., N-1, are referred to as the Fourier descriptors (FDs). In (5.3), the centroid, (x_c, y_c) is subtracted from the curve coordinates to make the resultant signature invariant to translation. The rotation invariance is achieved by considering only the magnitude of the FDs. Since the centroid distances are real-valued, there will only be N/2 unique FDs. In addition, lower frequency coefficients in DFT contain more relevant shape information as compared to higher frequency coefficients [184]. Therefore, the feature vector is formed using only the first N/2 FDs. Discarding higher frequency coefficients also has the effect of suppressing noise or distortions in the activity curve. To achieve scale invariance, the magnitude of each of the retained coefficients is divided by the DC component [185] and the final FD-based representation is generated as shown below:

$$f = \left[\frac{|FD_1|}{|FD_0|}, \frac{|FD_2|}{|FD_0|}, \dots, \frac{|FD_{N/2}|}{|FD_0|}\right]$$
(5.6)

In our approach, a total of 14 activity curves are formed from the input time and frequency domain signals and their corresponding FDs are computed. Fig. 5.4 shows discriminatory power of the FDs for activity recognition. In this figure, we have plotted FDs extracted from a total of ten signals, with two signals each belonging to five activities. In this case, the activity curves are generated using only the first set of time domain signals. It is evident that there is high intra-class and low inter-class similarity among the FDs.



Figure 5.4: Comparison of the centroid distance signature based FDs of different activities.

The novelty of the feature extraction method employed in this work can be summarized as follows. While the HG descriptor characterizes patterns in the time and frequency domain signals by capturing the gradient information into a histogram feature, the FD captures shape of the activity curves. The latter characterizes patterns of evolution of a time or a frequency domain signal in relation to the other signal involved in generation of the activity curve.

5.1.4 Information fusion

Information fusion can be performed in two ways namely, pre-classification (feature level) fusion and post-classification (score level) fusion. In this work, we have investigated both of these approaches. Specifically, we have performed fusion at the feature level by concatenating the two feature sets, as shown below:

$$F = [f_1 \ f_2] \tag{5.7}$$

and the score level fusion has been performed as follows:

$$S = w_1 s_1 + w_2 s_2 \tag{5.8}$$

where s_1, s_2 are the scores generated by the two classifiers for FSI and FSII, respectively and w_1, w_2 are the weight parameters, which are selected in such a way that $w_1 + w_2 = 1$.

5.1.5 Classifiers

In the proposed approach, a classifier is used to identify the user's activity based on the extracted features. We have explored two classifiers namely, the multiclass support vector machine (SVM) [187] and the k-nearest neighbor (k-NN) [172]. In k-NN classifier, the parameter k is set to 1, and the Euclidean distance is used for computing distance between the feature vectors. The multiclass SVM framework [187] is achieved by decomposing a multiclass classifier into multiple binary classifiers [188]. The most commonly used approaches are one-against-all and one-against-one. In the one-against-all approach, each class is discriminated from the remaining classes, which are considered as a single class. On the other hand, in the one-against-one approach, two classes are discriminated at a time, and likewise, all pairs of classes are discriminated. In this work, one-against-one approach is employed for building our multiclass SVM.

5.2 Experimental results and discussion

Performance evaluations have been carried out on two publicly available datasets namely, UCI HAR dataset [156] (Dataset-I) and the physical activity sensor data [157] (Dataset-II). In this section, we report the activity recognition performance in terms of sensitivity (SEN), specificity (SPF) and accuracy (ACC) [189]. While ACC, which is computed using the total number of correctly classified signal segments in the test set, is a measure of overall performance, SEN and SPF are determined for individual classes.

5.2.1 Evaluation on Dataset-I

Dataset-I contains activity data from 30 candidates collected using Samsung Galaxy S-II smartphone [156]. Specifically, it contains tri-axial accelerometer and gyroscope signals acquired while a user performs different activities with the phone fixed to the waist belt. The set of physical activities includes standing, sitting, laying down, walking, downstairs and upstairs. The raw signals are preprocessed and resampled at 50Hz. The signals are then segmented using a fixed-width sliding window of 2.56s with 50% overlap between successive segments. The dataset is partitioned into training and test sets, which contain 7352 and 2947 signal segments, respectively. Our experimental evaluations on Dataset-I

have been performed using these training and test sets.

In the first set of experiments, we have investigated the effectiveness of feature level and score level fusion techniques. The two features sets FSI and FSII are combined using the methods described in Section 5.1.4. For score level fusion, we have partitioned the training data equally into non-overlapping training and validation sets. The weights are then determined based on the performance of individual classifiers on the validation set. Specifically, the weight parameters have been set using the performance weighting method discussed in [190]. Using this parameter estimation method, we have set the weight parameters w_1 and w_2 to 0.508 and 0.492 for fusion of SVM classifiers, and 0.516 and 0.484 for fusion of k-NN classifiers. Finally, we have evaluated the performance of the proposed approach on the test set with feature level and score level fusion. The results of this set of experiments are presented in Table 5.1. As can be seen, the best average classification accuracies achieved using score level fusion clearly outperforms score level fusion and therefore, we have performed rest of the performance evaluations with feature level fusion.

Fusion method	SVM	k-NN
Feature level fusion	97.12%	91.75%
Score level fusion	96.44%	84.02%

Table 5.1: Activity classification accuracy with feature level and score level fusion

The objective of the second set of experiments is to compare the performance of the two classifiers considered in this work. To this end, we have evaluated their activity classification performance on the individual feature sets as well as their feature level combination. The results of this set of experiments are shown in Fig. 5.5. It can be seen that the SVM classifier provides consistently higher activity classification accuracies compared to the k-NN classifier. Therefore, further performance evaluations on Dataset-I have been carried out using only the SVM classifier.

The objective of the next set of experiments is to ascertain the improvement in recognition performance resulting from the inclusion of additional time and frequency domain signals. To this end, firstly, we have extracted the two feature sets FSI and FSII



Figure 5.5: Performance comparison of the classifiers on Dataset-I using FSI, FSII and their feature level fusion.

from the acquired tri-axial accelerometer and gyroscope signals and evaluated the performance of our approach using these features. In the literature, it has been shown that the activity classification performance can be improved by combining the features extracted from accelerometer and gyroscope sensor readings [140, 157]. We have observed similar performance trend (see classification accuracy for AT+GT in Fig. 5.6) in this set of experiments, results of which are summarized in Fig. 5.6 and Table 5.2. Thereafter, we have appended the features extracted from the magnitude signals to those extracted from the originally acquired signals. The same experimental protocol has been followed for investigating the usefulness of the frequency domain signals. Finally, we have combined the feature sets computed from time and frequency domain signals through concatenation. The bar graph in Fig. 5.6 shows the average activity classification accuracies with respect to the signals employed. These results indicate that the inclusion of additional signals leads to considerable improvement in performance. The individual performances of time and frequency domain signals are comparable. More importantly, the fusion of the corresponding information achieves higher activity classification accuracy.



AT: tri-axial accelerometer signals in time domain, GT: tri-axial gyroscope signals in time domain, MT: Magnitude of accelerometer and gyroscope signals in time domain, AF: tri-axial accelerometer signals in frequency domain, GF: tri-axial gyroscope signals in frequency domain, MF: Magnitude of accelerometer and gyroscope signals in frequency domain, AL: All the signals are considered.

Figure 5.6: Comparison of time and frequency domain signals with combined performance.

The average activity classification accuracies achieved using FSI, FSII and their combination are presented in Table 5.2. These results show that the fusion leads to improved performance in all the three cases and the best performance is achieved when both time and frequency signals are utilized for activity classification.

Signals	FSI	FSII	Fusion
Time domain	84.32%	82.01%	94.57%
Frequency domain	87.78%	92.21%	93.82%
Time domain + Frequency domain	92.67%	93.75%	97.12%

Table 5.2: Classification accuracy using feature sets FSI, FSII and their combination

We have also compared the performance of the proposed approach with those of the existing approaches [140, 156]. Tables 5.3, 5.4 and 5.5 present confusion matrices of the existing approaches (as reported in [156] and [140]) and the proposed approach. As can be observed, our approach correctly classifies more number of signal segments (see the diagonal elements) for majority of the activities. The performance of our approach is relatively poor when classifying segments belonging to the standing activity. It is interesting to note that considerable number of segments belonging to the standing

activity are misclassified as belonging to the sitting activity and vice versa. This is perhaps due to the non-repetitive nature of these activities. In addition, there exists only a slight difference in sensors' readings for sitting and standing activities. The average classification accuracies computed from the confusion matrices in Tables 5.3, 5.4 and 5.5 are 96.33%, 90.13% and 97.12%, respectively. A detailed comparison using the performance measures namely, SEN and SPF is presented in Table 5.6. We have excluded the method in [140] from this comparison, as it is evident (from the results presented in Tables 5.4 and 5.5) that our approach provides significantly better accuracy.

			Predicted as								
		Walking	Upstairs	Downstairs	Sitting	Standing	Laying Down				
ass	Walking	492	1	3	0	0	0				
	Upstairs	18	451	2	0	0	0				
C	Downstairs	4	6	410	0	0	0				
tual	Sitting	0	2	0	432	57	0				
Act	Standing	0	0	0	14	518	0				
	Laying Down	0	0	0	0	0	537				

Table 5.3: Confusion matrix of the approach presented in [156]

Table 5.4: Confusion matrix of the approach presented in [140]

		Predicted as								
		Walking	Upstairs	Downstairs	Sitting	Standing	Laying Down			
	Walking	473	2	21	0	0	0			
Class	Upstairs	17	444	10	0	0	0			
	Downstairs	28	54	338	0	0	0			
tual	Sitting	0	4	0	381	106	0			
Act	Standing	0	10	0	38	484	0			
	Laying Down	0	1	0	0	0	536			

Table 5.5: Confusion matrix of the proposed approach

		Predicted as							
		Walking	Upstairs	Downstairs	Sitting	Standing	Laying Down		
ass	Walking	491	0	5	0	0	0		
	Upstairs	Upstairs 2		1	0	0	0		
C	Downstairs	1	4	415	0	0	0		
tual	Sitting	0	0	0	446	45	0		
Act	Standing	0	0	0	27	505	0		
	Laying Down	0	0	0	0	0	537		

	SI	EN	SPF		
Activity	Anguita et al. [156]	Proposed approach	Anguita et al. [156]	Proposed approach	
Walking	99.19%	98.99%	99.10%	99.88%	
Upstairs	95.75%	99.36%	99.64%	99.84%	
Downstairs	97.62%	98.81%	99.80%	99.76%	
Sitting	87.98%	90.83%	99.43%	98.90%	
Standing	97.37%	94.92%	97.64%	98.14%	
Laying	100%	100%	100%	100%	

Table 5.6: Performance measures of the proposed and the existing approaches

5.2.2 Evaluation on Dataset-II

To ascertain the performance of the proposed algorithm, we have performed a similar set of experiments on Dataset-II. The accelerometer and gyroscope sensor readings in this dataset are collected using iPod Touch [157]. The data is collected from 16 subjects, aged between 21 and 60 years, using the device kept in the front pocket of subject's trousers. The activities performed by the users are slow walking (C1), normal walking (C2), brisk walking (C3), jogging (C4), sitting (C5), normal upstairs (C6), normal downstairs (C7), brisk upstairs (C8) and brisk downstairs (C9). The preprocessing includes noise removal and manual clipping of segments at the beginning and end of the acquired signals. For experiments, a set of signal segments. Since the signals are sampled at 30Hz, there are 60 sample values in a segment. A total of 2807 such segments are made available by Wu et al. [157]. We have performed 10-fold cross-validation on Dataset-II. This helps us make a fair comparison with their approach, for which 10-fold cross-validation results are available [157].

In the first set of experiments, we have evaluated the performance of SVM and k-NN classifiers for activity classification. The classification accuracies achieved using individual feature sets and their fusion are shown in Fig. 5.7. Interestingly, k-NN performs consistently better than SVM on Dataset-II and therefore, we have employed only the k-NN classifier in our further performance analysis. It may be recalled that we observed the opposite trend on Dataset-I. In general, it has been observed in the literature [191-193] that the performance of classifiers depends mainly on the characteristics of the dataset and often a single classifier does not achieve the best performance on all datasets.

As discussed previously, the datasets used in this work for performance evaluation are collected using different devices in different settings. Also, the size of the training sets and sampling rates of the signals are different in these datasets. All of these factors have possibly contributed to the inconsistency in performance of the classifiers. Despite this inconsistency in performance, the proposed method with SVM classifier outperforms the existing methods [156, 157] on both datasets.



Figure 5.7: Performance comparison of the classifiers on Dataset-II using FSI, FSII and their feature level fusion.

As was done in the case of Dataset-I, we have investigated the usefulness of additional time and frequency domain signals that are derived from the signals in Dataset-II. The bar graph in Fig. 5.8 shows the classification accuracies corresponding to the signals employed. The overall performance trend in this figure is very similar to that observed on Dataset-I. More importantly, the combination of FSI and FSII leads to considerable performance improvement in all cases. The results of this set of experiments are summarized in Table 5.7. These results clearly indicate that our approach provides the highest accuracy when both time and frequency domain signals are utilized.

Signals	FSI	FSII	Fusion
Time domain	86.86%	92.73%	93.52%
Frequency domain	87.03%	91.88%	94.29%
Time domain + Frequency domain	91.77%	95.26%	96.83%

Table 5.7: Classification accuracy using feature sets FSI, FSII and their combination



AT: tri-axial accelerometer signals in time domain, GT: tri-axial gyroscope signals in time domain, MT: Magnitude of accelerometer and gyroscope signals in time domain, AF: tri-axial accelerometer signals in frequency domain, GF: tri-axial gyroscope signals in frequency domain, MF: Magnitude of accelerometer and gyroscope signals in frequency domain, AL: All the signals are considered.

Figure 5.8: Comparison of time and frequency domain signals with their combined performance.

The objective of the next set of experiments is to perform a comparative evaluation of our approach with an existing approach [157], which has been evaluated on Dataset-II. Tables 5.8 and 5.9 present confusion matrices of the two approaches. As can be observed, the classification accuracy of the proposed approach is higher for all the activities except for sitting activity (C5). In Table 5.8, it may be seen that the misclassifications occur mostly in discriminating activities C1 through C4. This may be because these activities differ mainly in speed. In addition, the walking speed is likely to vary from user to user. This makes the discrimination of these activities a challenging task. The performance of our approach is quite encouraging as it greatly reduces the number of misclassifications. The results presented in Table 5.8 also indicate that a significant number of segments belonging to activities C6 through C9 are misclassified as belonging to C1, C2 and C3. As can be observed in Table 5.9, the proposed approach greatly reduces the number of misclassifications for these cases as well.

For a detailed analysis, we have computed SEN and SPF from the confusion matrices in Tables 5.8 and 5.9. These results are presented in Table 5.10. Overall, the proposed approach outperforms the existing approach with higher SEN and SPF for majority of the cases. It is noteworthy that our approach provides significant improvements in SEN for classification of activities C6 through C9. The average classification accuracies computed from the confusion matrices in Tables 5.8 and 5.9 are 90.2% and 96.83%, respectively.

			Predicted as								
		C1	C2	C3	C4	C5	C6	C7	C8	С9	
	C1	572	30	5	0	0	0	0	0	0	
	C2	29	602	13	0	0	4	5	0	1	
s	C3	7	17	475	25	0	0	1	0	2	
Clas	C4	0	1	32	389	0	0	0	2	0	
al (C5	0	0	0	0	266	0	0	0	0	
⊾ctu	C6	8	15	2	0	0	67	4	0	0	
A	C7	6	7	3	0	0	4	77	0	0	
	C8	1	8	10	1	0	0	0	50	1	
	С9	0	16	14	0	0	1	0	0	34	

Table 5.8: Confusion matrix of the approach presented in [157]

Table 5.9: Confusion matrix of the proposed approach

		Predicted as								
		C1	C2	C3	C4	C5	C6	C7	C8	С9
	C1	586	7	4	5	0	0	1	1	3
	C2	5	642	7	0	0	0	0	0	0
S	C3	4	10	511	1	0	0	0	0	1
Clas	C4	3	1	0	420	0	0	0	0	0
al (C5	3	0	0	0	260	0	1	0	2
\ctu	C6	5	0	0	0	0	91	0	0	0
Α	C7	5	0	0	0	1	1	86	1	3
	C8	1	1	1	0	0	4	1	63	0
	C9	1	2	2	0	0	0	1	0	59

Table 5.10: Performance measures of the proposed and the existing approaches

Activity		SEN	SPF		
	Wu et al. [157]	Proposed approach	Wu et al. [157]	Proposed approach	
C1	94.23%	96.54%	97.68%	98.77%	
C2	92.05%	98.17%	95.63%	99.02%	
C3	90.13%	96.96%	96.54%	99.39%	
C4	91.75%	99.06%	98.91%	99.75%	
C5	100%	97.74%	100%	99.96%	
C6	69.79%	94.79%	99.67%	99.82%	
C7	79.38%	88.66%	99.63%	99.85%	
C8	70.42%	88.73%	99.93%	99.93%	
C9	52.31%	90.77%	99.85%	99.67%	

5.3 Summary

In this chapter, we have presented a descriptor-based method for classification of daily physical activities using built-in sensors of a smartphone. Our approach utilizes accelerometer and gyroscope sensor readings from a smartphone or a similar device. The approach also utilizes a set of time and frequency domain signals that are derived from the originally acquired signals. The histogram of gradients and the centroid distance signature-based FD are employed for extracting discriminatory information. We have investigated feature level and score level fusion techniques for combination of the resultant information. The proposed approach has been evaluated on two publicly available datasets. Our experimental results show that the fusion at the feature level outperforms the score level fusion and that such a combination provides significant improvement in classification accuracy. Most importantly, the results from our experimental evaluations show that the proposed approach achieves state-of-the-art performance (with the highest average classification accuracy) on the two datasets, despite the fact that these datasets are collected using different devices with differences in the set of activities, sampling rates and placement of the devices. There are a large number of wearable devices that require the user to choose the activity they are performing. When integrated into such devices, our method will provide automatic recognition of activities as an additional feature.

Chapter 6

Conclusions & Future Research Directions

This chapter presents concluding remarks on the main contributions of the thesis and highlights the directions for future research.

6.1 Conclusions

The prime objective of this thesis is the analysis of human behavior based on the information extracted from the signals acquired using built-in sensors in the smartphones. We have proposed methodologies for smartphone-based biometric authentication, gender recognition and physical activity recognition. Our approaches for user authentication and gender recognition utilize the sensor data captured while the user interacts with the smartphone through its touchscreen. We have also investigated the problem of gender recognition in smartphones by analyzing the gait information acquired using the built-in accelerometer and gyroscope sensors. Furthermore, we have presented an approach that utilizes the same set of sensor readings to recognize various physical activities performed by the user.

In Chapter 2, we have presented an approach for biometric authentication in smartphones based on a set of behavioral characteristics captured using built-in sensors in the smartphone during the user's interaction with the device. The performance evaluation on a dataset of 104 users shows that the proposed MHD based matching achieves better performance (in terms of both the accuracy and computational time) than the existing DTW-based approach. Our experimental results also demonstrate that the information acquired from the built-in orientation and accelerometer sensors carry adequate discriminatory information for user authentication. This study indicates that the way user performs gestures on mobile phones is unique to some extent and the information can be exploited for user authentication. The performance of the proposed approach has also been ascertained on a second dataset of 30 subjects collected using another smartphone. The experimental results on the second dataset also show that the proposed approach

outperforms the existing DTW-based approach. The framework for user authentication developed in this thesis is very well suited for continuous authentication in smartphones.

In Chapter 3, we have presented an approach for gender recognition in smartphones using behavioral biometrics. Our approach utilizes a set of attributes acquired during the user's interaction with the smartphone's touchscreen. The performance of the proposed approach has been evaluated on datasets collected from a total of 126 subjects using two devices. Our experimental results show that the attributes considered in this study provide information useful for gender recognition in smartphones. Our experiments also demonstrate that the score level fusion performs better than the feature level fusion on both the datasets. Moreover, a set of experiments performed to evaluate different combinations of gestures show that the accuracy of gender recognition increases with the number of gestures combined. Most importantly, the proposed approach outperforms the existing work on both the datasets. This study suggests that the behavioral traits acquired during the user's interaction with the device provide reliable information on the gender of the smartphone user.

In Chapter 4, we have proposed another approach for smartphone-based gender recognition using gait biometrics. The gait information is acquired using the built-in accelerometer and gyroscope sensors in the smartphone and a novel approach based on the histogram of gradients is proposed for feature extraction. The performance of the proposed approach has been evaluated on datasets collected from 109 subjects using two different devices. Our experimental results show that the features extracted from accelerometer and gyroscope sensor readings carry information useful for gender classification in smartphones. More importantly, the feature level combination of accelerometer and gyroscope sensor readings significantly improves the performance of the proposed approach. The results of our cross-speed experiments indicate that variations in the user's walking speed have a minimal impact on the performance of the proposed approach, except for cases in which there is a significant change in walking speeds between the training and test sets. Our experimental results also demonstrate that the performance of the proposed approach is consistent across the two gait datasets and it achieves higher classification accuracy than the existing works. The results of this study are significant as they indicate that gait information captured using the smartphones'

built-in sensors can be used to derive the user's gender information reliably and unobtrusively.

In Chapter 5, we have presented a descriptor-based method for classification of daily physical activities using built-in accelerometer and gyroscope sensors in the smartphone. The approach utilizes a set of time and frequency domain signals for activity recognition. We have investigated the feature level and score level fusion techniques for the combination of the resultant information. The proposed approach has been evaluated on two publicly available datasets. Our experimental results show that the fusion at the feature level outperforms the score level fusion and that such a combination provides a significant improvement in classification accuracy. Most importantly, the results of our experimental evaluations demonstrate that the proposed approach achieves state-of-the-art performance (with the highest average classification accuracy) on the two datasets, despite the fact that these datasets are collected using different devices with differences in the set of activities, sampling rates and placement of the devices.

In summary, each of the approaches proposed in this thesis has been evaluated on datasets collected using two different devices. We have developed an Android application to acquire the behavioral data while a user performs the touchscreen gestures or while he/she walks with the smartphone in the trouser pocket. The results of our comparative evaluations indicate that the approaches proposed in this thesis achieve state-of-the-art performance consistently. The results presented in this thesis clearly indicate that the smartphone sensor readings provide information related to the user's behavior.

Human behavior-based intelligence for smartphones is a relatively new area of research, which has gained significant attention from the researchers in recent times. This thesis has presented an explorative study on the applicability of the smartphone sensor data for various tasks. However, the current work has certain limitations. This study used in-house datasets as there are only a few public datasets. Although these in-house datasets are larger than most of the datasets used for evaluation of the methods reported in the literature, our datasets can be expanded further to obtain better estimates of the performance of the proposed approaches. In addition, the participants in our data collection processes were in the age group of 19-36 years. A more diverse age group can be considered for data collection for evaluating generalizability of the proposed

approaches. Secondly, the current interface for database collection built using an Android application may not be completely transparent to the user, thus making the user conscious of the data collection process. Running a data collecting application in the background is likely to provide a more realistic user data. Finally, in this work, tri-axial sensor readings were used to compute features, and the resulting dimensionality of the features is considerably high. The storage of the feature templates may not be an issue as the storage capacity of the smartphones has increased significantly in recent times. However, the prediction may take longer due to the increased feature length, which may affect the real-time performance of the system.

6.2 Future research directions

The Android application developed in this work collects the touchscreen gestures and gait data while a smartphone user performs a set of predefined tasks. This data was analyzed by the proposed approaches, which were implemented in the MATLAB environment, for performing user authentication and gender recognition. Since our experimental results demonstrate that the performance of the proposed approaches is quite promising, these approaches can be employed to develop mobile applications that can be deployed on the end-user smartphones. Efforts can be taken for developing the mobile applications keeping in mind the computational constraints associated with the handheld devices such as limited computational power and storage capacity.

Apart from using touch gesture information, another way of performing user authentication can be based on the usage of mobile applications in the smartphone. The way in which a user interacts with the applications and the choice of the applications installed in the device are expected to be specific to an individual. This behavioral information can be exploited to perform user authentication by acquiring the data using the mobile phone. In addition, the daily physical activities performed by the user can be analyzed to achieve user authentication. Since each individual has a habit of performing a particular set of activities in a certain way, these behavioral characteristics of the user can be utilized as a soft biometric trait to perform the smartphone-based user authentication.

It has been demonstrated in the literature that the performance of the user authentication can be improved by supplementing the traditional biometrics with the soft biometrics such as gender, age and emotion. The experiments performed in this thesis demonstrate promising gender recognition performance. Thus, the gender information of the smartphone user can be incorporated with the touch gesture and gait-based biometrics to improve the user authentication performance. Similarly, other soft biometric traits such as age, emotion and stress can also be recognized by analyzing the touchscreen gesture data and gait signals acquired using the smartphone. In addition, the daily physical activities performed by the individual can be explored to determine the gender, age, and ethnicity of the individual.

In this thesis, the datasets were collected by instructing the subjects to perform a set of predefined tasks. In this condition, the user may become conscious while performing the tasks, affecting his behavior and hence, the behavioral data collected by the mobile application. For collecting the datasets in a more realistic scenario, the mobile application can be installed in the subject's smartphone. The subject can be asked to use the phone as he normally would in his routine life. Additionally, the sensor information collected in this study may change if the user walks or travels in a vehicle while performing the gestures on the touchscreen of the smartphone. These scenarios can be explored to provide more realistic touch gesture-based user authentication and gender recognition. On the other hand, in this thesis, the gait signals were collected by keeping the smartphone in the front pocket of the trouser as generally people carry it. In addition, one may also study gender recognition using gait data collected from the phone kept in different pockets of the trouser. Furthermore, it is imperative to develop intelligent systems that perform gender recognition in more realistic and challenging scenarios that involve walking on level, inclined and uneven surfaces in indoor and outdoor environments (crowded place, rainy/sunny day, etc.). A scenario, in which users carry different objects while walking, may also be considered. The datasets of activity recognition employed in this thesis include most commonly performed daily physical activities. However, other complex physical activities such as cooking, cleaning, sports, and meditation as well as overlapping of activities such as talking while having a meal or walking while talking over the phone can also be explored for activity classification.

Although it is difficult, it may be possible to imitate the behavior of a subject by another individual to gain illegitimate access to the biometric system. Therefore, a systematic study should be carried out to investigate the vulnerability of the behavioral biometric system. In future, suitable countermeasures should be devised to secure the biometric systems against such attacks. Additionally, it may also be noted that for smartphone user authentication biometric templates need to be stored either on an external server or in the smartphone. This may pose security threat to the user since his personal and sensitive information may be accessible to a hacker. Therefore, it is essential to develop a smartphone-based biometric template protection approach.

REFERENCES

- [1] Gartner, "Gartner Says Worldwide PC, Tablet and Mobile Phone Combined Shipments to Reach 2.4 Billion Units in 2013," available at: <u>https://www.gartner.com/newsroom/id/2408515</u>, [Accessed on 23 June 2018].
- [2] Emarketer, "Number of smartphone users worldwide from 2014 to 2020 (in billions)," available at: <u>https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/</u>, [Accessed on 23 June 2018].
- P. Chatterjee, "Constructing Mobile Multi-Sensor Systems Dominated by a Touch Display," available at: <u>https://www.digikey.com/en/articles/techzone/2011/sep/constructing-mobile-</u> <u>multi-sensor-systems-dominated-by-a-touch-display</u>. [Accessed on 23 June 2018],
- [4] Package Index, 2014. available at: <u>https://developer.android.com/reference/packages.html</u> [Accessed on 23 June 2018].
- [5] N. D. Lane, E. Miluzzo, H. Lu, D. Peebles, T. Choudhury and A. T. Campbell, "A survey of mobile phone sensing," *IEEE Communications Magazine*, vol. 48, no. 9, pp. 140-150, Sept. 2010.
- [6] Y. Zhang, G. Pan, K. Jia, M. Lu, Y. Wang and Z. Wu, "Accelerometer-based gait recognition by sparse representation of signature points with clusters," *IEEE Transactions on Cybernetics*, vol. 45, no. 9, pp. 1864-1875, Sept. 2015.
- [7] M. Ermes, J. PÄrkkÄ, J. MÄntyjÄrvi and I. Korhonen, "Detection of daily activities and sports with wearable sensors in controlled and uncontrolled conditions," *IEEE Transactions on Information Technology in Biomedicine*, vol. 12, no. 1, pp. 20-26, Jan. 2008.

- [8] O. Ojetola, E. I. Gaura and J. Brusey, "Fall detection with wearable sensors--safe (smart fall detection)," Seventh International Conference on Intelligent Environments, Nottingham, 2011, pp. 318-321.
- [9] B. Najafi, K. Aminian, F. Loew, Y. Blanc and P. A. Robert, "Measurement of stand-sit and sit-stand transitions using a miniature gyroscope and its application in fall risk evaluation in the elderly," *IEEE Transactions on Biomedical Engineering*, vol. 49, no. 8, pp. 843–851, Aug. 2002.
- [10] D. Impedovo and G. Pirlo, "Automatic signature verification: the state of the art," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 38, no. 5, pp. 609-635, Sept. 2008.
- S. Chen *et al.*, "Butler, not servant: A human-centric smart home energy management system," *IEEE Communications Magazine*, vol. 55, no. 2, pp. 27-33, Feb. 2017.
- [12] S. Mondal and P. Bours, "Person identification by keystroke dynamics using pairwise user coupling," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 6, pp. 1319-1329, June 2017.
- [13] Z. Liu, Z. Zhang, Q. Wu and Y. Wang, "Enhancing person re-identification by integrating gait biometric," *Neurocomputing, Elsevier*, vol. 168, pp. 1144-1156, Nov. 2015.
- [14] D. Lu, K. Xu and D. Huang, "A data driven in-air-handwriting biometric authentication system," *Proceedings of IEEE International Joint Conference on Biometrics (IJCB)*, pp. 531–537, Denver, CO, Feb. 2017.
- [15] K. Burda, "Authenticating users based on how they pick up smartphones," IIT SRC 2016: 12th Student Research Conference in Informatics and Information Technologies, April 2016.
- [16] M. Jin, Y. He, D. Fang, X. Chen, X. Meng and T. Xing, "iGuard: A real-time anti-theft system for smartphones," *IEEE Transactions on Mobile Computing*, Jan. 2018.

- [17] N. Kunnathu, "Biometric user authentication on smartphone accelerometer sensor data," *Proceedings of Student-Faculty Research Day*, CSIS, Pace University, 2015.
- [18] T. Feng, X. Zhao and W. Shi, "Investigating mobile device picking-up motion as a novel biometric modality," *IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, pp. 1–6, Sept. 2013.
- [19] M. Conti, I. Zachia-Zlatea and B. Crispo, "Mind how you answer me!: Transparently authenticating the user of a smartphone when answering or placing a call," *Proceedings of the 6th ACM Symposium on Information, Computer and Communications Security (ASIACCS)*, pp. 249–259, New York, NY, USA, 2011.
- [20] R. D. Findling, M. Muaaz, D. Hintze and R. Mayrhofer, "ShakeUnlock: Securely unlock mobile devices by shaking them together," *Proceedings of the 12th International Conference on Advances in Mobile Computing and Multimedia*, pp. 165–174, 2014.
- [21] H. Zhu, J. Hu, S. Chang and L. Lu, "ShakeIn: Secure user authentication of smartphones with single-handed shakes," *IEEE Transactions on Mobile Computing*, vol. 16, no. 10, pp. 2901-2912, Oct. 2017.
- [22] F. Hong, M. Wei, S. You, Y. Feng and Z. Guo, "Waving authentication: your smartphone authenticate you on motion gesture," *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pp. 263–266, 2015.
- [23] G. Ryu & S.-H. Kim and D. Choi, "Implicit Secondary Authentication for Sustainable SMS Authentication," *Sustainability*, MDPI, vol. 11, no. 1, pp. 1-14, January 2019.
- [24] Y. Li, H. Hu, G. Zhou and S. Deng, "Sensor-Based Continuous Authentication Using Cost-Effective Kernel Ridge Regression," *IEEE Access*, vol. 6, pp. 32554-32565, 2018.
- [25] M. Muaaz and R. Mayrhofer, "An analysis of different approaches to gait recognition using cell phone based accelerometers," *Proceedings of International Conference on Advances in Mobile Computing & Multimedia*, 2013.

- [26] C. C. Ho, C. Eswaran, K.-W. Ng and J.-Y. Leow, "An unobtrusive android person verification using accelerometer based gait," *Proceedings of the 10th International Conference on Advances in Mobile Computing & Multimedia*, pp. 271–274, 2012.
- [27] A. Primo, V. V. Phoha, R. Kumar and A. Serwadda, "Context-aware active authentication using smartphone accelerometer measurements," *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition Workshops* (CVPRW), pp. 98–105, 2014.
- [28] S. S. Arora, A. K. Jain and N. G. Paulter, "Gold fingers: 3d targets for evaluating capacitive readers," *IEEE transactions on information forensics and security*, vol. 12, no. 9, pp. 2067–2077, 2017.
- [29] A. K. Jain, A. Ross and S. Prabhakar, "An introduction to biometric recognition," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, no. 1, pp. 4-20, Jan. 2004.
- [30] L. C. F. Araujo, L. H. R. Sucupira, M. G. Lizarraga, L. L. Ling and J. Yabu-Uti, "User authentication through typing biometrics features," *IEEE Transactions on Signal Processing*, vol. 53, no. 2, pp. 851-855, Feb. 2005.
- [31] A. A. E. Ahmed and I. Traore, "A new biometric technology based on mouse dynamics," *IEEE Transactions on Dependable and Secure Computing*, vol. 4, no. 3, pp. 165-179, Aug. 2007.
- [32] C. Shen, Z. Cai, X. Guan, Y. Du and R. A. Maxion, "User authentication through mouse dynamics," *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 1, pp. 16-30, Oct. 2012.
- [33] A. A. Ahmed and I. Traore, "Biometric recognition based on free-text keystroke dynamics," *IEEE Transactions on Cybernetics*, vol. 44, no. 4, pp. 458-472, May 2013.
- [34] K. Venkataramani, S. Qidwai and B. Vijayakumar, "Face authentication from cell phone camera images with illumination and temporal variations," *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, vol. 35, no. 3, pp. 411-418, July 2005.

- [35] P. Abeni, M. Baltatu and R. D. Alessandro, "Implementing biometrics based authentication for mobile devices," *Proceedings of IEEE Global Telecommunications Conference (GlobeCom)*, pp. 1-5, Nov. 2006.
- [36] S. Chen, A. Pande and P. Mohapatra, "Sensor-assisted facial recognition: An enhanced biometric authentication system for smartphones," *Proceedings of the* 12th annual International Conference on Mobile systems, applications, and services (MobiSys), pp. 109-122, June 2014.
- [37] M. Baloul, E. Cherrier and C. Rosenberger, "Challenge-based speaker recognition for mobile authentication," *Proceedings of the International Conference of Biometrics Special Interest Group (BIOSIG)*, Darmstadt, 2012.
- [38] M. Kunz, K. Kasper, H. Reininger, M. Mobius and J. Ohms, "Continuous speaker verification in realtime," *Proceedings of International Conference of the Biometrics Special Interest Group*, pp. 79-87, 2011.
- [39] E. Miluzzo, C. Cornelius, A. Ramaswamy, T. Choudhury, Z. Liu and A. T. Campbell, "Darwin phones: The evolution of sensing and inference on mobile phones," *Proceedings of the 8th annual international conference on Mobile systems, applications, and services (MobiSys)*, pp. 5-20, June 2010.
- [40] M. O. Derawi, B. Yang and C. Busch, "Fingerprint recognition with embedded cameras on mobile phones," *Proceedings of MobiSec*, pp. 136-147, 2012.
- [41] X. Chen, J. Tian, Q. Su, X. Yang and F. Y. Wang, "A secured mobile phone based on embedded fingerprint recognition systems," *International Conference on Intelligence and Security Informatics*, vol. 3475, pp. 549-553, 2005.
- [42] N. L. Clarke, S. M. Furnell, B. M. Lines and P. L. Reynolds, "Keystroke dynamics on a mobile handset: A feasibility study," *Information Management and Computer Security*, vol. 11, no. 4, pp. 161-166, 2003.
- [43] P. Campisi, E. Maiorana, M. L. Boscoa and A. Neri, "User authentication using keystroke dynamics for cellular phones," *IET Signal Processing*, vol. 3, no. 4, pp. 333-341, 2009.

- [44] S. Hwang, S. Cho and S. Park, "Keystroke dynamics-based authentication for mobile devices," *Computers & Security*, vol. 28, no. 1–2, pp. 85-93, March 2009.
- [45] P. S. Teh, A. B. J. Teoh and S. Yue, "A survey of keystroke dynamics biometrics," *The Scientific World Journal*, vol. 2013, Article ID 408280, 24 pages, 2013.
- [46] K. Cheng and A. Kumar, "Contactless finger knuckle identification using smartphones," *Proceedings of the International Conference of Biometrics Special Interest Group (BIOSIG)*, Darmstadt, 2012.
- [47] K. R. Park, H. A. Park, B. J. Kang, E. C. Lee and D. S. Jeong, "A study on iris localization and recognition on mobile phones," *EURASIP Journal on Advances in Signal Processing*, pp. 1-12, 2008.
- [48] N. F. Moço, I. S. Técnico, I. de Telecomunicações and P. L. Correia, "Smartphone-based palmprint recognition system," 21st International Conference on Telecommunications (ICT), Lisbon, pp. 457-461, 2014.
- [49] Y. Han, T. Tan, Z. Sun and Y. Hao, "Embedded palmprint recognition system on mobile devices," *International Conference on Biometrics*, vol. 4642, pp. 1184-1193, 2007.
- [50] M. Franzgrote, C. Borg, B. J. T. Ries, S. Bussemaker, X. Jiang, M. Fieseler and L. Zhang, "Palmprint verification on mobile phones using accelerated competitive code," *International Conference on Hand-Based Biometrics, Hong Kong*, pp. 1-6, 2011.
- [51] D. J. Kim, K. W. Chung and K. S. Hong, "Person authentication using face, teeth and voice modalities for mobile device security," *IEEE Transactions on Consumer Electronics*, vol. 56, no. 4, pp. 2678-2685, Nov. 2010.
- [52] P. Tresadern *et al.*, "Mobile biometrics: Combined face and voice verification for a mobile platform," *IEEE Pervasive Computing*, vol. 12, no. 1, pp. 79-87, March 2013.

- [53] A. Hadid, J. Y. Heikkila, O. Silven and M. Pietikainen, "Face and eye detection for person authentication in mobile phones," *International Conference on Distributed Smart Cameras*, pp. 101-108, Oct. 2007.
- [54] D. J. Kim and K. S. Hong, "Multimodal biometric authentication using teeth image and voice in mobile environment," *IEEE Transactions on Consumer Electronics*, vol. 54, no. 4, pp. 1790-1797, Dec. 2008.
- [55] E. Khoury, L. E. Shafey, C. McCool, M. Gunther and S. Marcel, "Bi-modal biometric authentication on mobile phones in challenging conditions," *Image and Vision Computing*, vol. 32, no. 12, pp. 1147-1160, Dec. 2014.
- [56] S. H. Khan, M. A. Akbar, F. Shahzad, M. Farooq and Z. Khan, "Secure biometric template generation for multi-factor authentication," *Pattern Recognition*, vol. 48, no. 2, pp. 458-472, Feb. 2015.
- [57] C. McCool, S. Marcel, A. Hadid, M. Pietikainen, P. Matejka, J. Cernocky, N. Poh, J. Kittler, A. Larcher, C. Levy *et al.*, "Bi-modal person recognition on a mobile phone: using mobile phone data," *IEEE International Conference on Multimedia and Expo Workshops*, Melbourne, VIC, pp. 635-640, 2012.
- [58] W. Meng, D. S. Wong, S. Furnell and J. Zhou, "Surveying the development of biometric user authentication on mobile phones," *IEEE Communications Surveys* & *Tutorials*, vol. 17, no. 3, pp. 1268-1293, Dec. 2014.
- [59] A. Mahfouz, T. M. Mahmoud and A. S. Eldin, "A survey on behavioral biometric authentication on smartphones," *Journal of Information Security and Applications*, vol. 37, pp. 28-37, 2017
- [60] A. Alzubaidi and J. Kalita, "Authentication of smartphone users using behavioral biometrics," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 3, pp. 1998-2026, thirdquarter 2016.
- [61] T. J. Neal and D. L. Woodard, "Surveying biometric authentication for mobile device security," *Journal of Pattern Recognition Research*, vol. 1, pp. 74-110, 2016.

- [62] H. Saevanee and P. Bhattarakosol, "Authenticating user using keystroke dynamics and finger Pressure," 6th IEEE Consumer Communications and Networking Conference, pp. 1-2, Jan. 2009.
- [63] A. D. Luca, A. Hang, F. Brudy, C. Lindner and H. Hussmann, "Touch me once and I know it's you! Implicit authentication based on touch screen patterns," ACM Conference on Human Factors in Computing Systems, pp. 987-996, 2012.
- [64] M. Frank, R. Biedert, E. Ma, I. Martinovic and D. Song, "Touchalytics: On the applicability of touchscreen input as a behavioral biometric for continuous authentication," *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 1, pp. 136-148, Jan. 2013.
- [65] T. Feng, Z. Liu, K. A. Kwon, W. Shi, B. Carbunar, Y. Jiang and N. Nguyen, "Continuous mobile authentication using touchscreen gestures," *IEEE Conference* on Technologies for Homeland Security, pp. 451-456, Nov. 2012.
- [66] T. Feng, J. Yang, Z. Yan, E. M. Tapia and W. Shi, "TIPS: Context-aware implicit user identification using touch screen in uncontrolled environments," ACM *HotMobile*, Feb 2014.
- [67] Y. Meng, D. S. Wong, R. Schlegel and L. F. Kwok, "Touch gestures based biometric authentication scheme for touchscreen mobile phones," *Proc. 8th Int. Conf. INSCRYPT*, pp. 331-350, 2012.
- [68] N. Sae-Bae, N. Memon, K. Isbister and K. Ahmed, "Multitouch gesture-based authentication," *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 4, pp. 568-582, Jan. 2014.
- [69] M. Antal, L. Z. Szabo and Z. Bokor, "Identity information revealed from mobile touch gestures," 10th Joint Conference on Mathematics and Computer Science, Jan. 2014.
- [70] X. Zhao, T. Feng and W. Shi, "Continuous mobile authentication using a novel Graphic Touch Gesture Feature," *IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, pp. 1-6, 2013.

- [71] X. Zhao, T. Feng, W. Shi and I. Kakadiaris, "Mobile user authentication using statistical touch dynamics images," *IEEE Transactions on Information Forensics* and Security, vol. 9, no. 11, pp. 1780-1789, Aug. 2014.
- [72] J. Fierrez, A. Pozo, M. Martinez-Diaz, J. Galbally and A. Morales, "Benchmarking touchscreen biometrics for mobile authentication," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 11, pp. 2720-2733, Nov. 2018.
- [73] U. Mahbub, S. Sarkar, V. M. Patel and R. Chellappa, "Active user authentication for smartphones: A challenge data set and benchmark results," *IEEE International Conference on Biometrics: Theory, Applications, and Systems*, 2016.
- [74] R. Kumar, V. V. Phoha and A. Serwadda, "Continuous authentication of smartphone users by fusing typing swiping and phone movement patterns," *IEEE International Conference on Biometrics: Theory, Applications, and Systems*, pp. 1–8, 2016.
- [75] R. Kumar, P. P. Kundu and V. V. Phoha, "Continuous authentication using oneclass classifiers and their fusion," *IEEE International Conference on Identity*, *Security, and Behavior Analysis*, 2018.
- [76] Z. Sitová et al., "Hmog: New behavioral biometric features for continuous authentication of smartphone users," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 5, pp. 877–892, May 2016.
- [77] H. Zhang, V. M. Patel, M. Fathy and R. Chellappa, "Touch gesture-based active user authentication using dictionaries," *Proceedings of the IEEE Winter Conference on Applications of Computer Vision*, pp. 207–214, 2015.
- [78] A. Serwadda, V. V. Phoha and Z. Wang, "Which verifiers work?: A benchmark evaluation of touch-based authentication algorithms," *IEEE International Conference on Biometrics: Theory, Applications, and Systems*, pp. 1–8, 2013.
- [79] G. Li and P. Bours. 2018. "Studying WiFi and Accelerometer Data Based Authentication Method on Mobile Phones." 2nd International Conference on

Biometric Engineering and Applications. ACM, New York, NY, USA, pp. 18-23, 2018.

- [80] G. Li and P. Bours, "A Novel Mobilephone Application Authentication Approach based on Accelerometer and Gyroscope Data," *International Conference of the Biometrics Special Interest Group (BIOSIG)*, Darmstadt, pp. 1-4, 2018.
- [81] H. Lee, J. Y. Hwang, D. I. Kim, S. Lee, S.-H. Lee and J. S. Shin, "Understanding Keystroke Dynamics for Smartphone Users Authentication and Keystroke Dynamics on Smartphones Built-In Motion Sensors," *Security and Communication Networks*, vol. 2018, Article ID 2567463, 2018.
- [82] M. P. Centeno, A. v. Moorsel and S. Castruccio, "Smartphone Continuous Authentication Using Deep Learning Autoencoders," 15th Annual Conference on Privacy, Security and Trust (PST), Calgary, Alberta, Canada, pp. 147-1478, 2017.
- [83] M. P. Centeno, Y. Guan and A. v. Moorsel, "Mobile Based Continuous Authentication Using Deep Features." 2nd International Workshop on Embedded and Mobile Deep Learning (EMDL), ACM, New York, NY, USA, pp. 19-24, 2018.
- [84] P. S. Teh, N. Zhang, A. B. J. Teoh and K. Chen, "A survey on touch dynamics authentication in mobile devices," *Computers & Security*, vol. 59, pp. 210-235, June 2016.
- [85] C. B. Ng, Y. H. Tay and B. M. Goi, "Vision-based human gender recognition: A survey," *Computer Vision and Pattern Recognition*, arXiv preprint arXiv:1204.1611, 2012.
- [86] C. Chen and A. Ross, "Evaluation of gender classification methods on thermal and near-infrared face images," *International Joint Conference on Biometrics* (*IJCB*), pp. 1-8, Oct. 2011.
- [87] X. Lu, H. Chen and A. K. Jain, "Multimodal facial gender and ethnicity identification," *Proceedings Advances in Biometrics*, vol. 3832, pp. 554-561, 2006.

- [88] T. Danisman, I. M. Bilasco and J. Martinet, "Boosting gender recognition performance with a fuzzy inference system," *Expert Systems with Applications*, vol. 42, no. 5, pp. 2772-2784, April 2015.
- [89] P. Gnanasivam and S. Muttan, "Gender classification using ear biometrics," Proceedings of the 4th International Conference on Signal and Image Processing (ICSIP), vol. 222, pp. 137-148, 2012.
- [90] X. Li, X. Zhao, Y. Fu and Y. Liu, "Bimodal gender recognition from face and fingerprint," *IEEE Conference on Computer Vision and Pattern Recognition* (CVPR), pp. 2590-2597, June 2010.
- [91] M. Wu and Y. Yuan, "Gender classification based on geometry features of palm image," *Hindawi Publishing Corporation, The Scientific World Journal*, vol. 2014, pp. 1-7, April 2014.
- [92] V. Thomas, N. V. Chawla, K. W. Bowyer and P. J. Flynn, "Learning to predict gender from iris images," 1st IEEE International Conference on Biometrics: Theory, Applications, and Systems, Crystal City, VA, pp. 1-5, 2007.
- [93] J. E. Tapia, C. A. Perez and K. W. Bowyer, "Gender classification from the same iris code used for recognition," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 8, pp. 1760-1770, Aug. 2016.
- [94] P. Maji, S. Chatterjee, S. Chakraborty, N. Kausar, S. Samanta and N. Dey, "Effect of Euler number as a feature in gender recognition system from offline handwritten signature using neural networks," 2nd International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, pp. 1869-1873, 2015.
- [95] M. Hu, Y. Wang, Z. Zhang and D. Zhang, "Gait-based gender classification using mixed conditional random field," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 41, no. 5, pp. 1429-1439, Oct. 2011.
- [96] X. Li, S. J. Maybank, S. Yan, D. Tao and D. Xu, "Gait components and their application to gender recognition," *IEEE Transactions on Systems, Man, and*

Cybernetics, Part C: Applications and Reviews, vol. 38, no.2, pp. 145–155, March 2008.

- [97] J. Lu, G. Wang and P. Moulin, "Human identity and gender recognition from gait sequences with arbitrary walking directions," *IEEE Transactions on information Forensics and Security*, vol. 9, no. 1, pp. 51–61, Jan. 2014.
- [98] D. G. Childers, K. Wu, K. S. Bae and D. M. Hicks, "Automatic recognition of gender by voice," *International Conference on Acoustics, Speech, and Signal Processing, ICASSP-88, New York, NY, vol. pp. 603-606, 1988.*
- [99] M. Kotti and C. Kotropoulos, "Gender classification in two emotional speech databases," 19th International Conference on Pattern Recognition, Tampa, FL, pp. 1-4, 2008.
- [100] I. Tsimperidis, A. Arampatzis and A. Karakos, "Keystroke dynamics features for gender recognition," *Digital Investigation*, vol. 24, pp. 4-10, March 2018.
- [101] R. Giot and C. Rosenberger, "A new soft biometric approach for keystroke dynamics based on gender recognition," *International Journal of Information Technology and Management (IJITM), Special Issue on: Advances and Trends in Biometrics by Dr. Lidong Wang*, vol. 11, no. 1/2, pp. 35-49, 2012.
- [102] C. Peersman, W. Daelemans and L. V. Vaerenbergh, "Predicting age and gender in online social networks," *Proceedings of the 3rd international workshop on Search and mining user-generated contents* (SMUC '11), ACM, New York, NY, USA, pp. 37-44, 2011.
- [103] N. V. Balen, C. T. Ball and H. Wang, "A behavioral biometrics based approach to online gender classification," *SecureComm 2016*, pp. 475-495, 2016.
- [104] S. Yu, T. Tan, K. Huang, K. Jia and X. Wu, "A study on gait-based gender classification," *IEEE Transactions on Image Processing*, vol. 18, no. 8, pp. 1905-1910, Aug. 2009.
- [105] L. E. Shafey, E. Khoury and S. Marcel, "Audio-visual gender recognition in uncontrolled environment using variability modeling techniques," *IEEE International Joint Conference on Biometrics*, Clearwater, FL, pp. 1-8, 2014.

- [106] C. Shan, S. Gong and P. W. McOwan, "Fusing gait and face cues for human gender recognition," *Neurocomputing*, vol. 71, no. 10-12, pp. 1931-1938, June 2008.
- [107] D. Zhang and Y. Wang, "Gender recognition based on fusion of face and multiview gait," Advances in Biometrics Lecture Notes in Computer Science, vol. 5558, pp. 1010-1018, 2009
- [108] A. K. Jain, K. Nandakumar, X. Lu and U. Park, "Integrating faces, fingerprints, and soft biometric traits for user recognition," *Biometric Authentication Lecture Notes in Computer Science*, vol. 3087, pp. 259-269, 2004.
- [109] U. Park and A. K. Jain, "Face matching and retrieval using soft biometrics," *IEEE Transactions on Information Forensics and Security*, vol. 5, no. 3, pp. 406-415, Sept. 2007.
- [110] S. Z. Syed Idrus, E. Cherrier, C. Rosenberger, S. Mondal and P. Bours, "Keystroke dynamics performance enhancement with soft biometrics," *IEEE International Conference on Identity, Security and Behavior Analysis (ISBA 2015)*, Hong Kong, pp. 1-7, 2015.
- [111] O. O. Ogunduyile, K. Zuva, O. A. Randle and T. Zuva, "Ubiquitous healthcare monitoring system using integrated triaxial accelerometer, S_PO₂ and location sensors," *International Journal of UbiComp*, vol. 4, no. 2, pp. 1-13, April 2013.
- [112] Y. Ren, Y. Chen, M. C. Chuah and J. Yang, "User verification leveraging gait recognition for smartphone enabled mobile healthcare systems," *IEEE Transactions on Mobile Computing*, vol. 14, no. 9, pp. 1961-1974, Sept. 2015.
- [113] A. Minutolo, M. Esposito and G. Pietro, "Design and validation of a light-weight reasoning system to support remote health monitoring applications," *Engineering Applications of Artificial Intelligence*, vol. 41, pp. 232-248, May 2015.
- [114] M. Alhussein, Z. Ali, M. Imran and W. Abdul, "Automatic gender detection based on characteristics of vocal folds for mobile healthcare system," *Mobile Information Systems*, vol. 2016, Article ID. 7805217, 2016.

- [115] O. Miguel-Hurtado, S. V. Stevenage, C. Bevan and R. Guest, "Predicting sex as a soft-biometrics from device interaction swipe gestures," *Pattern Recognition Letters*, vol. 79, pp. 44-51, 2016.
- [116] Y. Wang, Y. Tang, J. Ma and Z. Qin, "Gender prediction based on data streams of smartphone applications," *International Conference on Big Data Computing and Communications*, pp. 115-125, 2015.
- [117] A. Dantcheva, P. Elia and A. Ross, "What else does your biometric data reveal? A survey on soft biometrics," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 3, pp. 441-467, March 2016.
- [118] L. Vergeest, "Uses, gratifications and gender in smartphone use. Academia.edu. Tilburg University.
- [119] E. Eidinger, R. Enbar and T. Hassner, "Age and gender estimation of unfiltered faces," *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 12, pp. 2170-2179, Dec. 2014.
- [120] A. Rattani, N. Reddy and R. Derakhshani, "Gender prediction from mobile ocular images: A feasibility study," *IEEE International Symposium on Technologies for Homeland Security (HST)*, Waltham, MA, pp. 1-6, 2017.
- [121] A. Agneessens, I. Bisio, F. Lavagetto and M. Marchese, "Design and implementation of smartphone applications for speaker count and gender recognition," *The Internet of Things: 20th Tyrrhenian Workshop on Digital Communication*, 2010.
- [122] C. Sarraute, P. Blanc and J. Burroni, "A Study of age and gender seen through mobile phone usage patterns in Mexico," *IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, pp. 836–843, 2014
- [123] Y. Choi, Y. Kim, S. Kim, K. Park and J. Park, "An on-device gender prediction method for mobile users using representative wordsets," *Expert Systems with Applications*, vol. 64, pp. 423-433, 2016.

- [124] A. Buriro, Z. Akhtar, B. Crispo and F. Del Frari, "Age, gender and operatinghand estimation on smart mobile devices," *International Conference of the Biometrics Special Interest Group (BIOSIG)*, Darmstadt, pp. 1-5, 2016.
- [125] M. Antal and G. Nemes, "Gender recognition from mobile biometric data," 11th International Symposium on Applied Computational Intelligence and Informatics (SACI), pp. 243-248, 2016.
- [126] M. Antal, Z. Bokor and L. Z. Szabó, "Information revealed from scrolling interactions on mobile devices," *Pattern Recognition Letters*, vol. 56, pp. 7-13, 2015.
- [127] L. Igual. A. Lapedriza and R. Borras, "Robust gait-based gender classification using depth cameras," *EURASIP Journal on Image and Video Processing*, 2013.
- [128] D. Kastaniotis, I. Theodorakopoulos, G. Economou and S. Fotopoulos, "Gaitbased gender recognition using pose information for real-time applications," 18th International Conference on Digital Signal Processing (DSP), 2013.
- [129] S. Sprager and M. B. Juric, "Inertial sensor-based gait recognition: A review," Sensors, vol. 15, no. 9, pp. 22089–22127, 2015.
- [130] H. M. Thang, V. Q. Viet, N. D. Thuc and D. Choi, "Gait identification using accelerometer on mobile phone," *International Conference on Control, Automation and Information Sciences (ICCAIS)*, pp. 344-348, 2012.
- [131] Y. Zhong and Y. Deng, "Sensor orientation invariant mobile gait biometrics," *IEEE International Joint Conference on Biometrics (IJCB)*, 2014.
- [132] Q. Riaz, A. Vögele, B. Krüger and A. Weber, "One small step for a man: Estimation of gender, age and height from recordings of one step by a single inertial sensor," *Sensors (Basel, Switzerland)*, vol. 15, no. 12, pp. 31999–32019, 2015.
- [133] G. M. Weiss and J. W. Lockhart, "Identifying user traits by mining smart phone accelerometer data," 5th International Workshop on Knowledge Discovery from Sensor Data, 61-69, Aug. 2011.

- [134] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann and I. H. Witten,
 "The WEKA data mining software: an update," *SIGKDD Explorations*, vol. 11, no. 1, pp. 10-18, 2009.
- [135] A. Jain and V. Kanhangad, "Investigating gender recognition in smartphones using accelerometer and gyroscope sensor readings," *IEEE International Conference on Computational Techniques in Information and Communication Technologies*, pp. 597-602, 2016.
- [136] W. Zjlstra and K. Aminian, "Mobility assessment in older people: New possibilities and challenges," *European Journal Aging*, vol. 4, pp. 3–12, 2007.
- [137] C. E. Matthews, S. M. George, S. C. Moore, *et al.*, "Amount of time spent in sedentary behaviors and cause-specific mortality in US adults," *American Journal Clinical Nutrition*, vol. 95, no. 2, pp. 437-445, 2012.
- [138] P. T. Katzmarzyk and I. Lee, "Sedentary behaviour and life expectancy in the USA: A cause-deleted life table analysis", *BMJ Open*, 2012.
- [139] J. Wannenburg and R. Malekian, "Physical activity recognition from smartphone accelerometer data for user context awareness sensing," *IEEE Transactions Systems, Man, and Cybernetics*, 2016.
- [140] A. Wang, G. Chen, J. Yang, S. Zhao and C. Y. Chang, "A comparative study on human activity recognition using inertial sensors in a smartphone," *IEEE Sensors Journal*, vol. 16, no. 11, pp. 4566-4578, 2016.
- [141] O. D. Incel, M. Kose and C. Ersoy, "A review and taxonomy of activity recognition on mobile phones," *BioNanoSci*, vol. 3, no. 2, pp. 145-171, 2013.
- [142] O. D. Lara and M. A. Labrador, "A survey on human activity recognition using wearable sensors," *IEEE Communication Surveys Tutorials*, vol. 15, no. 3, pp. 1192-1209, 2013.
- [143] T. Guha and R. K. Ward, "Learning sparse representations for human action recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 8, pp. 1576-1588, 2012.

- [144] L. Bao and S.S. Intille, "Activity recognition from user-annotated acceleration data," In: Ferscha A., Mattern F. (eds) *Pervasive Computer*, LNCS, vol. 3001, 2004.
- [145] J. Mantyjarvi, J. Himberg and T. Seppanen, "Recognizing human motion with multiple acceleration sensors," *IEEE International Conference on Systems, Man* and Cybernetics, vol. 2, pp. 747-752, 2001.
- [146] E. M. Tapia *et al.*, "Real-time recognition of physical activities and their intensities using wireless accelerometers and a heart rate monitor," *Proceedings* 11th IEEE International Symposium on Wearable Computer, 2007, pp. 37-40.
- [147] G. Bieber, J. Voskamp, B. Urban and C. Stephanidis, "Activity recognition for everyday life on mobile phones," *Lecture Notes in Computer Science, Springer*, vol. 5615, pp. 289-296, 2009.
- [148] J. R. Kwapisz, G. M. Weiss and S. A. Moore, "Activity recognition using cell phone accelerometers," *SIGKDD Explorations Newsletter*, vol. 12, no. 2, pp. 74-82, 2011.
- [149] M. Lv, L. Chen, T. Chen and G. Chen, "Bi-view semi-supervised learning based semantic human activity recognition using accelerometers," *IEEE Transactions* on Mobile Computing, January 2018.
- [150] S. Dernbach, B. Das, N. C. Krishnan, B. L. Thomas and D. J. Cook, "Simple and complex activity recognition through smart phones," 8th International Conference on Intelligent Environments, pp. 214-221, 2012.
- [151] Y. Kwon, K. Kang and C. Bae, "Unsupervised learning for human activity recognition using smartphone sensors," *Expert Systems with Applications*, vol. 41, no. 14, pp. 6067-6074, 2014.
- [152] Y. Chen and C. Shen, "Performance analysis of smartphone-sensor behavior for human activity recognition," *IEEE Access*, vol. 5, pp. 3095-3110, 2017.
- [153] M. Shoaib, S. Bosch, O. D. Incel, H. Scholten and P. J. M. Havinga, "Complex human activity recognition using smartphone and wrist-worn motion sensors," *Sensors*, vol. 16, no. 4, pp. 426, 2016.
- [154] X. Su, H. Tong and P. Ji, "Activity recognition with smartphone sensors," *Tsinghua Science and Technology*, vol. 19, no. 3, pp. 235-249, 2014.
- [155] M. Shoaib, S. Bosch, O.D. Incel, H. Scholten and P.J. Havinga, "A survey of online activity recognition using mobile phones," *Sensors*, vol. 15, pp. 2059– 2085, 2015.
- [156] D. Anguita, A. Ghio, L. Oneto, X. Parra and J. L. Reyes-Ortiz, "A public domain dataset for human activity recognition using smartphones," *Proceedings of 21st ESANN, Bruges, Belgium*, pp. 437-442, 2013.
- [157] W. Wu, S. Dasgupta, E. E. Ramirez, C. Peterson and G. J. Norman, "Classification accuracies of physical activities using smartphone motion sensors," *Journal Medical Internet Research*, vol. 14, no. 5, pp. 130, 2012.
- [158] P. A. Devijver and J. Kittler, Pattern Recognition: A Statistical Approach, London, GB: Prentice-Hall, 1982.
- [159] M. P. Dubuisson and A. K. Jain, "A modified Hausdorff distance for object matching," *Proceedings of the 12th International Conference on Pattern Recognition*, pp. 566-568, 1994.
- [160] A. Jain, K. Nandakumar and A. Ross, "Score normalization in multimodal biometric systems," *Pattern Recognition*, vol. 38, no. 12, pp. 2270-2285.
- [161] W. J. Scheirer, A. Rocha, R. Micheals and T. E. Boult, "Robust fusion: Extreme value theory for recognition score normalization," *European Conference on Computer Vision, Springer*, pp. 481-495, 2010.
- [162] A. Ross and A. Jain, "Information fusion in biometrics," *Pattern Recognition Letters*, vol. 24, no. 13, pp. 2115–2125, 2003.
- [163] J. Kittler, M. Hatef, R. P. Duin and J. G. Matas, "On combining classifiers," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 3, pp. 226–239, 1998.

- [164] A. Jain and V. Kanhangad, "Exploring orientation and accelerometer sensor data for personal authentication in smartphones using touchscreen gestures", *Pattern Recognition Letters*, vol. 68, pp. 351-360, Dec. 2015.
- [165] A. Oliva and A. Torralba, "Modeling the shape of the scene: a holistic representation of the spatial envelope," *International Journal of Computer Vision*, vol. 42, no. 3, pp. 145–175, 2001.
- [166] Y. Wang, Y. Li and X. Ji, "Recognizing human actions based on GIST descriptor and word phrase," *International Conference on Mechatronic Sciences, Electric Engineering and Computer (MEC)*, Shengyang, pp. 1104-1107, 2013.
- [167] G. Tanisik, C. Zalluhoglu and N. Ikizler-Cinbis, "Facial descriptors for human interaction recognition in still images," *Pattern Recognition Letters*, vol. 73, pp. 44-51, 2016.
- [168] J. Hayes and A. Efros, "Scene completion using millions of photographs," *SIGGRAPH*, 2007.
- [169] A. Torralba, K. P. Murphy, W. T. Freeman and M. A. Rubin, "Context-based vision system for place and object recognition," 9th IEEE International Conference on Computer Vision, Nice, France, vol.1, pp. 273-280, 2003.
- [170] M. Douze, H. Jégou, H. Sandhawalia, L. Amsaleg and C. Schmid, "Evaluation of GIST descriptors for web-scale image search," *Proceedings of the ACM International Conference on Image and Video Retrieval* (CIVR '09), ACM, New York, NY, USA, Article 19, 2009.
- [171] R. Kohavi and G. H. John, "Wrappers for feature subset selection," Artificial Intelligence, vol. 97, no. 1-2, pp. 273-324, Dec. 1997.
- [172] T. Mitchell, Machine Learning, McGraw-Hill, Inc., New York, NY, 1997.
- [173] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 886-893, 2005.

- [174] A. Albiol, D. Monzo, A. Martin, J. Sastre and A. Albiol, "Face recognition using HOG-EBGM," *Pattern Recognition Letters*, vol. 29, no. 10, pp, 1537–1543, 2008.
- [175] B. Su, S. Lu, S. Tian, J. H. Lim and C. L. Tan, "Character recognition in natural scenes using convolutional co-occurrence HOG," 22nd International Conference on Pattern Recognition, Stockholm, pp. 2926-2931, 2014.
- [176] S. K. Berkaya, H. Gunduz, O. Ozsen, C. Akinlar and S. Gunal, "On circular traffic sign detection and recognition," *Expert Systems with Applications*, vol. 48, pp. 67-75, 2016.
- [177] R. Arroyo, J. J. Yebes, L. M. Bergasa, I. G. Daza and J. Almazán, "Expert videosurveillance system for real-time detection of suspicious behaviors in shopping malls," *Expert Systems with Applications*, vol. 42, no. 21, pp. 7991-8005, 2015.
- [178] L. Cao, M. Dikmen, Y. Fu and T. S. Huang, "Gender recognition from body," Proceedings of the 16th ACM International Conference on Multimedia (MM '08), pp. 725-728, 2008.
- [179] L. Breiman, Bagging predictors, *Springer Machine learning*, vol. 24, no. 2, pp. 123-140, 1996.
- [180] R. Duda, P. Hart and D. Stork, Pattern classification, 2nd edition John Wiley & Sons, 2001
- [181] M. W. Tahir, N. A. Zaidi, R. Blank, P. P. Vinayaka, M. J. Vellekoop and W. Lang, "Fungus detection through optical sensor system using two different kinds of feature vectors for the classification," *IEEE Sensors Journal*, vol. 17, no. 16, pp. 5341-5349, Aug. 2017.
- [182] L. Wang, T. Tan, H. Ning and W. Hu, "Silhouette analysis-based gait recognition for human identification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 12, pp. 1505-1518, 2003.
- [183] C. T. Zahn and R. Z. Roskies, "Fourier descriptors for plane closed curves," *IEEE Transactions on Computers*, vol. C-21, no. 3, pp. 269-281, 1972.

- [184] H. Kauppinen, T. Seppanen and M. Pietikainen, "An experimental comparison of autoregressive and Fourier-based descriptors in 2D shape classification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, no. 2, pp. 201-207, 1995.
- [185] D. Zhang and G. Lu, "A comparative study on shape retrieval using Fourier descriptors with different shape signatures," *Proceedings 5th ACCV*, pp. 646-651, 2002.
- [186] Y. Ma, P. Ding, Q. Li, W. Lu, J. Yang and W. Yao, "A vision-based precipitation sensor for detection and classification of hydrometeors," *IEEE Sensors Journal*, vol. 16, no. 11, pp. 4546-4554, June 2016.
- [187] E. Allwein, R. Schapire and Y. Singer, "Reducing multiclass to binary: A unifying approach for margin classifiers," *Journal Machine Learning Research*, vol. 1, pp. 113–141, 2000.
- [188] C. Cortes and V. Vapnik, Support vector networks, *Machine Learning*, vol. 20, no. 3, pp. 273-297, 1995.
- [189] A. Wang, N. An, G. Chen, L. Li and G. Alterovitz, "Predicting hypertension without measurement: A non-invasive, questionnaire-based approach," *Expert Systems with Applications*, vol. 42, no. 21, pp. 7601-7609, 2015.
- [190] L. Rokach, "Ensemble-based classifiers," *Artificial Intelligence Review*, vol. 33, pp. 1-39, 2010.
- [191] M. Fernández-Delgado, E. Cernadas, S. Barro and D. Amorim, "Do we need hundreds of classifiers to solve real world classification problems?," *Journal of Machine Learning Research*, vol. 15, pp. 3133–3181, 2014.
- [192] J. Demšar, "Statistical comparisons of classifiers over multiple data sets," *Journal of Machine Learning Research*, vol. 7, pp. 1-30, 2006.
- [193] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry and Y. Ma, "Robust face recognition via sparse representation," *IEEE Transactions on Pattern Analysis* and Machine Intelligence, vol. 31, no. 2, pp. 210-227, Feb. 2009.