M.Sc. Engg. (CSE) Thesis

Leveraging a Wearable for Activity Recognition in Salat

Submitted by

Ishrat Jahan

1017052011

Supervised by

Dr. A. B. M. Alim Al Islam

Submitted to

Department of Computer Science and Engineering

Bangladesh University of Engineering and Technology

Dhaka, Bangladesh

in partial fulfillment of the requirements for the degree of

Master of Science in Computer Science and Engineering

May 2023

Candidate's Declaration

I, do, hereby, certify that the work presented in this thesis, titled, "Leveraging a Wearable for Activity Recognition in Salat", is the outcome of the investigation and research carried out by me under the supervision of Dr. A. B. M. Alim Al Islam, Professor, Department of CSE, BUET.

I also declare that neither this thesis nor any part thereof has been submitted anywhere else for the award of any degree, diploma or other qualifications.

Ishrat Jahan
1017052011

The thesis titled "**Leveraging a Wearable for Activity Recognition in Salat**", submitted by Ishrat Jahan, Student ID 1017052011, Session October 2017, to the Department of Computer Science and Engineering, Bangladesh University of Engineering and Technology, has been accepted as satisfactory in partial fulfilment of the requirements for the degree of Master of Science in Computer Science and Engineering and approved as to its style and contents on May 2, 2023.

Board of Examiners

1. ________________________

Dr. A. B. M. Alim Al Islam — Chairman (Supervisor)
Professor
Department of CSE, BUET, Dhaka

2. ________________________

Dr. Mahmuda Naznin — Member (Ex-Officio)
Professor and Head
Department of CSE, BUET, Dhaka

3. ________________________ 02.05.2023

Dr. Muhammad Masroor Ali — Member
Professor
Department of CSE, BUET, Dhaka

4. ________________________

Dr. Mohammed Eunus Ali — Member
Professor
Department of CSE, BUET, Dhaka

5. ________________________

Dr. Md. Abdur Razzaque — Member (External)
Professor
Department of CSE
University of Dhaka, Dhaka

Acknowledgement

First of all, I thank the Almighty for blessing me with the opportunity, strength, and resources to work on and complete this project. I would like to express my heartfelt gratitude to my supervisor, Dr. A. B. M. Alim Al Islam for his continuous guidance, encouragement, and kind support throughout this whole project. His thoughtful directions and supervision were invaluable.

I extend my appreciation to the Deputyship for Research & Innovation, Ministry of Education in Saudi Arabia for funding this research work through project 764 number (DRI-KSU-762). Special thanks to Prof. Dr. Najla Abdulrahman Al-Nabhan for her kind collaboration.

Besides, I would like to thank the honorable members of my thesis committee: Prof. Dr. Mahmuda Naznin, Prof. Dr. Muhammad Masroor Ali, Prof. Dr. Mohammed Eunus Ali, and especially the external member Prof. Dr. Md. Abdur Razzaque, for their encouragement, insightful comments, and valuable feedback.

I am also thankful to my fellow research-mates Masfiqur Rahaman (Master's student, CSE, BUET), Jannatun Noor (Ph.D. student, CSE, BUET), Md Shihabul Islam (Ph.D. student, CSE, BUET), and Tarik Reza Toha. They were always there whenever I sought any help or suggestions while undertaking this research study. I would also like to thank everyone who participated in this study for their valuable time and effort.

Last but not the least, I would like to take this opportunity to thank my loving parents for their unconditional love, amazing support, and unwavering encouragement throughout my whole life. They tried their best to arrange the best possible education for me. Their prayers for me are what brought me this far. Hope I can make them more proud in the upcoming days!

Dhaka
May 2, 2023

Ishrat Jahan
1017052011

Contents

List of Figures

List of Tables

List of Algorithms

Abstract

Salat, the most important worship of Muslims and the second pillar of Islam, is an integral part of the Muslim community. Salat involves a series of steady and transitional activities to be performed in a specific sequence. In the HAR literature, such activities are termed as complex activities [6, 7], which is the most challenging category of activities to recognize [6]. On top of that, Salat has variations based on time, priority, school of thought, etc. The activities in Salat, the postures of the same activity, etc., differ greatly based on these aspects. There also exist chances that people make different mistakes in Salat, such as forgetting an activity or adding an extra activity while performing Salat. Considering all these facts, it is understandable that recognizing activities in Salat is indeed a challenging task.

Existing research studies related to recognizing individual activities in Salat either demand capturing images by a camera for image processing or carrying a smartphone (sometimes in inconvenient places) while praying for capturing sensor signals. Both of the demands are not convenient or applicable in real cases. Besides, the existing studies lack user-independent accuracy analysis and fine-grained prediction. Therefore, the literature is yet to provide a convenient, and user-independent solution for activity recognition in Salat that is applicable in general. To address this gap, in this study, we first assess the requirement and acceptability of technological solutions for activity recognition in Salat by conducting an online survey. Here, the objective is to analyze whether people need and are willing to explore such solutions for their Salat. Our key findings from the analysis are that mistakes in prayer are common, and people, in general, are eager to explore convenient technological assistance for improving their Salat through corrective measures.

Upon establishing the requirement, to go further in addressing the gap in the literature, we propose an activity recognition methodology using a smartwatch to recognize different activities in Salat, such as standing, bowing, prostrating, etc., so that a user can assess different aspects of his prayer such as correctness, completeness, etc., of his prayer. We prepare a Salat activity dataset through collecting data from 30 subjects using a smartwatch. Upon preprocessing the collected raw data, we separate the steady and transitional states of Salat following two different ways - 1) using machine learning, and 2) using Signal Magnitude Area (SMA). Afterwards, through applying some semantic rules derived from domain knowledge, we recognize a couple of steady states, namely bowing and standing, with perfect accuracy. Next, we develop a pattern database for the transitional activities in Salat. We classify the transitional activities using Dynamic Time Warping (DTW) algorithm based on the best-matching template set. Finally, a

postprocessing algorithm detects misclassifications and applies domain knowledge again to correct them to provide the final predictions. In the process of performing all these tasks, we propose a new pipeline of methodology for activity recognition in Salat.

We perform user-independent analysis over the performance of our proposed methodology and achieve a near-perfect accuracy (99.3%). Along with this very high accuracy for the first time in the literature, our model offers fine-grained recognition of the individual activities in Salat. Besides, our methodology is robust enough to overlook the extra transitional activities a person performs while praying, which does not nullify Salat. To demonstrate the robustness of our methodology and the implication of each of these steps of the methodology, we perform rigorous experimentation and analysis. Findings on the experimentation and analysis demonstrate step-by-step advancement towards achieving near-perfect accuracy in real-life settings. Therefore, this research, covering the ground study on user requirements and developing a new methodology to achieve near-perfect accuracy in activity recognition in Salat, is expected to lead towards a comprehensive solution for monitoring Salat. The solution, in the future, could provide a detailed summary with appropriate feedback to a worshipper to help him improve his quality of Salat.

Chapter 1

Introduction

Human Activity Recognition (HAR) has become an active research area for more than a decade due to its numerous applications in various domains. By definition, HAR implies detecting and classifying human activities from time series sensor data [8]. Thus, this branch of literature aims to define and test novel approaches for accurately recognizing human activities using signals, images, etc., collected through different types of devices while performing those activities. Generally, human activities are composed of single or multiple postures, actions, and interactions. Throughout the last two decades, researchers have attempted to recognize activities of varying complexities by adopting various approaches. Because of the technological advancements during this time, HAR trends change quickly and necessitate an up-to-date perspective.

Over the years, researchers have explored different types of sensing technology in HAR studies such as body-worn sensors, cameras, wearable sensors, etc., to improve recognition performance. Early research on activity recognition focuses primarily on computer vision-based activity recognition [9, 10]. However, this approach requires placing capturing devices such as cameras in strategic positions to capture the activities of people. Although these activity monitoring methods can efficiently recognize activities, they are less suitable for many indoor environments, especially where privacy is a concern. Furthermore, the accuracy of vision-based approaches is greatly affected by occlusion, change in illumination or background, etc., which greatly limits their practical use [11]. These limitations eventually push researchers to focus more on sensor-based activity recognition. In sensor-based methods, data is collected from one or multiple sensors, and later processed and classified using different techniques such as machine learning, template matching, etc., In recent times, object sensors, environmental sensors, and wearable sensor devices have been widely used in HAR studies. Among the alternatives, wearable sensors require attaching physical sensors to humans so that they can still carry on necessary activities without any hamper or distraction. In this regard, recent studies are mostly geared toward leveraging the ubiquity, ease-of-use, and self-sufficiency of smart devices such as smartphones, smartwatches, etc., [12, 13], as they are equipped with all the essential sensors necessary for activity recognition. Utilizing the integrated sensors of these devices offers an unprecedented opportunity for an automated approach to perform activity recognition in our daily life. Consequently, in the last decade, many HAR research studies have been carried out using these devices as data collection tools.

Owing to the advancements in data collection tools, HAR research is now penetrating almost all aspects of our lives. HAR has become one of the emerging areas that have garnered the interest of researchers to apply it in different domains. Practical deployments of HAR have been done in fall detection [14, 15], behavioral monitoring [16], gait analysis [17, 18], Ambient Assisted Living (AAL) [19], surveillance systems [20], sports coaching [21], and many more [22]. Thus activities under HAR research have become ever spreading.

Different Types of Activities under Recognition

Activities that have been covered by HAR research so far are huge in number, and diverse in nature - from walking, running, and cycling to kitchen activities, construction activities, and whatnot. Moreover, they vary greatly in complexity. Researchers classify activities into two main categories - simple and complex. Simple activities refer to single repetitive actions, whereas complex activity refers to sequences of multiple simple activities or interleaved activities. Walking, cycling, sitting, etc., are examples of simple activities whereas driving a car, preparing a meal, shopping, taking a bus, etc., are examples of complex activities.

The most commonly studied activities in HAR research are walking, running, biking, jogging, remaining still, walking upstairs, and walking downstairs [23]. However, the set of activities under recognition differs from one study to another. Some studies focus on exercise activities [24], while others approach construction activities [25], kitchen activities [26], and many more. On the other hand, gait analysis [17] and fall detection [14, 15] are some other widely explored HAR domains due to their important applications in healthcare.

Simple activities are explored extensively in the literature with greater success to date. On the other hand, complex activity recognition is yet a less-explored area [6]. Moreover, it can be easily understood as well as proven from the literature that complex activities are more challenging to recognize than the simpler ones [27].

Salat - A Complex Religious Activity under Recognition

In this study, we focus on a particular human activity, which is Muslim prayer or Islamic prayer activity known as Salat. Salat consists of repeating units called Rakah (plural Rakat) consisting of several predefined steps such as standing, bowing, prostrating, etc., all of which must be performed maintaining the exact sequence and postures [28, 29]. Salat involves several different activities, covering both static/steady and transitional activities, in an alternating manner. Table

Table 1.1: Different activities in Salat

State	Activities
Static/Steady (S)	Standing (St) Bowing (B) Short-standing (Sh) Prostrating (P) Sitting (S)
Transitional (T)	Standing to bowing (St-B) Bowing to short-standing (B-Sh) Short-standing to prostrating (Sh-P) Prostrating to sitting (P-S) Sitting to prostrating (S-P) Prostrating to standing (P-St) Takbeer (Tk)

1.1 presents a list of these activities. A person has to perform several simple activities (as shown in the table) one after another in Salat [6, 30]. Accordingly, Salat, by its very nature, falls under the category of complex activities. Moreover, Salat exhibits variations in its steps, the number of units to pray, and the postures performed based on different factors such as time, priority, school of thought, etc. These variations further increase the complexity of activity recognition in Salat to a great extent.

The HAR literature focuses on Salat through adopting different approaches. Some of the research studies in this regard adopt computer vision-based approaches [31, 32], whereas other studies utilize the inertial sensors of smartphones [3, 33], and body-worn sensors [34]. Between the alternatives, computer vision-based approaches demand the prayer to be captured by a camera, whereas sensor-based approaches require a person to carry the smartphone or other sensors in specific positions while praying. Among the existing research studies, some researchers target recognizing the major steps of Salat [3, 33], whereas some simply target to recognize whether the signal pattern obtained from smartphone sensor stands for a prayer pattern or not [35]. Many of these existing studies adopt traditional machine learning approaches [2, 3, 33, 36], whereas few others apply deep learning [31, 32].

Importance of Salat among Muslim Community

The importance of Salat is immense among Muslims, as it is the second pillar of Islam among the five pillars [28, 29]. Regardless of gender, health, and income, it is mandatory upon every adult and sane Muslim to pray five times a day as long as s/he is conscious. Thus, it is the most universal and regular daily worship among Muslims [28, 29].

In Islam, the importance of prayers lies in the fact that the most important thing for a Muslim is his relationship with God, i.e., his faith, God-consciousness, sincerity, and worship of Allah.

This relationship with God is both demonstrated and put into practice, as well as nourished and enhanced, by Salat. These timely meetings with their God refresh their faith and recharge them to follow Islamic guidance with renewed efforts. Therefore, according to the Muslim faith, if the prayers are sound and proper, the rest of the deeds will be sound and proper [37].

In our current world, Salat is also practiced widely among Muslims. For example, a survey [38] of Pew Research Center, which interviewed 38,000 Muslims face-to-face across 39 countries and territories found that, in four of the six regions included in the study – the Middle East and North Africa, Southeast Asia, South Asia, and sub-Saharan Africa – a majority of Muslims in most of the countries pray several times a day. The study also found that three-quarters of the participants or more in 12 countries report performing all five prayers daily. Thus, Salat carries immense importance in the life of Muslims, especially religious Muslims.

Necessity of Activity Recognition in Salat

Considering the importance of Salat among Muslims and the complexity of the activities performed in Salat, earlier HAR studies focusing on Salat emphasized on monitoring the prayer and checking for its completeness and correctness to help the worshippers with their prayers. The primary conception presented in these studies is that people tend to make mistakes in Salat. The study in [39] states that, with the repetition of the basic steps in each prayer, some people might forget some steps or might be confused about whether their prayer was correct or not. Thus, they argue that recognizing the activity sequences of prayer can help detect its correctness. Besides, the studies in [31, 32] aim at evaluating the correctness of the postures of their prayers as people, due to lack of knowledge, might perform Salat with incorrect postures. Overall, the studies in [3, 31, 32] emphasize the idea of having a system that can be used as an educational tool for learning how to offer Salat correctly. Additionally, the study in [2] states that their Salat Activities Recognition Model (SARM) might help Alzheimer's patients or a person with a lack of concentration in prayer so that they can offer their prayers without any confusion. In addition to that, the studies in [33, 35] emphasize designing a system that can recognize the prayer pattern so that mobile phones can go automatically into silent mode while Salat is offered to reduce the chance of distraction due to the ringing of the phone. On top of that, the study in [33] mentions that alerting people about prayer presents another usage of such solutions, as each Salat has to be offered in a fixed time window [33] within the day. Unfortunately, none of these studies reports any statistics about the frequency of mistakes people generally make in Salat, nor they explore the acceptability of such solutions among general people.

Leveraging Smartphone for Activity Recognition in Salat

Sensor-based HAR has got much popularity with the rapid development of mobile and wearable technologies. Over the last decade, HAR research has got significantly inclined towards deploying these devices, especially smartphones as data collection tools due to their portability, unobtrusive sensing, high public adoption, numerous features, and the high processing power that permits analysis in real-time [13]. Consequently, we see almost all sensor-based HAR studies focusing on Salat deploy smartphones for data collection [3,33,36,39]. These studies require a smartphone to be carried or attached to specific positions of the body while praying to collect data from different sensors of the smartphone. Then, the preprocessed raw data are fed into different machine learning classifiers to train them and finally make the prediction about various steps of Salat. These studies differ in the number of sensors they use, placements of the smartphone, number of steps recognized, and classification techniques. Most of the studies use only one sensor and that is accelerometer [3, 33, 39], while some use gyroscope [35], and few studies use gyroscope and magnetometer both along with the accelerometer [36]. Most of these studies report achieving above 90% cross-validation accuracy [2, 3, 33, 36], and thus, establish the proof that prayer activities can be recognized using smartphones.

Potential of Leveraging a Smartwatch in Recognizing Activities in Salat

To date, all earlier sensor-based HAR studies focusing on Salat use smartphones for data collection to the best of our knowledge. Among these studies, in many cases, the placements of smartphones lack convenience and are impractical for daily use [2, 3]. Other studies require the smartphone to be kept in the pocket at specific locations [33, 36, 39]. However, not all types of garments have pockets, and even though they would have pockets, the pockets are not at specific locations or of specific sizes. Besides, carrying a phone while praying is not natural for all.

As Salat is a very frequent worship, the device to be used for its activity recognition, should be practical for daily use. Unfortunately, we do not see such devices being used in the existing related studies. Hence, there is a gap in the existing literature in exploring convenient wearables for activity recognition in Salat that would be applicable for all. On this ground, we assess the potential of using smartwatches for this purpose as they seem to be the only viable alternative to smartphones as a convenient data collection tool. Moreover, as activities in Salat involve movements of hands, where the smartwatch is naturally worn, leveraging a smartwatch in recognizing activities in Salat should have a high potential. Even after having all these promises, the potential of leveraging a smartwatch to recognize activities in Salat is yet to be focused on in the literature to the best of our knowledge.

However, in recent times, smartwatches have already become a promising tool for activity recognition applications due to their specific advantages over other wearable inertial sensors and smartphones [40]. The primary advantage is that they are ubiquitous as people are accustomed to wearing watches and they can be worn for a long time, i.e., at home, at night, etc., [40]. Besides, the wrist placement of smartwatches presents one of the least intrusive placements to wear a device [41] for monitoring activities and data collection. Additionally, the battery life of smartwatches is more durable than smartphones [42], whereas they can combine almost all features of smartphones for data collection and even continuous monitoring [43]. Furthermore, the notifications delivered through the smartwatches to the user are more easily observed than that through smartphones, due to their proximity to the user's line of sight [43]. In addition, smartwatches can easily be worn during any kind of activity in general. Hence, smartwatches stand out among all the marketable wearable devices that can be used for data collection in HAR. Moreover, we find many studies in the literature that successfully use smartwatches for various activity recognition tasks [44, 45]. Therefore, considering the availability, pervasiveness, convenience in use, wide adoption in HAR literature, etc., we find it worth attempting to recognize activities in Salat using smartwatches.

Our Research Contributions

In this study, we propose a new methodology to recognize the activities in Salat using a smartwatch. Before that, we confirm the necessity of recognizing activities in Salat by performing an exploratory study. To do so, we conduct an online survey and ask participants regarding their frequency of making various mistakes in prayer as well as whether they would be willing to explore technological assistance to improve their prayers. Here, the key finding from our survey is that mistakes in Salat are common. Besides, people are interested to explore convenient technological assistance and pray wearing wearable devices if that helps them with their prayer. Thus, we establish the requirement and acceptability of such solutions in the real world.

Afterwards, in this study, we use Samsung Galaxy Watch Active 2 [46] as the data collection tool, and develop a wearable app for this purpose. We prepare our own dataset comprising prayer data from 30 subjects and 3,50,762 data samples. Then, we preprocess the collected raw data and separate the steady and transitional states using two alternative approaches. Subsequently, we detect some of the steady states using semantic rules derived from domain knowledge. Next, we use Dynamic Time Warping (DTW) algorithm for predicting the transitional states. Finally, we apply some postprocessing on the predictions to detect potential misclassifications and correct them using domain knowledge to further enhance the accuracy. In this process, we achieve up to 99.3% overall accuracy. We compare our methodology with baseline machine learning-based methodology and show that our proposed methodology outperforms the baseline methodology as well as that of the earlier studies. To summarize, the main contributions of this study are as follows.

- First of all, we perform an exploratory study by conducting an online survey to understand the opinion of people regarding the idea of helping them to improve their prayer through technological assistance. Our key findings from the survey responses provide potential directions for shaping this technology for real-life adoption.

- We propose a methodology for recognizing activities in Salat with a smartwatch which is, to the best of our knowledge, the first work that makes use of a wearable that is a convenient and non-distracting option for regular use during the prayer.

- We prepare a dataset by collecting data using a smartwatch. The dataset consists of data collected from 30 subjects and has 3,50,762 data samples pertinent to activities in Salat. The dataset can help future researchers to explore new techniques in this regard.

- We divide a prayer unit into more granular steps and achieve near-perfect overall accuracy (around 99.3%) by using semantic rules derived from domain knowledge and Dynamic Time Warping (DTW) algorithm followed by postprocessing steps through integrating domain knowledge.

- We further devise strategies to make our system robust to overlook the extra activities that do not nullify the prayer. To the best of our knowledge, we are the first to deal with the extra activities in recognizing activities in Salat.

In the process of making these contributions, we focus on the following set of research questions.

RQ1 How prevalent are mistakes in Salat among people? What types of mistakes are more common among them? Can we find the mistakes and their frequencies through a mixed-method analysis?

RQ2 Do people need technological assistance for improving their prayer? How willing are they to accept such technological assistance? Can we find the need for technological assistance and its acceptability through a mixed-method analysis?

RQ3 If people welcome technological interventions or assistance for their Salat, then can we help them in improving their prayers by leveraging a more convenient device, with improved accuracy compared to that of the solutions existing in the literature?

We will elaborate on each of these research questions later in this study in Section 5.1.

Fitting Our Study in the Literature

In the process of answering the first two research questions, i.e., RQ1 and RQ2, we conduct a mixed-methofd analysis over data collected by a self-reporting based online survey. Our survey results contribute to a self-reported dataset. Besides, our analysis comprises both quantitative and qualitative explorations. Such outcomes contribute to the research in the domains of Human Computer Interaction (HCI) and Computer Supported Cooperative Work (CSCW). Besides, in the process of answering the third research question, i.e., RQ3, we develop a new methodology for recognizing activities in Salat leveraging a smartwatch. Here, we prepare a dataset using a smartwatch. To recognize activities in Salat from the dataset, we divide a prayer unit into granular steps and use semantic rules as well as DTW algorithm followed by custom postprocessing. Such outcomes contribute to the research in the domain of HCI and ubiquitous computing. Thus, this study emplaces itself in a broad cross-section of HCI, CSCW, and ubiquitous computing.

Chapter 2

Backgrounds and Preliminaries

In this chapter, we describe some theoretical concepts used in our study. We first start with activity classification and then introduce the concept of HAR followed by a discussion on Salat which is the activity that is our focus in this study.

Human Activities and Their Classifications

Many of the previous research studies [6, 7, 27, 47] have classified human activities into two categories, namely simple and complex activities. Simple activities consist of a single repeated action. Some studies defined the notion of action or atomic activity as a unit-level action that cannot be broken down further [6, 7]. Examples of action include such as a step in walking, raising a hand, etc. On the other hand, examples of simple activities are walking, running, sitting, etc., where actions are repeated back to back over time. Besides, complex activities compile a series of multiple actions [6, 11]. Complex activities frequently entail numerous concurrent, interleaved, or overlapping behaviors. Examples include playing a game, cooking, cleaning, buying, etc. Such activities are sometimes termed as composite activities [48]. They last for a long time and have high-level interpretations. They demonstrate realistic representations of people's daily lives [6].

Human Activity Recognition

Human activity can be defined as any bodily movement produced by skeletal muscles resulting in energy expenditure above resting level [49]. Human Activity Recognition (HAR) implies mapping a set of collected time series sensor data to a specific human activity [50]. The advent of ubiquitous and pervasive computing has resulted in the development of several sensing technologies for capturing sensor data relevant to information related to human activities. Besides, the advances in artificial intelligence have revolutionized the ability to extract deeply hidden

Figure 2.1: Basic steps of Salat [1]

information within the sensor data for accurate human activity detection and interpretation. As a result, in the last two decades, the literature has got enriched with numerous HAR studies focusing on different activities done in real life from different dimensions.

HAR can be divided into two major categories - computer vision-based and sensor-based. Computer vision-based studies require an activity to be captured by a camera and later use image processing on the captured images. On the other hand, sensor-based HAR aims at recognizing and analyzing the activity of an individual based on one or multiple sensor data. During the past decade, sensor technologies have got much developed in terms of power, size, accuracy, costs, etc. These developments enable a wide range of sensors to be integrated into smartphones and other portable devices to make them smarter and more useful, which in turn contributed to expediting HAR research. Even then, in today's world, HAR is extensively and successfully used in numerous domains such as surveillance systems [20], sports coaching [21], gesture recognition [51], gait analysis [52], smart home [53], patient monitoring systems [54] and a variety of healthcare systems [21], etc.

Activity under Study: Salat

Salat is the second pillar of the Muslim faith. It is the most essential and fundamental religious ritual and the most regular compulsory worship of a Muslim's life. A Muslim has to pray five times a day: Fajr (dawn prayer), Dhuhr (afternoon prayer), Asr (late afternoon prayer), Maghrib (evening prayer), and Isha (night prayer) [28, 29]. Salat consists of repeating units, called Rakah (plural Rakat). In each prayer, every person must fulfill a mandatory count of Rakats. For example, 'Fajr' or the dawn prayer, which is the first prayer of the day, consists of two units, "Dhuhr" and "Asr" consist of four units, and so on. A Rakah consists of a series of postures that must be executed in a predefined sequence such as standing, bowing, prostrating, etc. Figure 2.1 depicts a complete prayer cycle.

As multiple postures or simple activities need to be performed one after another in Salat, it falls under the category of complex activity by definition [6, 11]. To elaborate a bit more, Salat starts with Takbeer which means raising hands up to the ear or shoulder and then lowering immediately.

After Takbeer, a person remains standing while placing his hands either on the chest or belly and has to recite the first chapter from the Qur'an. In some prayers, this is followed by the recitation of (part of) some other chapter. Generally, a person remains longer in standing and sitting positions than in other steps. Once the recitation is done, a person bends down keeping his hands on the knees as shown in Figure 2.1 which is called Ruku or bowing. In this state, the person recites some supplication, which is short in general. Therefore, this step, in general, takes less time than Standing. After the recitation in the bowing stage is done, a person goes up and stands straight with hands on both sides and recites some supplications. We refer to this as short standing, as the posture of this stage is different from standing in terms of placement of hands and duration. Next, a person goes to the Sujud or prostration phase and touches his head to the ground, and recites some supplications. In each Rakah, a person has to prostrate twice. After the first prostration is done, a person rises up from prostration and sits down for a short time, and then goes again for the second prostration. Besides, at every even Rakah a person has to sit again after the second prostration to recite a specific supplication called Tashahhud. If the Rakah is the last one of the prayer, then a person has to recite some more supplications. Consequently, this sitting also takes a bit longer. However, once done, a person ends the prayer with Taslim, which means turning his head to the right and then to left reciting specific a supplication.

Variations in Salat

Based on priority, time, school of thought, capability, etc., the number of units or sequence of steps in Salat varies greatly. First of all, based on priority, there exist different types of prayer [28, 29, 55]. For example, one category is obligatory or Fard, which is compulsory for Muslims to offer within a prescribed time frame. Next, some prayers belong to the Sunnah category, which is rooted in the practices of the Prophet of the Muslims and needs to be offered if there is no strong excuse. Then, there are Nafil prayers and this category is totally optional. There exist some other such categories too.

Secondly, the number of Rakah to offer varies based on the time of the prayer [28, 29, 55]. For example, in the Fajr prayer, a person has to pray two Rakat of Sunnah prayer followed by two Rakat of Fard prayer, whereas, in the Maghrib prayer, three Rakat of Fard and two Rakat of Sunnah prayers are prescribed. The steps in Salat also differ based on the number of units and priority. For example, the way to offer four units of Sunnah prayer differs from that of the Fard prayer. In the Sunnah prayer, a person has to recite (part of) a new chapter after the first chapter of the Holy Qur'an in each Rakah. However, while standing in the Fard prayer, he needs to do this only for the first two Rakat. In the last two Rakat, he should only recite the first chapter and nothing else. Additionally, another prayer, called Witr, which needs to be prayed after the Ishaa prayer and before Fajr, differs from other prayers. This happens as, in the last Rakah, one has to give Takbeer after the recitation is done and recite some other supplications.

Furthermore, there are several schools of thought in Islam regarding Islamic canonical laws, and prayer postures vary from one school to another. For example, in most schools of thought, after the Takbeer, one has to place his hands either on their chest or on the belly. However, in one school of thought (Maliki), it is said to keep hands on both sides floating low while in the standing position. Besides, male and female prayer vary in some schools of thought, whereas others claim prayers of both genders to be the same. For example, in one school (Hanafi), a male places his hands on his belly and a female on her chest. The postures for bowing and prostration also vary greatly in this school based on gender [28, 56, 57].

Even if two persons belong to the same school of thought, prayer postures vary from person to person. For example, while raising hands for Takbeer, some raise their hands up until their ear, some raise up until their neck, some keep their hands close to their face, some keep them a bit far, and so on. Differences also happen due to the lack of knowledge of the standard ways of performing Salat. Similarly, the sitting position and the prostrating position vary a lot. Postures vary due to differences in capability too. If a person is unable to make some postures completely due to illness or so, then he does as much as possible and need not go beyond his capacity. For example, if someone cannot offer prayer in standing due to any disability, he is allowed to offer the whole prayer through sitting.

Potential Mistakes in Salat

A Muslim, while offering Salat, has to repeat the specific postures mentioned above, systematically and recite from the Holy Qur'an or other specific supplications at specific points. Along with these, he has to keep track of the counts of Rakat, prostrations, etc. Hence, due to the variations in different types of Salat or difference within the units, people tend to make mistakes in Salat [28, 58]. A lack of knowledge is also responsible for making mistakes. Moreover, in case praying with proper concentration becomes difficult, and mistakes happen more frequently. It is more common, especially for beginners and elders. For example, forgetting to recite another chapter from the Qur'an after the first one or forgetting specific supplications for a specific step is common. Besides, out of forgetfulness, a person may either forget to perform any obligatory step such as forgetting to bow or missing one prostration, etc., or he may add anything extra to Salat such as performing five Rakat instead of four. Apart from that, a person may also become confused about the counts. The most common mistake in Salat is to forget the count of Rakat or the count of prostration.

Chapter 3

Related Work

We situate our research work in a body of related studies exploring the recognition of various human activities. In the literature, numerous attempts have been made to develop different HAR approaches using different types of sensors. Such studies vary from the perspectives of sensors, their number, locations, activities to be tracked, pre-processing techniques, feature extraction and selection methods, classification approaches, and even the research purpose itself. In the following, we go through such relevant topics and shed some light on the literature.

Research on Recognizing Human Activities

Human activities recognized in the literature, exhibit different complexity, nature, and domains. Depending on the complexity, various techniques and types of signals are explored as follows.

Simple Activities

Simple activities refer to a repetitive occurrence of atomic activities or actions. Examples are - walking, running, sitting, etc. In the literature, simple activities are mostly attempted to be recognized using different techniques and sensors [27]. Many research studies focused on Activities in Daily Life (ADL) such as walking, running, jogging, etc., [59–61]. Other studies focused on transitional activities too [62, 63]. The research study in [64], explored walking (forward, left, right, upstairs, and downstairs), running forward, jumping, sitting, standing, sleeping, elevator up, and elevator down using wearable sensors and built a benchmark dataset with these activities. The study in [24] attempted exercise activities such as bench dips, squat upright row, triceps extension, and many others. Besides, kitchen activities such as chopping, peeling, slicing, dicing, coring, spreading, etc., were recognized by the study in [26]. This study collected data from 20 subjects using four specially-designed kitchen utensils incorporating embedded 3-axis accelerometers. The study testified that a broad set of food preparation actions can be reliably recognized using sensors embedded in kitchen utensils. On the other hand, a research study [65] recognized ten different dance micro steps. Another study [25] explored construction activities such as hammering, sawing, filling, drilling, sanding, and grinding using the accelerometer data. Moreover, recently fall detection has also gained much interest from researchers because of its vast application in healthcare [14, 15]. These studies present a few examples of the wide variety of simple activities covered by the HAR researchers of different application domains.

Complex Activities

Techniques for recognizing atomic activities such as gestures or actions are mostly mature for now, however, complex activity recognition still remains a challenging issue [6]. In this regard, the study in [27], attempted to recognize simple activities as well as some complex activities such as cooking, cleaning, etc., through a smartphone. Using supervised machine learning algorithms, this study found that while simple activities can be easily recognized with an accuracy of around 93%, the performance of the prediction models in recognizing complex activities appears to be poor (50%). Besides, the study, in [7], developed and validated a Context Driven Activity Theory (CDAT) to recognize complex activities. This study used probabilistic analysis and Markov chains to reveal signatures of complex activities, assign weights to the signatures and update the definition of complex activities. Their algorithm achieves an accuracy of 95.73% while recognizing complex activities that are concurrent and interleaved in nature.

Another research study [66] presented a machine learning approach to correctly classify highly-correlated and imbalanced nursing activities. The model presented in this study appeared to be very simple and can be trained with low computational costs achieving around 72% accuracy. Besides, the study in [11] proposed an algorithm capable of mining temporal patterns from low-level actions to represent high-level human activities. Another study [67], proposed a model to recognize and classify complex at-home activities through wearable sensing leveraging selective multi-modal sensor suites from wearable devices. It further enhances the richness of sensed information for activity classification by carefully leveraging the placement of the wearable devices across multiple positions on the human body. Additionally, the study in [6], built a dictionary of time series patterns, called shapelets, to represent atomic activities. The study then presented three shapelet-based models to recognize sequential, concurrent, and generic complex activities. However, to the best of our knowledge, complex activities are still less explored compared to simple activities [6, 11].

Data Collection in HAR

HAR allows machines to analyze and predict human activities from different input data sources, such as sensors, multimedia content, etc. There exist different methodologies of data collection in HAR. The major and most widespread categories are computer vision-based and sensor-based data collection. We briefly present some studies for each of these methods below.

Computer Vision-based Data Collection

Computer vision-based data collection require data capturing through one or more camera and activities are recognized by processing captured images or recorded video sequences. The recognition of human activities from static images or video sequences exhibits applications in many fields. For example, computer vision-based HAR is utilized in monitoring applications in industries [68], fraud detection [69], extraction of information from videos [68], video assistance and surveillance [68], and public security [70] where crowds' movements are tracked for detecting anomalous, violent, or criminal activities. There also exist applications in surgical operations, patient monitoring, and interpretation of language signs [71]. Besides, it is possible to classify static signs of the sign language using computer vision-based HAR [72]. Additionally, authors in computer vision-based HAR systems can help a teacher to control a multi-screen and multi-touch teaching tool such as swiping right or left to access the previous or next slide, calling the eraser tool to rub out the wrong content, etc., [73]. However, as image processing is unavoidable in computer vision-based HAR, it mostly relies on DL, and thus, generally appear to be computationally expensive. Besides, privacy and security issues get entailed with this approach.

Sensor-based Data Collection

Sensor-based data collection in HAR generally covers two prominent data collection devices. They are smartphones and smartwatches. These devices differ in their data collection approaches as well as applications.

Smartphone-based Data Collection

Smartphones have been extensively studied for recognizing different physical activities in recent years [13]. Smartphones are being used because of their wide availability due to mass adoption among people. Besides, another advantageous aspect of smartphones is the fact that they are equipped with different sensors such as accelerometers, gyroscopes, magnetometers, etc., that can be used in different types of activity recognition.

In the initial phase of developing smartphone-based approaches, the developed approaches mostly worked offline. For example, the study in [74], presented Centinela, a system consisting of a chest unit composed of several sensors to measure acceleration data and vital signs (e.g., heart rate, breath amplitude, etc.). Here, a smartphone is wirelessly connected via Bluetooth for accumulating the collected data and processing the accumulated data. Besides, the study in [59] worked on recognizing common human activities such as walking, jogging, ascending stairs, descending stairs, sitting, and standing. They used the inertial accelerometer sensor of Android-based smartphones. In [75], a system for recognizing five transportation activities was developed by combining labeled and unlabeled data collected by inertial sensors of smartphones. The study in [76] attempted to classify basic activities in real-time while addressing the issues regarding transitions and unknown activities. This system allows the incorporation of new elements (e.g., inertial sensors) and provides an easily exportable output to other ambient intelligent systems requiring activity information. Several other studies performed gait analysis [17, 18] using smartphone sensors for elderly healthcare. In this regard, the study in [77], focused on energy efficiency and introduced an activity-sensitive strategy for continuous activity recognition. The proposed system achieved energy saving by tuning the accelerometer sampling frequency and by extracting classification features separately. While, most of the studies in this regard, focus on simple activities such as locomotion, the study in [27] covered more complex activities such as cleaning, cooking, medicating, sweeping, washing hands, and watering plants. Nevertheless, the study in [78], focused on online recognition of activities using smartphones together with performing the training phase within the smartphones.

Smartwatch-based Data Collection

With the recent emergence of smartwatches, HAR research arguably achieved a new dimension. Smartwatches, unlike most other technologies, offer new levels of ubiquitous computing by being physically attached to the users and by offering a unique set of features [79]. Samrtwatch-based data collection appears to be a promising alternative due to the fact that smartphone-placement issues noticeably affect performances of recognition [80]. More importantly, smartphone-based data collection [81, 82] often becomes difficult in long-term activity monitoring. On the other hand, studies conducted with both smartwatches and smartphones found that smartwatches can achieve superior performance for a wider range of activities than smartphones [44]. Besides, as explored in a smartwatch-based study [45], activity recognition using Restricted Boltzmann Machines (RBM), can cover a variety of typical behavior and tasks demanding no additional resource other than smartwatch-class hardware. Moreover, the study in [83] explored obtaining keyboard usage information of a laptop using the accelerometer and gyroscope sensors of the Samsung Galaxy Live smartwatch. In addition to that, the study in [84], classified eight different daily human activities with a Moto 360 smartwatch, using PCA and Random Forest.

Data Analysis in HAR

After collecting data, the next task generally performed in HAR, is data analysis. Data analysis can be done using classical machine learning, Deep learning, template matching techniques such as Dynamic Time Warping (DTW), etc. We present research studies in this regard in the following.

HAR using Classical Machine Learning

Currently, the most exploited and probably the most mature approach for data analysis in HAR is using classical machine learning methods such as Decision Tree, Random Forest, SVM, KNN, Naive Bayes, etc. The applicability of the methods varies based on the application under exploration.

Perhaps the most widely-used classifiers for data analysis in HAR are different variants of Decision Tree. A Decision Tree is a classifier with a tree structure in which one feature is evaluated at each traversed node and each leaf of the tree corresponds to one class label. In this regard, J48 algorithm presents an open source java implementation of the C4.5 Decision Tree algorithm [85] and is often used for HAR using motion sensor data [86, 87]. Other studies reported that decision-based classifiers outperform Na¨ıve-Bayes [88] in their setup. Besides, Random Forest is another widely used classifier in HAR [89, 90].

Support Vector Machine (SVM) is also a popular classifier in HAR [87, 91, 92]. In this case, hyperplanes are used to create decision boundaries for the data in a high-dimensional space. The study in [93], used a multiclass support vector machine (SVM) to create nonlinear classifiers by adopting the kernel trick for training and testing purposes. Similarly, using an implementation of SVMs, another study recognized six indoor activities [92].

Instance-based Learning (IBL) algorithms, though computationally expensive, are also used in HAR because of their ability to adapt to new data. KNN is the mostly used IBL classifier that is used in a variety of activity recognition approaches [86, 94]. Other studies proposed a real-time classification system using KNN [94, 95] to classify basic human movements.

Many studies investigated the potential of ensemble classifiers, which combine the outputs of several classifiers of the same type in order to get better results. For example, bagging and boosting ensemble meta-algorithms were used in [74]. Such studies are generally computationally expensive, as more base-level algorithms need to be trained and evaluated.

Several studies in the literature compared different classification techniques. For example, the study in [87], compared different classification techniques using a combination of time and frequency domain features. The study also attempted to fuse different combinations of low-power classifiers. Another study [96] compared different techniques and created a code base for an Android-based operating system. Similarly, the study in [59] compared different classifiers using a smartphone-based platform having an emphasis on the diversity of a 43-element feature set.

HAR using Deep Learning

Recent studies on human activity recognition are now inclining towards using deep learning models due to their capability of simulating high-level features in the supplied data. This happens as the models include multiple neural networks that imitate the function of human neurons. The neural networks can represent features from low-level to high-level. Convolutional Neural Networks (CNN) [97] and Recurrent Neural Networks (RNN) [98] are the two most popular deep learning models in this regard. Between these two alternatives, learning spatial representation from sensor data is a strength of CNN. Therefore, CNNs are suitable for discovering relationships in the spatial dimension. Different research studies [99, 100] applied CNN to the field of activity recognition. However, CNN lacks the ability to capture temporal relationships in the time-series sensor data. To overcome this limitation, RNNs are designed to model the time-series data and are suitable for discovering relationships in the temporal domain. As the performance of activity recognition is expected to improve with longer context information and longer temporal intervals, several research studies have applied RNNs for the purpose of complex activity recognition [101, 102].

The major weakness of RNNs lies in the vanishing or exploding gradient problems, which is addressed in Long Short-Term Memory (LSTM)-based RNNs [103]. The LSTMs, with their capability of memorizing and modeling the long-term dependency in the supplied data, have taken a dominant role in the HAR domain [104, 105]. Nevertheless, in recent times, hybrid deep learning models combining both CNNs and RNNs are also explored for activity recognition tasks [106, 107].

HAR using Dynamic Time Warping (DTW)

Dynamic Time Warping (DTW), though extensively applied in speech recognition, has also been proven effective in HAR research. While most of the recent research studies focus on extracting complex features to achieve high accuracy in activity recognition, several research studies use a template selection approach based on DTW to avoid the complex feature engineering step. For example, the study in [108], modified DTW to improve computational efficiency and similarity measure accuracy. They use it for motion data clustering and activity template construction and classification. Here, each template is the time series average representation of a cluster. Another study [109] proposes a new ensemble classifier based on DTW and uses combined information from multiple time-series sensors to map them with corresponding activities. The training data used in this study consisted of a set of short-time samples used as templates for DTW and the time series for each sensor is classified by assessing similarity to these templates.

Besides, in this study, results from separate classifiers are combined using a voting ensemble for recognizing six activities indicating the DTW classification as a promising one. The study in [110], classifies light sport exercise activities such as walking, push up, sit up, and squat jump using the accelerometer sensor on a smartphone and smartwatch that is placed on the left hand of the user. The study uses the k-Nearest Neighbor algorithm and DTW as the main classification algorithms and analyzes the similarity of time series data. Another research applied DTW to process different shapes of foot movements, which was captured using wearable sensors [111]. Here the proposed method detects early signs of Alzheimer's disease. Besides, the study in [112], performs HAR for six different human activities through exploring time-phased data and the signal magnitude of an on-body creeping wave.

HAR Research Related to Salat

In the literature, we find a handful of HAR research studies that focus on recognizing the activities in Salat. Though the first HAR work targeting Salat came to light over a decade ago, only a few studies were conducted on this topic in the meantime. These studies include both computer vision-based and sensor-based approaches. In this section, we will shed some light on these studies presenting their contributions, strengths, and weaknesses.

The first study targeting automatic recognition of prayer movements was conducted in 2009 [113]. They adopted a computer vision-based approach to solve this problem and used a camera to capture the side view of the prayer. Later, the study in [30], investigated motion tracking for Salat activity recognition leveraging two Kinect devices. They place the devices at pre-defined positions and angles to obtain multiple views in a single space. The system gets the skeleton information from Kinect Software Development Kit. They selected some joint movements that had significant changes during Salat activities and classified the activities using Hidden Markov Model.

The first study using deep learning for recognizing the basic gestures of Salat was conducted in [31]. They built an image dataset for the basic Salat positions and trained a YOLOv3 neural network for recognizing the gestures. Besides, the most recent computer vision-based study on Salat [32], proposed an assistive intelligent framework that guides worshippers to evaluate the correctness of their prayer's postures. The methods proposed in this study for image comparison and pattern matching, utilized several algorithms such as Euclidean Distance, Template Matching, and Grey-Level Correlation in combination to compare the images of the user and those stored in the database. The study reported some wrong predictions resulting from insufficient lighting. Additionally, in this study, the camera must be deployed at a fixed position keeping that static and the camera angle should also be fixed to a certain value for correct classification.

The first sensor-based approach for Salat activity recognition was proposed in 2016 [3]. The study used a smartphone accelerometer for prayer activity monitoring. The study utilized three machine learning classifiers: J48 Decision Trees, IB1 (Instance-Based Learning) Algorithm, and Naive Bayes. However, the suggested placement of the smartphone in this study (at the upper back of the user) is not practical or convenient. Besides, manual cleaning of data and ignoring the gravity factor as well as temporal information are some other limitations of this study.

Another study [39] presented a framework for activity recognition using a smartphone accelerometer sensor to recognize simple daily-life activities such as standing, sitting, bowing, prostrating, etc., in order to detect the correctness of a more complex activity that could be Salat. The necessary placement of the smartphone in this study is in the shirt's pocket. On the other hand, the study in [35] developed a pattern for the whole prayer and used DTW for matching the test pattern against the developed pattern to decide whether the whole activity is a prayer pattern or not. The study tested prayer patterns against walking patterns and found a significant difference between them. The study claimed its proposed system to be independent of smartphone placement, which makes the proposed scheme more robust. Another study [34] distinguished between congregational prayer and individual prayer as well as between silent prayer and loud prayer using two body-worn sensors. However, neither [35] nor [34] attempt to recognize different steps and activities within the prayer.

A similar study was conducted in [33] using mobile accelerometer data and performance comparison was done among three feature extraction approaches and eight machine learning classifiers. The study found that Random Forest is the most appropriate classifier for their intended task. The study also explored a two-level classifier to solve the confusion between two similar prayer stages. Besides, the study investigated the effect of personal characteristics such as height and age on the performance of the classifier. However, the main limitation of the study is that it demands placement of the smartphone over the hip area which might be inconvenient for many users.

Another research study [2], utilized a smartphone's accelrometer to help Alzheimer's patients in their prayers. The study used a bunch of machine learning classifiers to classify three steps of prayer such as standing, standing to bowing, and standing to prostration. The study investigated placements of smartphones at four body positions, i.e, left-hand, right-hand, trouser right pocket, and trouser left pocket, and finally suggested the first two positions to be better ones. However, these positions for smartphone placements are not realistic. Another study [36] recognized six steps in prayer using three smartphone sensors such as accelerometer, gyroscope, and magnetometer using SVM. However, this study was conducted by placing a smartphone in the pant's pocket, which is not always common, especially for female users.

Chapter 4

Motivations behind Our Study

From the existing literature, it is evident that the practical usage of the sensing device is mostly overlooked in the case of HAR in Salat. This happens as most of the existing studies either demand video capturing or require placement(s) of the sensing device(s) at an unconventional or inconvenient place(s). Besides, the existing studies lack the exploration of recognition at the fine granularity for Rakat counting and steps in Salat. Additionally, personalized testing and enhancing recognition accuracy present two other aspects that are still worth investigating for HAR in Salat. Nonetheless, developing a dataset with a smartwatch for HAR in Salat is yet to be focused on in the literature. All these vacancies in the literature motivate us to perform this study. We elaborate more on these vacancies in the literature in the next subsections.

Natural Usage of the Sensing Device used for HAR in Salat

Existing studies on activity recognition in Salat are yet to provide a convenient solution for practical use. Both the computer vision-based approaches and sensor-based approaches using smartphones, proposed for recognizing the steps in Salat, have their own challenges limiting their applicability in real-life usage.

First of all, the computer vision-based approaches are not suitable in practical cases, as they demand good video recording facilities and sufficient lighting [31], which might not be always available. This might also be a cause of distraction for people, as they have to pray under the coverage of a camera. In addition to that, accuracy might also get affected greatly due to any occlusion and background change owing to other nearby worshipers. Moreover, as they require image processing, they are computationally expensive. Last but not the least, the video capturing raises a serious privacy concern especially for Muslim women, as Muslim women who observe modesty will not be comfortable using such a solution.

Figure 4.1: Suggested smartphone placement positions in [2] and [3]

On the other hand, the sensor-based approaches demand placement of a smartphone with the body to capture intended signals. The placements proposed by several of the studies are not convenient [3] at all. Figure 4.1 portrays some of the inconvenient placements required by some earlier studies. Other studies [35, 36, 39] need a person to place the phone in his/her pocket. However, we should consider the fact that not all types of garments have pockets. Even if the garments have pockets, it is not confirmed that the pockets are placed in the same place or even of the same size. Therefore, the solutions based on the placement of smartphones in a pocket are not applicable for all. Nevertheless, it is not natural for all to carry smartphones with them while praying.

On the other hand, smartwatches, due to their specific advantages over other wearables, have recently become a promising tool for activity recognition applications. They are more ubiquitous as people are accustomed to wearing watches and they can be worn anytime and anywhere, even at night, during exercise, etc., [40]. Besides, the wrist placement of smartwatches presents the least intrusive placement for wearing a device [41] while monitoring activities and data collection. Additionally, the notifications delivered to the users through smartwatches are more easily observed than that delivered through smartphones, owing to the proximity of the smartwatches to the user's line of sight. Furthermore, smartwatches can easily be worn during any kind of activity. Though the smartwatches combine most of the features of smartphones, the battery life of smartwatches is generally more durable [42], which makes the smartwatches more suitable for continuous monitoring [43]. Hence, smartwatches stand out among the marketable wearable devices that can be used for data collection in HAR. Moreover, several research studies in the literature successfully use smartwatches for various activity recognition tasks [44, 45]. In the context of Salat also, smartwatches can be more convenient and practical to wear while praying than carrying smartphones. More importantly, smartwatches pose no (or little) privacy concern. Therefore, considering the availability, pervasiveness, convenience in use, wide adoption in HAR literature, etc., we find it worth investigating to recognize activities in Salat using a smartwatch.

Recognition at the Fine Granularity for Rakat Counting

Another important thing is that, in the existing research studies, as the accuracy is not 100% (or very close to it), they might provide some wrong predictions. However, in the case of prayer activity recognition, our target is to provide the users with the perfect (or near-perfect) count of Rakat, bowing, prostration, etc. To achieve this goal, appropriate postprocessing in addition to applying machine learning classifiers is of utmost significance. Postprocessing is necessary for detecting and correcting the intermittent wrong predictions made by the classifiers. Unfortunately, we are yet to find any research study that goes up to this depth of detecting and correcting the wrong predictions even though these tasks are unavoidable in predicting the activities in Salat correctly.

Recognition of Steps at More Granular Level

Many of the existing studies recognize some major steps in Salat, however, not all of the steps. For example, the study in [2] considers three steps i.e., standing, standing to bowing, and standing to prostration. However, a person goes to the second prostration from the sitting position, which is not investigated in this study. Another issue with this approach is when its classifier model receives signals of a complete prayer, the non-recognizable steps need to be removed, and only then the classification can be performed. Similarly, the study in [3], segments the signals based on the intersection of the axes, which generally works for the steps performed in one unit. However, when a person goes up from sitting to standing, then there appears a back-to-back intersection of the axes, which is not considered in this approach. Thus there is a high chance of getting erroneous segmentation and recognition. Apart from that, none of the previous studies make any attempt to recognize Takbeer, which is an important task, as a person needs to start his prayer with Takbeer as well as the need to perform it inside the Witr prayer while standing after completing his recitation. Besides, in some schools of thought, a person needs to perform Takbeer multiple times inside one unit prayer [28, 29]. Thus, without recognizing Takbeer, the prayer of the people from some schools of thought and also the Witr prayer cannot be recognized properly. Furthermore, as all of the earlier studies use smartphones for data collection, they cannot differentiate between the postures of standing and short-standing, as the smartphone is either kept in the pocket [36, 39] or tied to the body [2, 3]. However, recognizing these steps of Salat is crucial for assessing the completeness and correctness of the prayer.

Personalized Testing through Leave-One-Subject-Out

A very important concern regarding the existing sensor-based studies for Salat is that most of the studies (except [3]) use cross-validation to measure the performance of their models [2, 33, 36].

However, in the case of HAR, investigating only cross-validation accuracy is not a very good estimate as already reported in [114]. The study in [114], demonstrates that k-fold cross-validation artificially increases the performance of the recognizers by about 10%, even by 16% when overlapping windows are used. This happens as the samples produced by the same subject are likely to be correlated due to many factors and k-fold cross-validation may thus, overestimate the performance of an activity recognizer. The problem appears to be more prominent when overlapping sliding windows are used. Therefore, the authors in [114], concluded that Human Activity Recognition systems should be evaluated by Subject Cross Validation or Leave-one-subject-out validation.

This problem gets evident in the study presented in [3]. Here, the cross-validation accuracy is 100% using Naive Bayes, however, with the test data of subjects who are absent in the training data, the maximum accuracy drops to 91.8% using Naive Bayes. This further establishes the fact that the models under exploration need to be investigated under personalized testing. Thus, as voiced in [114], the models need to be tested under the leave-one-subject-out (LOSO) scenario. As this is mostly overlooked in the literature, the accuracies reported in the existing studies, will not mostly reflect the realistic outcomes, and the accuracies in real usages are expected to be substantially lower than the reported accuracies.

Accuracy of Recognition

Assistive technologies developed to help people in their religious worship should be designed with utmost care and should be made as much reliable as possible. Otherwise, there will be a high chance that people will not use them if, instead of helping them, the solutions lead to mistakes even in a small fraction of the time.

To exemplify, it will not harm much if we miss one or two steps in the case of a pedometer in regular use, whereas, it will destroy the credibility of the HAR solution if it reports the Rakat count to be 3 where the actual count is 4. Therefore, although the accuracies of HAR models focusing on Salat are mostly around 90% [3, 33, 36], therefore, there still remains room for further improvement in the accuracy of recognition. In this regard, only [2], reports obtaining 97% average accuracy with Random forest. However, this is obtained by placing a smartphone in the arm position, which is not practical and convenient at all. On top of that, we already have mentioned, the existing studies mostly use cross-validation accuracy, therefore, the mentioned accuracies do not truly reflect their models' realistic performance, and the models are likely to perform worse in real-life cases [114]. Besides, the accuracy of HAR solution for Salat might hamper if a person performs any extra activity in Salat, which does not nullify prayer [115]. Therefore, the model should be robust enough to ignore these activities, i.e., perform uniformly in the presence of these activities.

Development and Exploration of Large Dataset

The size of the dataset, with which the earlier studies conducted their activity recognition for Salat, is often very small. The majority of these studies cover only ten subjects [2, 3, 39] or even less [36]. Only one existing study [33] collects data from 20 subjects. Among these studies, the study in [3], works with training data having only 118 samples. As activity patterns inevitably vary from person to person [116], the training data might not be capable to capture the variations among the activities performed by different people accurately when the dataset is small. Accordingly, it is natural that such models might not perform well in real cases.

Applications of HAR in Salat

A HAR solution that is capable of recognizing activities in Salat correctly, can provide a worshipper with detailed information about his prayer, such as the activities performed by him, the sequence in which those activities were performed, the count of Rakat as well as the count of other activities, etc. These can potentially help a worshipper to determine errors in Salat (for example, error in Rakat count), whether he missed out any activity, or added anything extra, etc. Such solutions are especially helpful for beginners and elderly people or people with diseases such as Alzheimer's, etc. Moreover, in the month of Ramadan, Muslims are prescribed to pray extra prayer after the Isha prayer, which covers either 8, 20, or more Rakat [28, 29] to be prayed. A person, while praying alone, may find it hard to keep track of the count of Rakat, i.e., how many Rakat he has offered and how many are left. In such a scenario, a HAR solution capable of providing an accurate Rakat count might come in handy. Additionally, such a system could also provide the duration of each step of Salat. This is important as Islam emphasizes praying in a slow and steady manner spending enough time on each state, and not rushing through prayer [28, 29]. Therefore, if a worshipper can get the information about the time he spent in each of the steps in Salat, he can determine which step(s) he spends less time on. However, through all this information, he can understand where he needs to pay more attention to improve his overall prayer. Thus, improving the individual steps of Salat such as standing, bowing, etc., through spending more time in the steps with more appealing recitations present another usage of HAR in Salat.

Besides, determination of the error in Salat is particularly important in the case of performing individual Salat in crowds (for example, in Makkah or Madinah during Hajj/Umrah as shown in Figure 4.2. However, if mistakes or confusions lead a worshipper to repeat his Salat, he stays some more time in the crowd. In this regard, a HAR solution in Salat can assist in determining the mistake and removing the confusion.

(a) Masjid-al-Haram, Mecca [117] (b) Masjid-an-Nabawi, Medina [118]

Figure 4.2: Muslims praying at different places during Hajj and Umrah times

Figure 4.3: Examples of extra activities performed during Salat [4]

Taking Care of Irregular Activities Permitted in Salat

While performing Salat, people might perform some extra activities such as fixing garments, rubbing eyes, etc., as shown in Figure 4.3. These activities, though not regular, do not invalidate the prayer [115]. The movements induced by these activities produce additional signals and these additional signals will be captured by the sensing device engendering noises in the sensed signals. A system designed for activity recognition in Salat needs some special handling to mark such noises induced by these irregular activities and carefully ignore them while recognizing ritual activities in Salat. However, to the best of our knowledge, this issue is yet to be addressed in the literature.

Being motivated by the gap in the existing literature as well as the potential of HAR in Salat as delineated above, in this study, we explore a new approach to activity recognition in Salat using a smartwatch. Here, we attempt to leverage the smartwatch considering it a convenient wearable, easy-to-use and non-distracting. Besides, we aim to make our approach robust by addressing the variations in signals for different subjects, taking care of noises induced by irregular activities, and predicting the complete prayer details with reliable accuracy.

Chapter 5

Problem Formulation and Research Challenges

From the aforementioned research gap in the literature, it is clear that the existing literature is yet to provide a convenient and robust solution for activity recognition in Salat. This motivated us to conduct this study, which evolves with three research questions. In this chapter, we will state the research questions and specify corresponding research challenges.

Our Research Questions

In this study, we focus on the following set of research questions.

RQ1 How prevalent are mistakes in Salat among people? What types of mistakes are more common among them? Can we find the mistakes and their frequencies through a mixed-method analysis?

RQ2 Do people need technological assistance for improving their prayer? How willing are they to accept such technological assistance? Can we find the need for technological assistance and its acceptability through a mixed-method analysis?

RQ3 If people welcome technological interventions or assistance for their Salat, then can we help them in improving their prayers by leveraging a more convenient device, with improved accuracy compared to that of the solutions existing in the literature?

Research Challenges of Our Study

In the process of answering the research questions, we envision some research challenges that seem to inevitably entail our study. The potential research challenges are as follows.

Revealing Common Mistakes in Salat and Acceptability of Technological Assistance for Salat

The existing HAR studies on Salat urged that, owing to the inherent complexity of the process in Salat, people tend to make mistakes in Salat for diversified reasons such as lack of knowledge, having sickness (e.g., Alzheimer's), etc. Thus monitoring Salat to check for completeness or correctness [2, 31, 39] is of utmost significance, as reported by the studies. However, from these studies, unfortunately, we do not get any insight into how frequently people make mistakes in prayer and how willing they are to accept this solution. Religion, being a sensitive and private topic, as religious worship is between a person and the Almighty, we need to explore whether people would allow technological assistance in their worship. Here, we need to perform methodical studies to explore these aspects which will eventually answer RQ1 and RQ2. Performing the methodical studies presents one of the main challenges in this research.

Variability in Sensed Signals

After establishing the requirement, our task is to find a convenient alternative for technological assistance enabling HAR in Salat. For the purpose of HAR, other than smartphones, the next most ubiquitous wearable used is the smartwatch. Even though smartphones come first from the perspective of ubiquity, smartwatches score far more than smartphones, as the smartwatches are worn on the wrist posing no challenge in offering Salat [42]. Nevertheless, with their adoption in activity recognition in Salat, some challenges entail, which need special handling.

Although smartwatches are more comfortable to wear continuously, due to their placement on the wrist, they record a higher variability in the sensed signals [119]. This happens as smartwatches record arm movements, which vary substantially more than body movements in general. Accordingly, if we want to recognize the body movements, then arm movements introduce additional variability. This variability can mask the whole-body activities giving rise to a significant degradation in HAR performance [119]. In the case of Salat, only a few activities such as Takbeer entail only arm movements. For most of the activities in Salat (such as bowing, prostrating, etc.), body movements appear the most having some hand movements. Hence, activity recognition in Salat could be more challenging using smartwatches than smartphones.

Handling Variations in Salat

We have already discussed the variations in the worship Salat in Section 2.3.1. As such, the variations in postures during Salat for men and women can sometimes become highly significant, which is very likely to introduce substantial variations in the signal patterns sensed from Salat of a different gender. The model intended for HAR in Salat needs to be able to handle such variations. Besides, for the purpose of accurate recognition of the Witr Salat as well as the

Salat of the people who give Takbeer (or raise hands) before or after the short-standing period, Takbeer needs to be recognized. However, Takbeer is not yet recognized by any of the existing research studies on HAR to the best of our knowledge. Moreover, the same activity can be performed differently (with different hand orientations, with a different pose, etc.) by different persons, which eventually results in a difference in the sensed signal patterns [116]. To handle this, having a large dataset consisting of data from people with different demography is of utmost significance. To sum up, it is a crucial challenge to make the HAR model capable of recognizing activities correctly in the presence of these variations in the activities during Salat.

Handling Extra Activities

As stated in Section 4.8, it is common among people to perform some extra activities that are not regular such as fixing garments, rubbing eyes, etc. as shown in Figure 4.3, which do not invalidate the prayer [115]. However, as they are mostly performed by hand, and the worshipper wears a smartwatch, then these extra activities are likely to introduce additional signals and these additional signals will be captured by the smartwatch engendering noises in the sensed signals. In the HAR literature, these types of non-relevant activities or non-activities are referred to as *Null* activity or *Null* class [120]. In our study, we also refer to these extra activities as Null activities. We need some special handling to mark these noises induced by Null activities and ignore them while recognizing ritual activities in Salat. Marking such noises, which may arrive now and then, poses a challenge in our study.

Demand for Near-Perfect Accuracy

As Salat is a very important worship for Muslims, having its Rakat count and count of other activities with a near-perfect accuracy has no alternative when it comes to assessing technological assistance. Thus, even though the HAR model for Salat gives high accuracy for individual activity recognition (for example, around 90%), errors in the model can give wrong counts undermining the credibility of the model to the users. To overcome such cases, blending domain knowledge in a post-processing step perhaps has no alternative. Here, achieving such a delicate blending presents a noteworthy research challenge.

Chapter 6

Common Mistakes in Salat and Acceptability of Technological Assistance to Overcome Them

To reveal the common mistakes in Salat and to assess the acceptability of technological assistance to improve Salat, we conduct an online survey. The survey contains questions related to mistakes in Salat as well as preferences in adopting technological assistance during Salat. The survey is completely anonymous and the participants voluntarily fill it out.

Survey Goals

The first goal of our survey is to get a sense of the frequency and nature of mistakes people usually make in prayer. Then, we check whether people would appreciate technological assistance in their prayers to improve their prayers. We also assess the willingness of the participants to wear a wearable (such as smartwatch, fitness band, etc.) while being in Salat to avail the technological assistance.

Justification behind Adopting A Self-Reporting based Survey

In our study, we adopt a self-reporting based survey to collect responses from the participants on common mistakes in Salat and the acceptability of technological assistance to overcome them. The notion of adopting such a self-reporting based survey is common in the literature. Existing research studies on exploring various types of religious experiences [121–124], judging computer efficacy [125], assessing social desirability [125,126], measuring personality [125,127], exploring digital well-being [128], criminology [129], psychopathology assessment [130], assessing openness in research content sharing [131], investigating psychological disorders [132, 133], etc., have already utilized self-reporting based surveys. Moreover, the research communities on HCI [131], CSCW [128], and ubiquitous computing [132, 133] often leverage self-reporting based surveys. Accordingly, in our study, we utilize a self-reporting based survey. Nonetheless, a longitudinal study over performing Salat by different persons could be conducted in the future to explore the common mistakes in Salat through an alternative lens.

Overview of Our Questionnaire

We collect demographic information at the beginning of the survey. The demographic information covers age, gender, country, education, occupation, and religion. Next, we ask the participants about their usage and ownership of technological devices such as smartphones and smartwatches. Afterwards, we try to asses their regularity in prayer and to what degree they are willing to improve their prayer quality and quantity. Then, we ask them about their frequency of mistakes or confusions during prayers. We ask them about five specific types of mistakes such as forgetting the counts of Rakat, bowing, and prostrating as well as forgetting to recite another chapter from the Qur'an after the first chapter, i.e., Surah Fatiha, and forgetting to sit for Tashahhud. We take responses for these questions in a 5-point Likert scale [134].

Next, we take their opinions regarding availing technological assistance for improving prayers, i.e., whether they would welcome if their devices such as smartphones, smartwatches, etc., assist them to improve their prayer quality. There is also a qualitative question so that they can express the reasons behind their opinions. Finally, we ask them about their willingness to pray wearing a smartwatch or fitness band if they have to do so to avail the above-mentioned assistance. There is a qualitative question too so that they can share the reasons behind their extent of willingness.

Survey Participant Demography

The sampling strategies used in our survey are convenience sampling [135, 136], referral sampling [135, 136], and list-based sampling [135, 136]. First of all, we disseminate the survey through email and social media to the people accessible to us, which falls under convenience sampling. Besides, we request each of them to refer this to other people they think are eligible or circulate it among their own networks, and this covers referral sampling. Furthermore, we email our questionnaire to the faculty members of different universities in Egypt, India, Indonesia, Iran, Iraq, Malaysia, Saudi Arabia, the United Arab Emirates, and some other countries. We collect their email addresses from their institutional web pages, and therefore, this stands for list-based sampling.

Table 6.1: Demography of the survey participants

Gender	#Participants	%
Male	84	67.7
Female	40	32.3
Age (Years)	**#Participants**	**%**
Child (1-14)	0	0
Young adult (15-24)	23	18.5
Middle-age (25-44)	57	46
Older adult to average retirement age (45-64)	35	28.2
Retired (65+)	9	7.3
Literacy (Highest level of educational degree achieved)	**#Participants**	**%**
High school or equivalent	9	7.3
Some college but no university degree	8	6.5
Diploma or equivalent	1	0.8
Bachelor's or equivalent	34	27.4
Master's or equivalent	14	11.3
PhD or equivalent or above	56	45.2
Prefer not to disclose	2	1.6

Country of Living	#Participants	%
Bangladesh	54	43.5
United Arab Emirates	27	21.8
Egypt	13	10.5
United States	11	8.9
Others	9	15.3
Occupation	**#Participants**	**%**
Student	45	36.3
Teacher	59	47.6
Homemaker	7	5.6
Others	13	10.5

We get responses from 126 participants in total who are from 15 different countries. Among them, two participants report that they are not religious or spiritual. Therefore, we had to discard their responses. We do so as we only target the Muslim population in this survey since this study is all about one of their religious worships. Therefore, we do not ask any further questions to the participants who reported their religion to be anything other than Islam and discard their responses. Thus, the count of our responses becomes 124.

The majority of our participants are male and educated. Regarding age diversity, we get responses from different age ranges except for children. As Salat is not mandatory for children, we do not attempt to reach them either. As per occupation, our participants cover students, teachers, IT professionals, engineers, homemakers, etc. Table 6.1 presents the demography of our participants.

Quantitative Analysis over the Survey Responses

We analyze the quantitative survey using descriptive statistics such as frequency, percentage, mean, etc. For correlation analysis, i.e., to analyze the relationships between demographic and other factors with any variable of interest, we use Chi-squared test [137]. Additionally, we use Mann-Whitney U test [138] and Kruskal-Wallis test [139] to compare whether there is any statistically significant difference in the dependent variable for the independent groups.

Here, in response to the question regarding the experience of using technological devices (namely smartphone and smartwatch), we find that all of the survey participants (100%) have experience in using a smartphone and currently own a smartphone. On the other hand, 43% of the survey participants have experience of using smartwatches, and currently, 31% of them own a smartwatch. Therefore, the pervasiveness of the smartwatch is somewhat less in comparison to that of the smartphone, which portrays the reality as the smartwatch was introduced later [79].

Regarding regularity in prayer, we find responses from almost all types of people such as regular, somewhat regular, not regular at all, and so on. However, the majority of the participants are

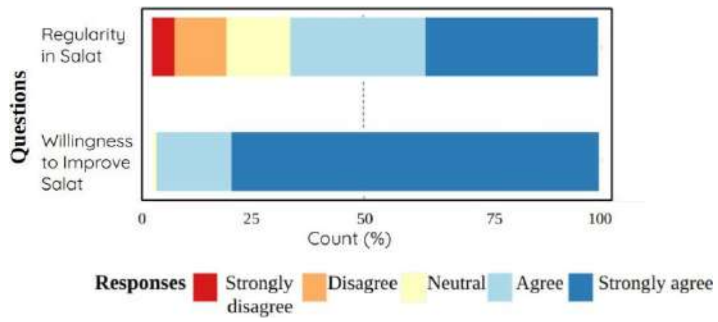


Figure 6.1: Summary of participants' regularity in prayer and willingness to improve prayer

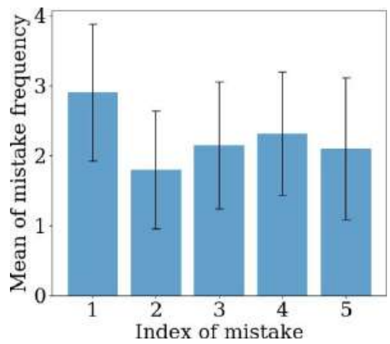
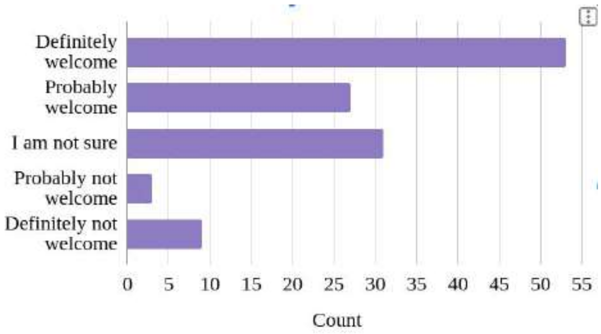


Figure 6.2: Summary of the frequency of different mistakes in Salat (1 = forgetting count of Rakat, 2 = forgetting count of bowing, 3 = forgetting count of prostration, 4 = forgetting to sit for Tashahhud, and 5 = forgetting to recite another Surah after Surah Fatiha)

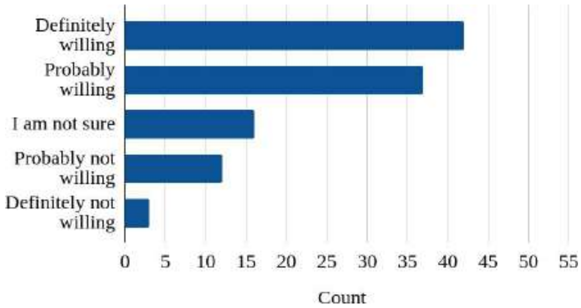
regular in their prayers. Interestingly, when it comes to the willingness to improve prayer quality and quantity, almost all of them (above 80%) respond to be willing to do that. Figure 6.1 depicts these findings.

Regarding mistakes, we find the most common mistake is forgetting the count of Rakat, i.e., forgetting how many Rakat a person has prayed. On the contrary, forgetting or getting confused about Ruku or bowing is the least frequent mistake as reported by the participants. Other types of mistakes also happen with varying frequency. Figure 6.2 presents the mean frequencies of the mistakes taken in the 5-point Likert scale (1 = never, 2 = rarely, 3 = occasionally, 4 = often, 5 = very frequently).

We observe that 42 participants ($\sim$ 34%) report making at least one type of mistake often or very frequently. This further clarifies that mistakes in Salat are real and common among people. We perform Mann-Whitney test to find whether there is any statistically significant difference in the mistake frequency among the male and female participants and find there is none ($W = 1852$, $P = 0.832 > 0.05$). Similarly, we do not find any significant difference among people of different age groups or different levels of education, regularity in Salat, etc., with their frequency of various types of mistakes.



(a)



(b)

Figure 6.3: Summary of the responses to the questions regarding (a) eagerness to explore technological assistance in Salat and (b) willingness to pray wearing a convenient wearable (smartwatch, fitness band, etc.)

The next important finding is the majority of the participants ($\sim$ 70%) express their eagerness towards welcoming technological assistance to help them in their prayers. We do not find any statistically significant correlation between the response to this question, i.e., welcoming technological assistance in Salat, and the demographic factors, as well as regularity in Salat. Figure 6.3a portrays the summary of the response to this question.

Regarding the willingness of the participants to pray while wearing a smartwatch (or a similar wearable) to avail the technological assistance in Salat, the majority of the participants ($\sim$ 85%) express their willingness to do so. Here again, we find no significant association between willingness to wear a wearable with mistake frequency and demographic factors. Figure 6.3b presents the summary of the responses of the participants to this question.

Qualitative Analysis of the Survey Responses

In our survey, we have a few qualitative questions. We collect free-text responses of the participants in response to these questions and perform thematic analysis on these responses [140, 141]. To do so, we go through the responses several times and systematically identify and cluster the themes or codes that are found to be present in the responses. We generate the codes, compare them, and reiterate the responses until we have a consistent codebook. We further organize our codes into higher-level categories. Finally, we give each of these high-level categories appropriate names, which eventually conclude the themes.

Eagerness towards Exploring New Technologies for Improving Prayer

Irrespective of gender, age, occupation, or regularity in prayer, the participants express their keenness to explore technologies that would help them to improve the quality and quantity of their prayers by overcoming the mistakes or at least reducing the mistake frequency. As per their responses, the primary reason behind the keenness is the fact that technological devices have already become an indispensable part of today's life and people rely on them for various purposes. As these technological devices are already around people almost all the time, and therefore, if they are capable of providing any good regarding Salat, according to them, people should embrace it. In this regard, some of the participants state the following:

> *"This will really be great if my devices help me to improve my Salah, because we use devices like smartphones or smartwatches on regular basis. So, it will help me to track my improvement easily."* (P7)
>
> *"It might provide handy, easy-to-access information."* (P40)

Beneficial Initiative for the Muslims

The comments of most of the participants regarding the idea of helping people to assist in their Salat are inspiring. Many express their hope that success in this work would benefit the Muslims greatly. As Salat is the most important religious worship for Muslims, and at the same time people are now highly dependent on technological devices, making these devices capable of helping the Muslims in their prayers should help them to a great extent. Below, we quote the comments of three participants in this regard.

> *"It's going to be a very beneficial tech assistance insha'Allah"* (P12)
>
> *"If these can be made possible, it will be extremely helpful"* (P81)
>
> *"Thanks a lot for having these concerns in your study and I wish you a successful outcome."* (P91)

Permissible or Not - A Bit of Suspicion

Just like we find most people welcoming the idea of technological assistance for improving prayer, some people express their confusion regarding the permissibility of such solutions. This means they want to be confirmed at first whether using such solutions would be permissible by Islam or not. Some of the participants suggest checking with the scholars and taking clearance regarding the permissibility of using such solutions from them. Moreover, few of the participants, think that the Islamic guidance regarding the mistakes in Salat is enough. We present responses relevant to this regard below.

> "*I am interested, but I would want to make sure it is permissible to use a device to assist my prayer.*" (P87)
>
> "*It needs an Islamic scholar Fatwa (advice)*" (P91)
>
> "*It feels like a disturbance. Islam has provided a solution by doing sujud sahw.*" (P103)

No Disturbance or Interference during the Prayer

Quite a few participants raise their concerns that technological support for Salat should not mean any disturbance or interference while a person is praying, i.e., no need for real-time notification during prayer about any mistake or so. As Salat in Islam means connecting with the Almighty, they are afraid that such notifications during Salat might break their concentration and divert their attention towards the device, which can hamper the essence of Salat. We receive the following comments pertinent to this concern.

> "*... there is a possibility that my concentration would be diverted toward the smartphone/watch. The subconscious part will be focusing on the device's instructions. During Salat, 100% concentration should be on Salat. If you come up with a solution that does not occupy a part of my mind towards that device, then the solution is probably welcome.*" (P67)
>
> "*If the devices can assist properly, no problem; but in times of praying, I don't want the natural environment to be harmed.*" (P29)
>
> "*During prayer, it is the only time that I choose to leave my phone in another room; or, if I go to the mosque, I leave it in the house or my office. I don't want it to ring/vibrate or light up and ruin this unique feeling.*" (P80)

Smartwatch Sounds Good

The majority of the participants express their willingness to pray wearing a wearable (smartwatch, fitness band, etc.) to avail the technological assistance for improving their prayers. Different participants mention different reasons for supporting this. Some express that, if such devices can actually help, then people should welcome the solution. To many, smartwatches are the most convenient option for this purpose, as they are used to wear smartwatches most of the time. People also prefer smartwatches over smartphones due to the proximity, ease of access to notifications, and less distraction compared to smartphones. The following comments portray the findings.

> "*I would prefer smartwatch because that's on me for a lot of time of the day*" (P52)
>
> "*If that helps, why not*" (P40)
>
> "*It depends how. If it's a display on a smartwatch that I can glance into it if I got confused, then maybe. If it will require me to take more movements in my prayers, then I won't prefer to use it. Also, yes for a smartwatch but no to a smartphone...*" (P37)

Summary of Our Findings

To summarize, this survey helps us to gain deeper insight into our research topic and helps us to formulate our problem in a better way. Through this survey and a comprehensive analysis over the survey data, we get assured that mistakes in Salat are real and in fact pretty common among people. People in general, whether regular or irregular in their prayers, want to reduce the frequency of mistakes, and desire to improve their prayer quality and quantity of their prayers. Therefore, if technological devices, which they already use in their daily lives, offer help in achieving these objectives, they are mostly willing to explore that. However, this needs to be non-distracting and convenient. Therefore, we attempt to recognize activities in Salat using a smartwatch, in such a way that, that fulfills these expected requirements.

Chapter 7

Proposed Methodology

This chapter contains a step-by-step description of our proposed methodology for activity recognition in Salat, including the details of our data collection and analyses. Figure 7.1 depicts the pipeline of our methodology from a high level.

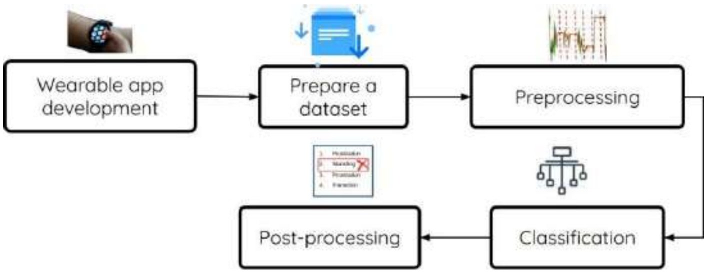


Figure 7.1: Pipeline of our proposed methodology

As per our proposed methodology, we first develop an app for our wearable to use in data collection. Then, we perform data collection using the wearable, having the app installed in it, and prepare a dataset accumulating our data collected from different users. We preprocess the raw data in our dataset and use multiple approaches for classification over the dataset. Subsequently, we perform some postprocessing to augment the classification results by incorporating domain knowledge and predict the final activity recognition results. We elaborate on each of these steps of operations in the following sections.

Wearable App Development

The wearable device we leverage in this study is a smartwatch. The model of the smartwatch is Samsung Galaxy Watch Active 2 [46] as shown in Figure 7.2. It is programmable, widely available, and equipped with the sensors needed for activity recognition. As it runs Tizen OS, we develop a Tizen service app in Tizen studio and install it in the watch exclusively for logging the sensor data while a person is praying wearing the smartwatch.

Figure 7.2: Samsung Galaxy Watch Active 2 [5] - smartwatch

This app records the sensor values from the accelerometer, gyroscope, and magnetometer pertaining to all three axes, with the timestamp, in files. Data is collected at a sampling rate of 25 Hz. This rate is sufficient for capturing human body motion, since 99% of the energy of human motion is contained below 15Hz [142].

Data Collection

Due to the unavailability of any study as well as any dataset for activity recognition in Salat using smartwatches, we prepare a dataset on our own and use it for our study. We collect data from 30 subjects individually. We use convenience or opportunity sampling strategy for recruiting the subjects [135, 136]. All subjects have agreed to the usage of the recorded data for scientific and research purposes. They all are from Bangladesh and currently living here. The demography of the subjects is given in Table 7.1.

We request the subjects to wear a smartwatch on the wrist of their left hand as per their convenience and perform four units of prayer. Figure 7.3 shows a subject under data collection. Here, as Takbeer is only performed once in the prayer only at the beginning, we request the participants to repeat Takbeer four additional times after completing their regular prayers for the purpose of our data collection. We capture and record video of the whole procedure using a

Table 7.1: Demography of our subjects who participated in data collection

	Count (Percentage)
Gender	
Female	13 (43%)
Male	17 (57%)
Age (years)	
12-25	6 (20%)
25-35	12 (40%)
35-45	4 (13%)
> 45	8 (27%)

Figure 7.3: Snapshots captured during our data collection

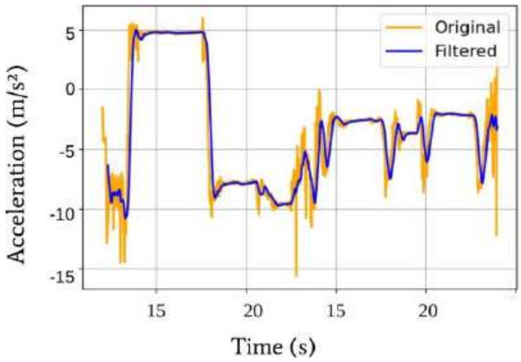


Figure 7.4: Comparison of the original (raw) and filtered data captured during our data collection

timestamp camera app to facilitate the task of ground-truth labeling. We explicitly take consent from the participants regarding the video capturing.

Preprocessing of Raw Data

The collected accelerometer signal is preprocessed before it is fed into the classifiers. Preprocessing includes different tasks such as denoising, labeling, and segmentation.

Denoising

Due to various issues such as calibration problems, device malfunction, deployment issues, etc., wearable sensor data frequently contain noises. Therefore, it is a common practice to filter the data and denoise accordingly before going to the next steps of classification. As such, to smooth out the raw data, we use the notion of moving average filter [143]. Many primary research studies on activity recognition use this notion for the purpose of denoising and smoothening [3, 144, 145]. Figure 7.4 shows the denoising process using a moving average filter.

Labeling

As our target is to recognize all the steady states of Salat such as standing, bowing, sitting, etc., along with Takbeer, we divide each prayer unit into seven steps - Takbeer, standing, bowing, short-standing, prostrating, sitting, and transition - and label accordingly. Here transition includes all transitional activities such as going from standing to bowing, sitting to prostration, and so on. We perform the labeling task manually with the help of the recorded timestamped video.

Segmentation

Segmentation of the collected signals is a very crucial step for activity recognition. Segmentation refers to dividing the signals into chunks of windows for further processing. The chunk size is generally problem-specific [146]. There are three basic types of windowing used in HAR - activity-defined window segmentation, event-defined window segmentation, and sliding window segmentation [146]. In activity-defined windows, the initial and end points of each window are picked by recognizing patterns of activity changes, whereas the window is constructed around a detected event in event-defined windowing. However, in the case of sliding window, data is divided into fixed-size windows with no gaps between them, and in certain circumstances, the data can even overlap. Among the three alternatives, the sliding window is the most used segmentation method in HAR [147]. Accordingly, in our study, we use the overlapping sliding window technique.

As the size of the window directly impacts the segmentation accuracy, windows should be large enough to ensure that at least one cycle of activity is contained and the comparable movements are distinguishable [48]. Keeping this in mind, we analyze the activities in Salat and find that the steady activities may take from a second or so (such as short-standing and sitting between two prostrations) to a few minutes (such as standing). This time varies from person to person. Accordingly, we choose our sliding window length in such a way that we can capture all the steady states correctly. The algorithm for preprocessing is presented in Algorithm 1. Similar steps, as adopted in the Algorithm 1, have also been used in other existing studies on activity recognition [148, 149].

Classification

The most conventional way of classifying sensor data in HAR studies is adopting various supervised machine learning classifiers, either classical machine learning or deep learning, for training a model and using it for final prediction [12, 150]. Therefore, we adopt machine learning-based classification as our baseline methodology for activity recognition in Salat. Rigorous analyses of the performance of the baseline methodology help us to understand its limitations.

Algorithm 1 Algorithm for preprocessing raw signal

Input: D accelerometer data of x, y, and z axes with labels,
n total number of the collected data
Output: Set of Segments S
for each axis in D **do**
apply moving average filter of window k
end for
$samplesPerWindow \leftarrow frequency \times windowSize$
$start \leftarrow 0$
$end \leftarrow 0$
while $end \leq n$ **do**
if $end \quad 0$ **then**
$start \leftarrow end - samplesPerWindow \times overlap$
else
$end \leftarrow start + samplesPerWindow$
end if
$s \leftarrow D[start : end]$
$S.append(s)$
end while

After realizing the limitations of the baseline methodology, we come up with an improved methodology for activity recognition in Salat. For both of these cases, we first perform the preprocessing as stated above is performed beforehand. In the next subsections, we elaborate steps of these methodologies.

Baseline Methodology using Machine Learning Classifiers

We use a bunch of classical machine learning classifiers as well as a deep learning model on our collected data after preprocessing. The classical machine learning classifiers need feature extraction after the preprocessing stage, whereas the deep learning model does not require anything so. We present the classical classifiers and deep learning model under our investigation below.

Feature Extraction for Classical Machine Learning Classifiers

Pinpointing the most important attributes in each segment of preprocessed data is an important task for classical machine learning classifiers. This task is called *Feature Extraction* [151], which presents an important aspect of developing HAR systems [12]. The use of features rather than raw data generally enhances classification accuracy as reported in the literature [151].

Accordingly, in our study, we summarize each resulting segment in the preprocessed data to a fixed number of features to feed the classical machine learning classifiers, i.e., we summarize one feature vector per segment. The feature vector contains a number of statistical measures.

Table 7.2: Features extracted for classical machine learning classifiers

Domain	Features
Time	Mean, Max, Median, Standard deviation, Variance, Skewness, Kurtosis, RMS, Inter-quartile range, Zero-crossing rate, AUC (Area under the curve), MAD (Mean absolute deviation), Number of peaks, Peak-to-peak distance, Entropy, and Absolute energy
Frequency (FFT)	Mean, Median, Standard deviation, Inter-quartile range, and Power
Others	Pairwise correlation between all three axes

Examples of statistical measures include mean, median, standard deviation, etc. We extract the features from both time and frequency domains. We extract 16 features in the time domain and 5 features in the frequency domain. Additionally, we consider pairwise correlations between the three axes and include them as features. Moreover, we extract all these features from the three accelerometer axes a_x, a_y, and a_z. Table 7.2 lists down all these features extracted from each of the axes.

Thus, in total, we extracted 21 features in total for each of the three accelerometer axes in addition to three pairwise correlations, which sum up to 66 features per segment.

Classical Machine Learning Classifiers used in Our Baseline Methodology

We use the features extracted from the raw inertial data sensed by the smartwatch corresponding to the user's activities, to train and test different supervised machine learning classifiers. Here, we use four prominent classical supervised machine learning classifiers for the classification of our feature vectors namely Random forest [152], J48 decision tree [85], Naive Bayes [153], and Logistic regression [154]. We select these classifiers considering their high accuracies in the existing HAR studies [88, 90, 155, 156].

Deep Learning Model used in Our Baseline Methodology

Recently, Deep Learning (DL) methods such as recurrent neural networks (RNN), LSTMs, autoencoders, and their variations have been proven to provide state-of-the-art results on challenging activity recognition tasks with little or no data feature engineering [48]. This inspires us to explore a deep learning model for our task. In our case, the diversity of the signal varies from person to person. Accordingly, as found from our investigation over the classical machine learning classifiers, the diversity exhibits to be the main factor responsible for lower classification accuracy. Considering the fact, we adopt the model presented in the study [116].

The model adopted from the study in [116] learns to automatically disentangle domain-agnostic and domain-specific features. Between these two types of features, the domain-agnostic features are expected to be invariant across various persons. Here, the domain stands for a specific person's data. To incorporate the variations across domains, in this study, a generative method is developed on top of the variational autoencoder (VAE) framework [157]. Here, two probabilistic encoders are used to induce two groups of latent representations, i.e., domain-agnostic and domain-specific representations. The domain-agnostic representations capture the common information pertinent to conducting a certain class of activity. Besides, the domain-specific representations can reflect the unpredictable factors that induce variations among training domains such as different environments, physical conditions, etc. To effectively disentangle these two latent spaces, a novel Independent Excitation mechanism was developed. Besides, by removing domain-specific representations, the resulting domain-agnostic latent space is made to be more invariant to different domains than the original data. As a result, the model can generalize better to new unseen target domains. In this manner, the model in [116] proves itself to be robust for cross-person or user-independent HAR. This model undergoes experimentation with three benchmark datasets [120, 158, 159] and yields better accuracy than many state-of-the-art DL models. Therefore, we keep all the parameters of this model intact to check how it performs in our case.

Different Approaches of Classification using Machine Learning

Salat involves both static/steady and transitional activities in an alternating manner.

We are interested to recognize the steady states along with Takbeer in isolation. Pertinent to the recognition, the literature reports that examining the transition period before a steady/immobile state can improve the performance of steady state recognition [160]. Being motivated by this, we attempt to explore two different approaches with our machine learning classifiers. In the first approach, (Approach-1), we do not attempt to recognize the transitions individually, but rather group them together into a single class 'Transition'. Therefore, in this case, we have seven classes - five steady states, Takbeer, and Transition (T) as different individual classes to recognize. In the second approach (Approach-2), we recognize all the steady and transitional activities mentioned in Table 1.1 as different individual classes.

Besides, in the HAR literature, when the classification problem involves both steady and transitional activities, many studies often separate the steady and transitional activities first and then perform more granular classification over their activities of interest [145, 161, 162]. Therefore, in both of our approaches, we first attempt classification altogether and then we attempt classification in a hierarchical fashion similar to the existing studies [145, 161, 162]. In the former one, we classify using a single classifier for recognizing all the classes. However, in a hierarchical way, we first recognize steady and transitional activities. For this purpose, we first classify the segments into steady and transitional using classical machine learning classifiers. Afterwards, we use separate classifiers to classify the steady and the transitional states and then combine their outcomes to determine the final classes.

Classification with Improved Methodology using Semantic Rules andDTW

Our experiments reveal some limitations and low accuracies of our baseline ML-based approaches, which we are going to present in detail in our next chapter. As Salat is a religious worship, any technology for Salat should provide near-perfect accuracy. However, the ML-based approaches do not meet this expectation. The underlying reason behind this is the inter and intra-class variabilities which present a well-known challenge in the HAR literature [163]. Intra-class variation refers to the fact that the same activity can be performed differently (e.g., at different speed and style) by different subjects, which result in a variation in the signals belonging to the same activity class [163, 164]. On the other hand, sometimes different action classes have similar patterns, which is related to inter-class variation [163]. Considering both aspects, our target is to design a new methodology that would be generic enough with higher discriminative power to have a clear realization over these variations.

However, in our problem, recognizing the transitions in Salat is necessary for the purpose of having enough context information to infer the steady states as the context information can substantially improve the recognition of steady states, [160]. To recognize the transitions with better accuracy, a viable alternative to ML-based approaches can be template-matching [22]. Template matching finds the distance or correlation of a given signal segment with some pre-defined templates. Based on the distance or correlation, template matching finds out the class of the test signal [165]. We find many prior research studies adopting template matching [26, 110, 165–167], especially for transitions [166, 167]. The reason behind this adoption is that transitions generally span a very short time. In the case of short-time samples, traditional features are unstable and cannot describe the actions effectively [168]. In fact, the study in [169] observes a higher generalization ability of the template-based methods compared to several ML classifiers while classifying activity data collected with a wrist-worn accelerometer. Therefore, in our proposed methodology, we adopt Dynamic Time Warping (DTW) [170], which is a famous template-matching algorithm for classifying transitions. The use of template matching using DTW has the advantage that, it works well even when the training data is limited [169]. However, to recognize transitions, we first distinguish the steady and transitional states. At the same time, using domain knowledge, we develop some semantic rules for recognizing some of the steady states. Finally, the results of both of these stages are postprocessed incorporating domain knowledge to detect and fix misclassifications and enhance accuracy. Figure 7.5 presents the pipeline of our proposed improved methodology integrating all these stages.

State Recognition

As stated earlier, it is often practiced in the HAR literature to distinguish the static and transitional states first and then, do further classification. Therefore, at the top layer of the

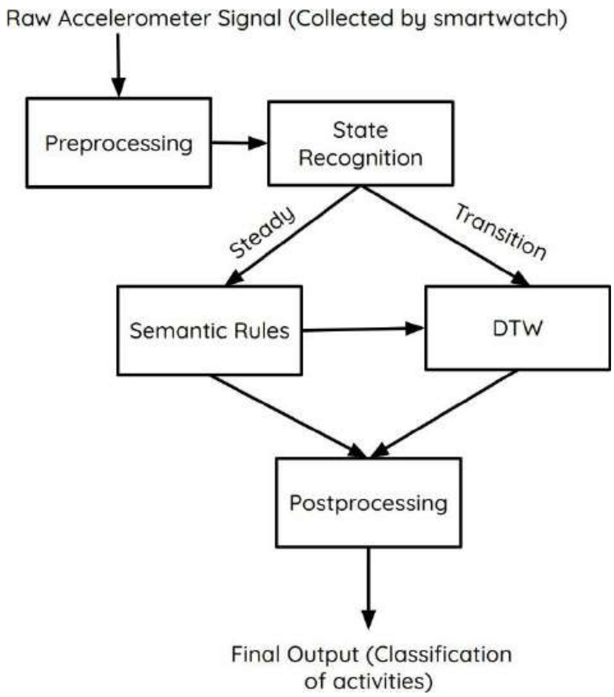


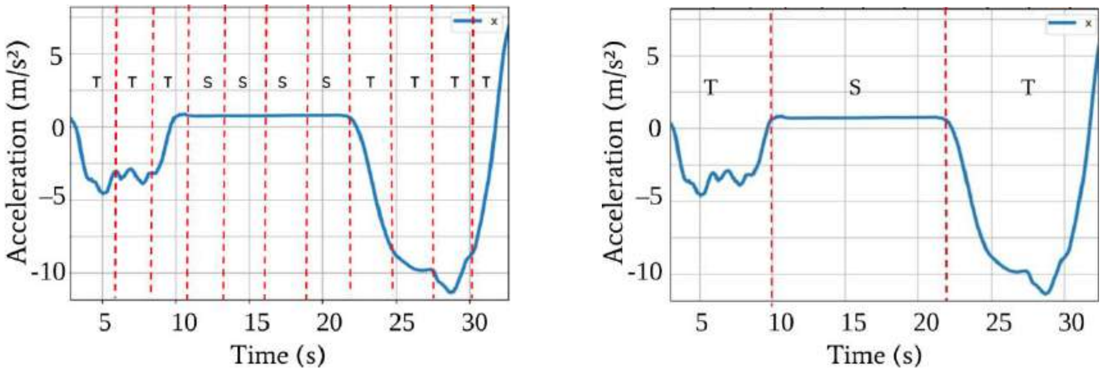
Figure 7.5: Pipeline of our proposed hierarchical methodology

proposed hierarchical methodology, we have our state recognition stage. Besides, the DTW-based classification stage demands the transitions be separated fully from the steady states. Accordingly, after preprocessing the signal, we determine the state to which a signal segment belongs, i.e., whether it is part of a steady activity such as bowing, prostrating, etc., or a transitional activity such as standing to bowing, prostrating to sitting, etc. Existing research studies often employ this step at the beginning of their pipeline to distinguish the static and dynamic activities [145,161,162,171]. The mean, range, and variance of the possible acceleration values as well as periodicity in the acceleration data many a time differ slightly between consecutive activities in Salat, however, they differ substantially over different states. Hence, as a classical approach, we can deploy machine learning to train a model to learn the characteristics of steady and transitional states. Then, we can feed our preprocessed signal segments to such a model to predict the state of each segment.

However, in the literature, state recognition has also been done using normalized signal magnitude area (SMA) [162,171]. In our study, we also explore this approach. Here, to calculate SMA, linear acceleration, i.e., acceleration due to body movement is separated from the total acceleration signal by discarding the gravity component. This signal is used to calculate the normalized SMA using Equation 7.1 [171].

$$SMA = \frac{1}{t}\int_0^t |x(t)|dt + \int_0^t |y(t)|dt + \int_0^t |z(t)|dt \tag{7.1}$$

Here, $x(t)$, $y(t)$, and $z(t)$ refer to the body components of the x, y, and z-axes of the



(a) Before merging adjacent segments (b) After merging adjacent segments

Figure 7.6: Impact of merging adjacent segments based on prediction labels (S = Steady and T = Transition)

accelerometer, respectively. However, the separation of body components is typically performed using a high-pass Butterworth filter of low order with a cutoff frequency of 1Hz [172]. We use the same in our study as we also deal with human body movement. An appropriate threshold value is determined such that a normalized SMA value below the threshold will refer the user to be in a steady state and the user to be in a transitional state otherwise [171].

As mentioned already, in the next stages, our target is to classify each transition between the steady states using DTW. Here, the input is a complete set of activities covering the transitions in Salat such as bowing to short-standing, prostrating to sitting, etc. However, due to the usage of fixed-length sliding window segmentation, these transitions are segmented into multiple chunks. To better distinguish each individual transition and steady state, we merge a segment with its neighboring segment(s), if its predicted label is the same as its neighbor(s). This eventually groups contiguous transitional segments into one complete transition and the same for the steady segments. This type of merging of segments exists in the literature [173]. Figure 7.6 portrays this process in our case. As shown in the figure, after the completion of this stage, the signal is segmented into alternative steady and transitional activity segments where each segment represents either a complete transitional activity or a complete steady activity in Salat such as bowing, standing, etc.

Classification of Steady States using Semantic Rules

Next, we take a deeper look into the prayer postures of an individual and corresponding accelerometer signals. We do so to come up with some rules for distinguishing some of the steady states. For example, based on the postures of the standing and bowing phases as well as the transition between them, we find some correlations of the values of the accelerometer in different axes, using which we can distinguish the steady states. Accordingly, we can set some semantic rules to distinguish the states. The term semantics refers to the study of meaning. In the HAR literature, semantic approaches refer to incorporating the human understanding

Figure 7.7: The three axes of a smartwatch

of an activity [164]. More specifically, semantics interpret an action as a relation between its features (e.g., body parts, corresponding objects, scenes, etc.). In activity recognition, semantic understanding enables users to apply prior knowledge in the recognition process [164]. In our case, we derive some rules based on the understanding of the activities and corresponding sensor signal patterns as well as prior knowledge, etc., and therefore, we term them as semantic rules.

Before we go to the details of these rules, it is worth mentioning that, while performing activities wearing a smartwatch, the accelerometer of the smartwatch measures the acceleration in m/s^2. This applies to the watch on all three physical axes (x, y, and z) as shown in Figure 7.7. The acceleration covers the force of gravity too. Sustaining these aspects, we derive semantic rules applicable to different states and positions.

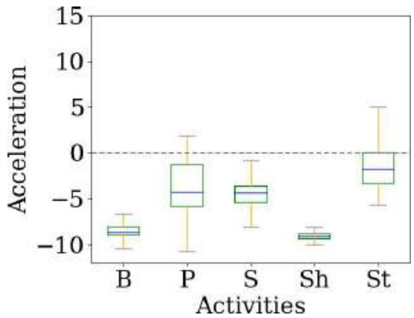
When a person is in the standing position, his hands are placed either on his chest or belly, as shown in Figure 7.8. Therefore, when he goes from standing to bowing, his hand first moves slightly outward and then moves straight downward. As per the axes of the watch shown in Figure 7.7, it is clear that the outward movement will result in acceleration towards the negative x-axis and the downward movement towards the negative z-axis. Here, the acceleration along the y-axis is not very significant. Besides, whatever the placement of the hands is while standing, the resulting acceleration from standing to bowing always follows the same pattern. Next, while in the bowing position, the value of the x and z-axes of the accelerometer should always be negative. This intuition complies with our findings from the boxplots depicted in Figure 7.9a and Figure 7.9c. Here, Figure 7.9 shows the summary of the real values collected from 30 subjects. Besides, as the acceleration towards y-axis is not much significant compared to the other two, the differences of y from x and y to z are always positive in the bowing position. We find that, if we develop rules combining these conditions, we can correctly recognize bowing among all the steady states.

In Salat, a person is supposed to go to the bowing from the standing position. He can also (wrongfully) go to the bowing position from the short-standing position, however, he can never go to the bowing position from the sitting position. From the sitting position, he has to stand up and then bow down. Therefore, the previous steady state of bowing should be either standing or short-standing. When a person goes to the standing position and places his hands on his chest or belly, an acceleration towards the positive x-axis takes place because of the inward direction

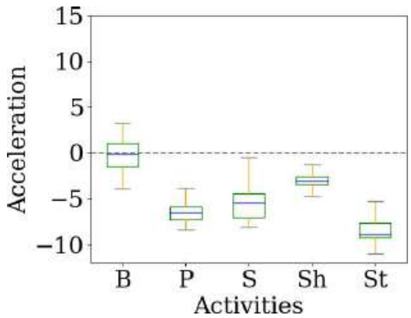
(a) Standing

(b) Bowing

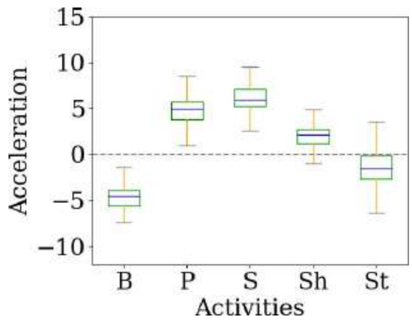
Figure 7.8: The placements of hands in (a) standing and (b) bowing positions



(a) x-axis



(b) y-axis



(c) z-axis

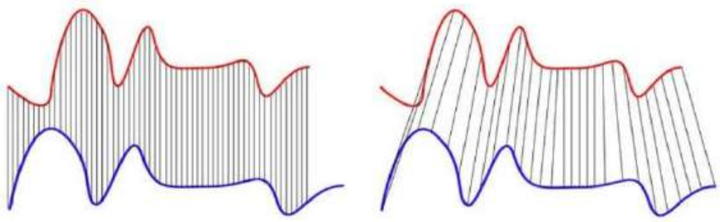
Figure 7.9: Comparison of the values of x, y, and z axis at different steady states (St = Standing, B = Bowing, Sh = Short-standing, P = Prostrating, and S = Sitting)

of the movement of our hands. Besides, the same inward direction also generates negative acceleration along the y-axis. On the contrary, while in the short-standing position, we keep our hands floating on both sides. Therefore, while going to the short-standing position, acceleration takes place slightly along the outward direction, and thus in short-standing, the x-axis value is always negative. Therefore, we can infer that the steady state before bowing is standing, if the x-axis value is greater than y-axis value, and short-standing otherwise. In this fashion, we can recognize the bowing and standing states with confidence, as per the semantic rules developed based on domain knowledge. The corresponding algorithm is presented in Algorithm 2.

Algorithm 2 Algorithm for recognizing steady states through semantic rules

Input: S Signal segments labeled as Steady or Transition by the state recognizers
Output: L new labels array

```
count ← 0
L ← []
for each segment s in S do
   if s is steady then
      if (s.x − s.y).mean < 0 and (s.z − s.y).mean < 0 and s.x.mean < 0 and s.z.mean < 0
      then
         L[count] ← Bowing
         ps ← previous Steady segment of s
         if (ps.x − ps.y).mean > 0 then
            L[count − 2] ← Standing
         else
            L[count − 2] ← ShortStanding
         end if
      else
         L[count] ← Unknown
      end if
   else
      L[count] ← Transition
   end if
   count = count + 1
end for
```



(a) Euclidiean match (b) DTW match

Figure 7.10: A comparison of Euclidean and DTW matching

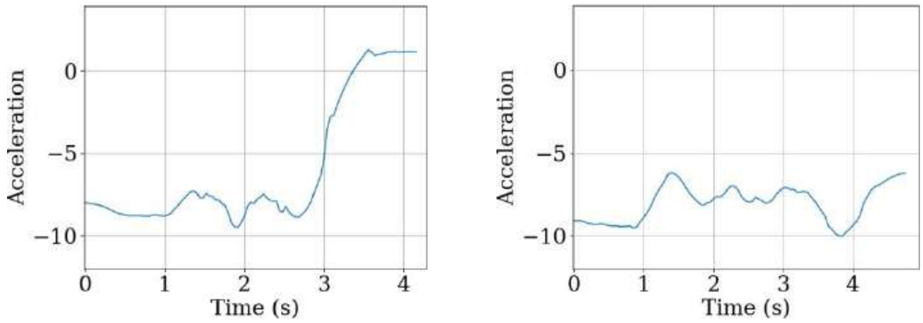
Classification of Transitions using Dynamic Time Warping (DTW)

In this stage of our proposed methodology, we aim to classify the transitions that occur between the steady states. This classification, in turn, helps us to infer the steady states. We use Dynamic Time Warping (DTW) [170], more specifically a variant of DTW called FastDTW [174], for classifying the transitions.

Dynamic time warping (DTW) [170] is a widely used and robust template matching algorithm for time series data. DTW seeks the optimal temporal alignment, which means a matching between time indexes of the two time series. The matching minimizes the Euclidean distance between the aligned series. Non-linear mapping is its primary strength. In contrast to Euclidean distance, which is extremely restrictive and matches point to point, DTW allows the two series to evenly match up even though the X-axes (i.e., time) are not necessarily in synchronization. Figure 7.10 presents the matching technique of DTW in contrast to the Euclidean matching.

A well-known application of DTW has been in automatic speech recognition, to cope with different speaking speeds [175–177]. Besides, it is used in partial shape matching applications [178]. However, we also find this algorithm being adopted in HAR research [26, 110, 111, 169], as this approach is beneficial for the analysis of real-world time series data. Besides, DTW is also robust against variation in speed or style in performing transitions. For instance, similarities in walking could be detected using DTW, even if one person walks faster than the other, or if there were accelerations and decelerations during the course of an observation. Considering these aspects, we utilize DTW to classify all the transitional activities in Salat as given in Table 1.1.

For the classification of transitional activities, we first create a template database with templates for each of the transitions. We slice out the transitions from each individual's data in the training sample and save the representative ones as templates. Along with the transitions mentioned in Table 1.1, three other transitions, namely bowing to short-standing to prostrating (B-Sh-P), prostrating to sitting to prostrating (P-S-P), and standing to short-standing (St-Sh) are taken into consideration. As for B-Sh-P and P-S-P, we include these two transitions as we find in our dataset that, many subjects do not spend much time in the short-standing phase in B-Sh-P and in the sitting phase in P-S-P. Therefore, the short-standing in B-Sh-P and the sitting in P-S-P



(a) Male (b) Female

Figure 7.11: Comparison of two representative templates of Sh-P from (a) Male and (b) Female template databases

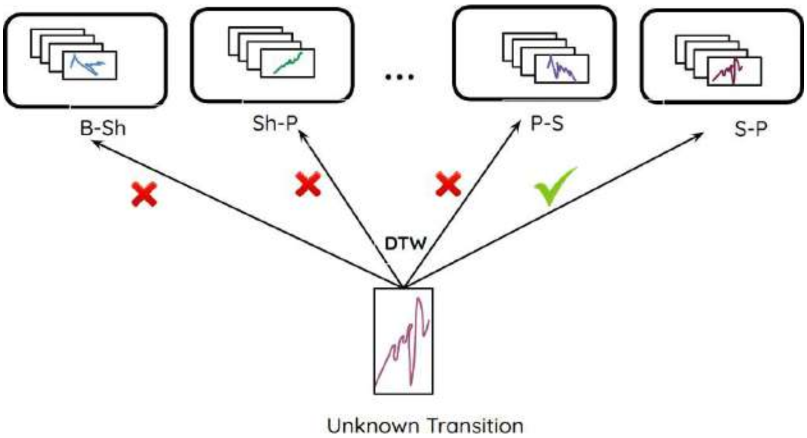


Figure 7.12: Classification of an unknown transition using DTW

can not be recognized individually as steady states for not having enough time spent. Here, the short-standing and sitting are considered as a part of the transition from the previous steady state to the next steady state. Regarding St-Sh, this transition is performed by some of our participants while performing the extra Takbeer. Hence, these patterns are stored in our databases too.

Thus, we get a total of 10 sets of templates pertinent to all 10 types of transitions. These sets are maintained for males and females separately as male and female prayer patterns vary significantly in our dataset. Figure 7.11 depicts an example case of the variation in the prayer patterns. For an unknown transition, we utilize the DTW scores for all the templates in the template sets. We take the average of the DTW distances from our unknown transition to all templates of a template set. In this way, we find the average distance of our unknown transition from all template sets. The set with the minimum distance indicates that our unknown or test transition belongs to this set.

As the male and female prayer patterns vary significantly in our dataset, therefore, we maintain separate template databases for males and females. One example of the difference in the transitions between male and female is depicted in Figure 7.11.

To elaborate our approach further, let X be an array of sample accelerometer values labeled as a transition by the state recognition stage and we want to classify X using DTW. To do so, first, we need to find out the distance of each template set from X. The formula to measure the distance of X to the k-th transition set T_k is as follows.

$$d(X, T_k) = \frac{1}{N} \sum_{i=1}^{N} DTW\,(X, T_i^k) \tag{7.2}$$

Using this equation, we will get the distances from X to all template sets. Now, the set with which X will give the minimum distance will be the class of X. Accordingly, if there are s template sets, the distance of X to the X's class will be as follows.

$$d = \min_{k=1..s} d(X, T_k) \tag{7.3}$$

In this manner, we can classify the transitions using DTW. Besides, from a transition, we can infer the next and previous steady states. Another important thing is, while classifying the transitions, we utilize the knowledge of bowing and standing recognized already using the semantic rules. This happens because, when we know that a steady state is bowing, we do not match the next transition with all ten transitions but only the ones which start with bowing, i.e., B-Sh and B-Sh-P. In this fashion, we can leverage the template-matching task through a delicate blending between semantic rules and DTW.

We perform this classification stage with (WT) and without (WOT) incorporating the knowledge of the recognition done using semantic rules. Here, WT means we incorporate the knowledge of the recognition of bowing and standing states while performing the DTW-based classification. Through applying the semantic rules in this way, we can detect the bowing and standing stages. Such detection can eventually facilitate recognition of the next transitions. This happens as, for example, if we know that a steady state is bowing, then the next transition should start with bowing, i.e., the next transition should be either B-Sh or B-Sh-P. Therefore, we can only match with these two types of templates to classify the unknown transition after bowing. On the other hand, in WOT, we do not incorporate the knowledge about the steady states. Therefore, in the case of WOT, to classify the transition after bowing, we have to match with all the possible ten types of template sets.

It is worth mentioning that the extra activities or Null activities cause some noise in the data. The noises are also segmented as transitions and if we try to match these noises with existing transitions, we will get very high DTW distances. Therefore, to distinguish the noises, we set a threshold value. While computing the distance of an unknown transition, if we get the distance beyond the threshold, then we predict that transition to be a noise resulting from a Null activity. In this way, we recognize the noises introduced by the extra or Null activities.

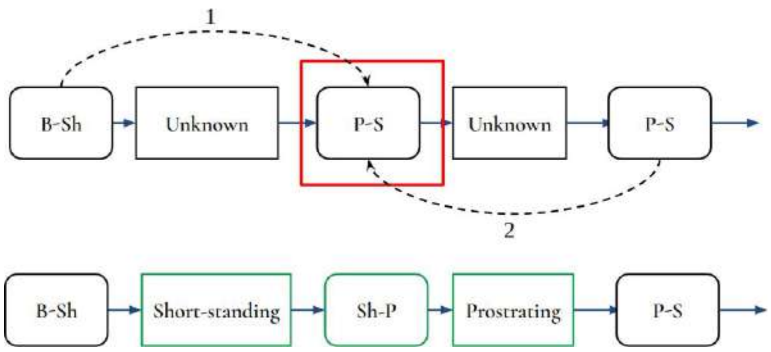


Figure 7.13: An example of fixing misclassified transition using neighboring transitions

Post-processing

In this study, our focus is to classify the steady states of Salat along with Takbeer. However, after the classification of the transitions is done using DTW, each of the predicted transitions provides us information about its previous and next steady states. In other words, each steady state has two transitions associated with it - one leading to that steady state from the previous steady state and another is the immediate next transition starting from that steady state. Besides, some steady states such as bowing and standing have already been classified using semantic rules. Therefore, they provide us with some extra information about the context upfront. Accordingly, by giving a second pass over the results obtained up to this stage, and combining these available results about the steady states, some misclassifications can be detected. The detected misclassifications can substantially be fixed using domain knowledge. The postprocessing stage performs this task and attempts to correctly predict each steady state through making necessary corrections. This type of postprocessing is also found in the existing HAR literature [179, 180]. We describe our postprocessing techniques below in detail.

Postprocessing based on Neighboring Transitions

We iterate through the predicted transitions by DTW and detect and fix the inconsistent transitions. Here, by being consistent, we mean if one transition is leading to a specific steady state, the next transition should start from that particular steady state. For example, if a person has gone through a transition A-B, that means he was in the steady state A, and from A, he has gone to the steady state B. Therefore, the next transition should be B-X, where X is any other steady state. We detect such inconsistent transitions and fix them based on their neighboring transitions.

For example, in the example presented in 7.13, the transition after bowing is found to be B-Sh, and the next transition is found to be P-S. Therefore, we call the next transition inconsistent with its predecessor. In this case, we fix the inconsistent transition based on its previous and next transitions. In our example, P-S is the inconsistent one. Its previous transition is B-Sh, and the next transition is P-S. Then both of the neighboring transitions indicate that the middle one should be Sh-P.

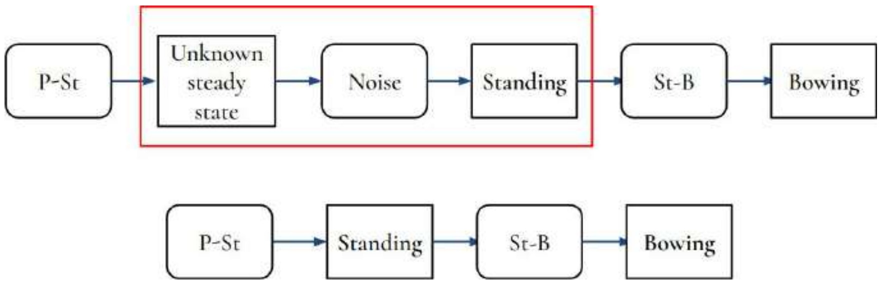


Figure 7.14: An example of ignoring extra movements

Postprocessing based on Duration of the Transitions

In our experimentation, we find B-Sh-P to be confused with B-Sh or Sh-P and P-S-P with P-S or S-P. However, both B-Sh-P and P-S-P consist of two transitions, and therefore their duration should be substantially higher than the individual transitions. We calculate the mean duration taken in each of these transitions and use this knowledge to resolve these confusions in the postprocessing stage. For example, if at some point, we find a transition P-S with unusual duration, i.e., duration much higher than expected then we assume that this is a P-S-P and correct accordingly.

Detection and Elimination of Extra Movements

The extra movements, i.e., Null activities, result in some extra transitions, which usually span a very short time duration and can be considered as noises. As mentioned earlier, if for any transition we find all the template sets having a distance greater than the threshold value, we will consider the transition as a noise. Therefore, if we find some steady states with noises in between, then those noises will be ignored and we will merge all these steady states as a single steady state. Figure 7.14 presents such a scenario of detecting and eliminating extra movements.

The flowchart of postprocessing is presented in Figure 7.15. Here, we first remove the Null activities or extra movements. Then, if we find a transition to be inconsistent with its previous one, then we fix it based on its neighboring transitions as stated above. And when two consecutive transitions are consistent, we can easily infer the steady state between them. For example, if the two consecutive transitions are B-Sh and Sh-P, then the steady state between them should be short-standing. In this way, the unknown steady states can be inferred.

Therefore, we can say that by incorporating our domain knowledge, we can detect and correct some misclassifications. Such detections and corrections of the misclassifications help us to improve our classification accuracy to a great extent.

Validation Protocol and Evaluation Metrics

We use both k-fold cross-validation [181] and Leave-One-Subject-Out (LOSO) [8] as our validation protocols. Besides, for the evaluation metrics, there exist several metrics to measure the performance of activity recognition. Examples of the metrics include accuracy, precision, recall, F-measure or F1-score, etc. These metrics are widely used in the evaluation of HAR models [8]. Definitions of the evaluation metrics are given below.

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \tag{7.4}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{7.5}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{7.6}$$

$$\text{F1-score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{7.7}$$

Here, TP is True Positive, TN is True Negative, FP is False Positive, and FN is False Negative. We use accuracy, precision, recall, and F1-score to analyze the performance of our methodology in recognizing each of the activities separately.

In our study, by accuracy and other metrics, we refer to the accuracy or respective metrics pertinent to classifying the activities through any of the above-mentioned methodologies. Thus, the accuracy and metrics correspond to how accurately the activities performed in Salat are recognized by the proposed methodology. Here, it is worth mentioning that we do not assess the correctness of Salat through these metrics, and therefore, accuracy does not refer to the accuracy of Salat in any way.

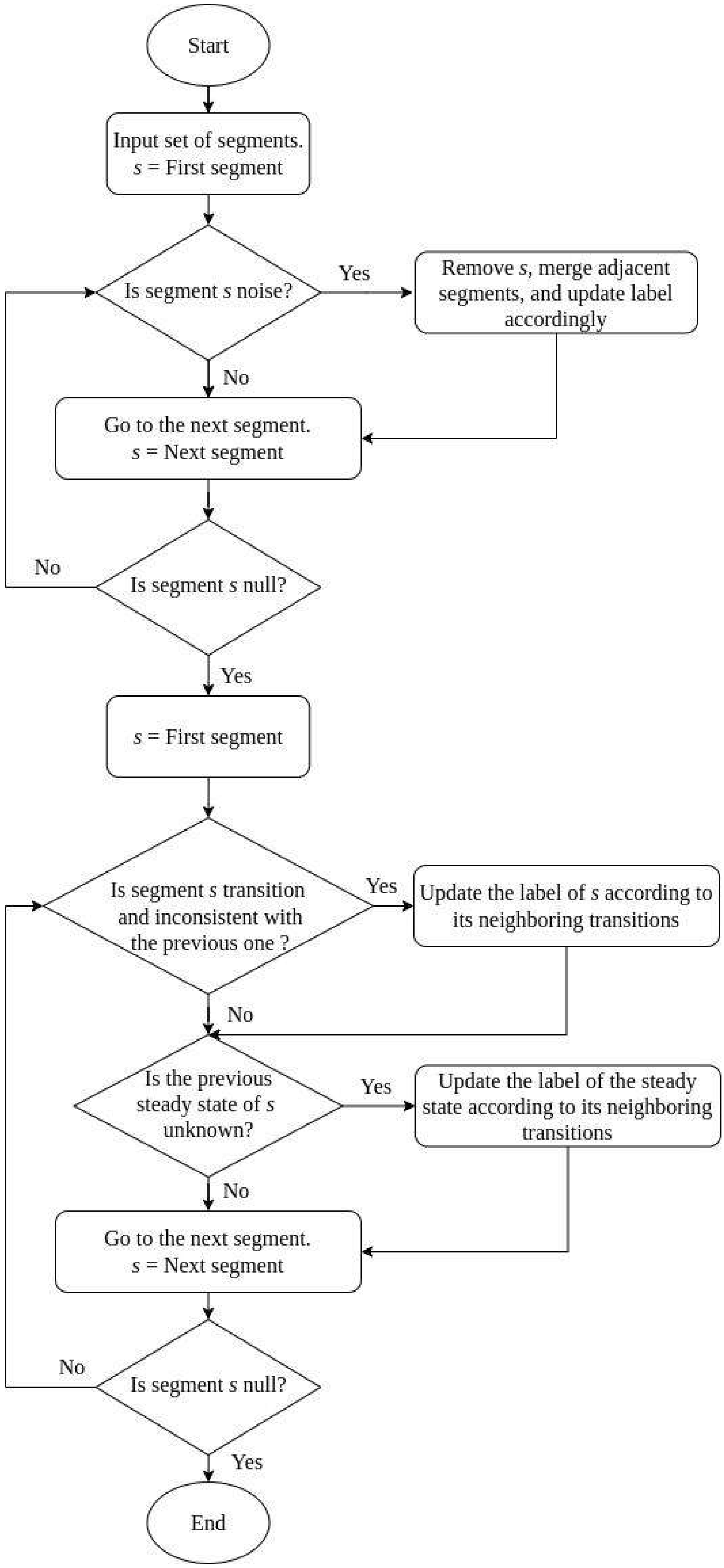


Figure 7.15: Flowchart for postprocessing

Chapter 8

Findings

In this chapter, we present the experimental evaluation of the baseline methodology as well as our proposed methodology presented in Chapter 7, for activity recognition in Salat. We perform the experimental evaluation using our prepared dataset. Therefore, first, we describe the details of the dataset, and then we present the performance at each stage of our proposed methodology. We also present a comparison between the performance of the baseline methodology and that of our proposed methodology.

Dataset Details

As mentioned earlier, we collect data from 30 subjects and prepare a dataset for the purpose of this research study. Our dataset contains 3, 50, 762 samples in total. Figure 8.1 summarizes the sample distribution of the activities in the dataset. Here, Null activity refers to the extra activities performed in Salat that do not nullify prayer.

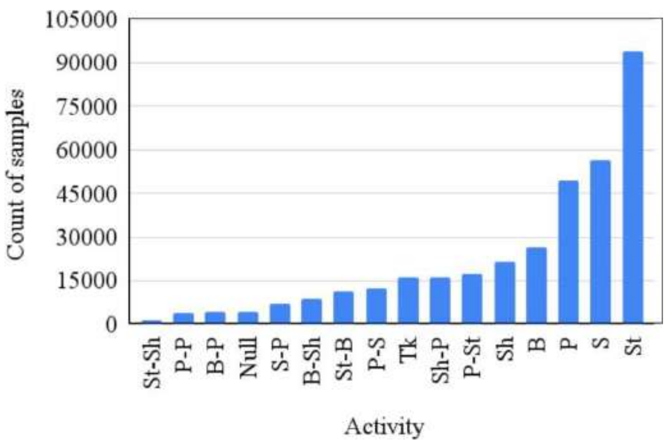


Figure 8.1: Distribution of activities in our dataset

As mentioned in the study of [182], in the context of human activity recognition, the diversity

Table 8.1: Statistics of the demographic factors of the subjects in our study

	Age (years)	Height (cm)	Weight (kg)
Range	15-67	149-175	35-91
Mean	38.3	160	59.5
SD	14.1	5.8	8.5

of the subjects enrolled includes the following four factors: (1) gender, (2) age, (3) height, and (4) weight. Accordingly, to cover gender diversity, we collect data from 13 female and 17 male subjects. Besides, to present the diversity over the other three factors, we present the statistics of age, height, and weight of our subjects are listed in Table 8.1. We expect that the diversity in each of these four demographic factors covers a wide range of populations.

Results and Findings

In this section, we will state our findings in detail obtained from our experimentation. Here, first, we present findings from the baseline methodology and then from our proposed methodology.

Baseline Analyses using Machine Learning Classifiers

We employ four classical machine learning classifiers as well as a deep learning classifier for our classification tasks. The four classical machine learning classifiers are Random forest [152], J48 decision tree [85], Naive Bayes [153], and Logistic regression [154]. Here, in all cases, we achieve very good k-fold cross-validation accuracy ($k = 5$) for Approach-1 and Approach-2 as mentioned in Section 7.4.1 in both single classifier and hierarchical fashion. However, when we perform LOSO for testing i.e., the model is trained with all but excluding one subject's data, and that particular excluded subject's data is used as the test data, accuracy varies substantially from person to person. For some people, the LOSO accuracy is found to be high, and their activities are recognized correctly. On the other hand, for some people, the LOSO accuracy gets below 70% and the classifiers become much confused between individual activities in Salat. Table 8.2 quantitatively presents details of these findings.

We use a sliding window of 1 sec for classical machine learning classifiers with 50% overlap. On the other hand, for GILE [116], the sliding window length is 1.28 sec with 50% overlap. For both types of classifiers, these values give us the best performances tuned up through our experiments.

From the results shown in Table 8.2, we can see that, for both approaches, accuracy increases, in general, in a hierarchical manner as each classifier can specialize in its own domain. The increase in accuracy also happens for LOSO which is our main focus. Therefore, from now on, we will only consider the results obtained in the hierarchical fashion. Besides, accuracy is much higher in Approach-1. This is expected, as in Approach-2, we classify the transitions and

Table 8.2: Accuracy (%) of machine learning classifiers in Approach-1 and Approach-2

Classifier	Single classifier				Hierarchical			
	Cross-validation	LOSO			Cross-validation	LOSO		
		Avg	Max	Min		Avg	Max	Min
RF	94.05	80.63	92.54	61.29	95.65	81.93	96.80	62.68
LR	91.65	82.04	93.70	65.09	92.88	83.10	95.01	66.78
NB	87.54	78.12	91.20	52.08	88.11	78.90	93.44	54.84
J48	91.63	76.35	89.39	58.09	92.63	76.88	90.34	58.78
GILE	96.1	83.68	93.75	69.14	93.52	83.96	95.35	69.55

(a) Approach-1

Classifier	Single classifier				Hierarchical			
	Cross-validation	LOSO			Cross-validation	LOSO		
		Avg	Max	Min		Avg	Max	Min
RF	91.22	74.99	92.91	54.32	92.49	78.93	94.80	55.38
LR	85.12	75.77	90.01	65.63	87.48	77.5	90.31	66.78
NB	75.24	66.12	43.20	30.08	76.32	43.90	68.44	31.14
J48	86.54	73.35	58.39	58.09	87.32	72.22	85.13	51.38
GILE	90.85	73.18	90.77	50.14	88.01	77.96	92.32	63.51

(b) Approach-2

steady states altogether, and thus, there arises a large number of classes making the recognition task more challenging. Among the two approaches, the best LOSO performance is found by the deep learning model GILE in the hierarchical fashion. The reason behind this finding is the fact that the model is specialized for cross-person HAR, i.e., for extracting latent characteristics of activity independent of any person. This, in turn, increases the cross-person generalization capability of the model. Besides, it is worth mentioning that for both the approaches, RF obtains the highest cross-validation accuracy among all the classical machine learning classifiers under consideration. Nonetheless, LR achieves the highest LOSO accuracy in most of the cases among all the classical machine learning classifiers.

Another important thing is that, in our dataset, the prayer pattern of male and female subjects vary substantially, as they belong to a specific school of thought (Hanafi [183]) and the school of thought prescribes to do so. Therefore, we separate male and female data, maintain a separate database for each of them and evaluate them independently. We also attempt training and testing with the combined dataset. Here, we find that the accuracy drops in many of the cases. Figure shows a comparison in this regard.

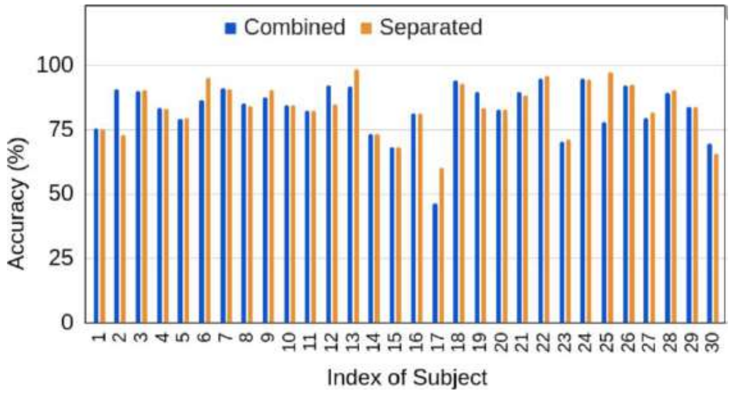
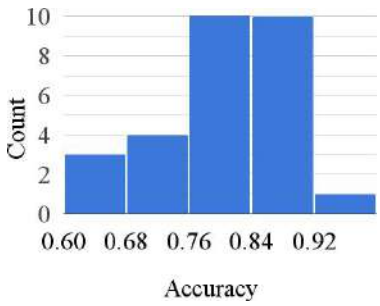
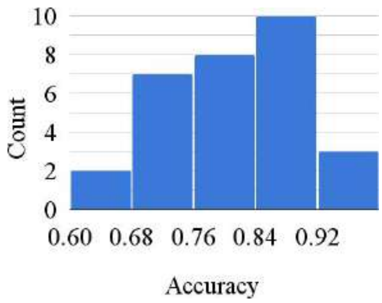


Figure 8.2: LOSO accuracy of GILE when datasets from male and female subjects are considered in combined and separated manners for Approach-1 in the hierarchical fashion



(a) LR



(b) GILE

Figure 8.3: Histogram of LOSO accuracies obtained by (a) LR and (b) GILE for Approach-1 in the hierarchical fashion

Limitations of Machine Learning Classifiers

It is clear from Table 8.2, that the LOSO accuracies are not uniform across the subjects and substantially vary among different subjects for all the classifiers. To analyze the variation in depth, we present the histogram of the LOSO accuracies found by LR and GILE for Approach-1 in the hierarchical fashion in Figure 8.3.

It is evident from Figure 8.3 that even if we achieve high cross-validation accuracy, the LOSO accuracy is not satisfactory for a considerable number of subjects. However, the confusion matrices can give us another insight into the predictions of the individual activities along with the types of errors that are being made by the classifiers. Therefore, we present confusion matrices obtained with LR and GILE for individual steady activities along with all transitions combined in a single class in Approach-1 in the hierarchical fashion in Figure 8.4.

From the confusion matrices in Figure 8.4, we find that bowing and standing activities are less confused by the classifiers. However, for the other activities such as prostrating, sitting, and short-standing, we do not notice any deterministic error pattern. For example, for some people, prostrating is sometimes confused with short-standing and sometimes with sitting. On the other

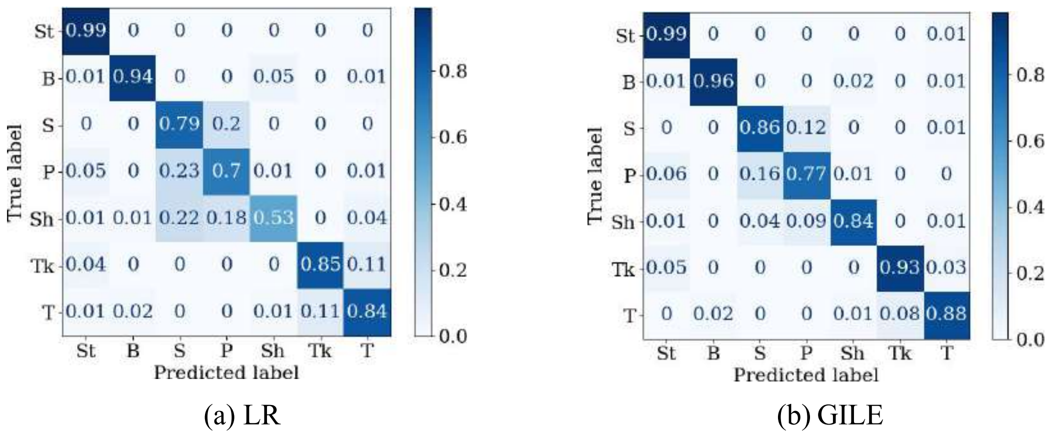


Figure 8.4: Confusion matrices obtained by (a) LR and (b) GILE with Approach-1 in the hierarchical fashion (St = Standing, B = Bowing, S = Sitting, P = Prostrating, Sh = Short-standing, Tk = Takbeer, and T = Transition)

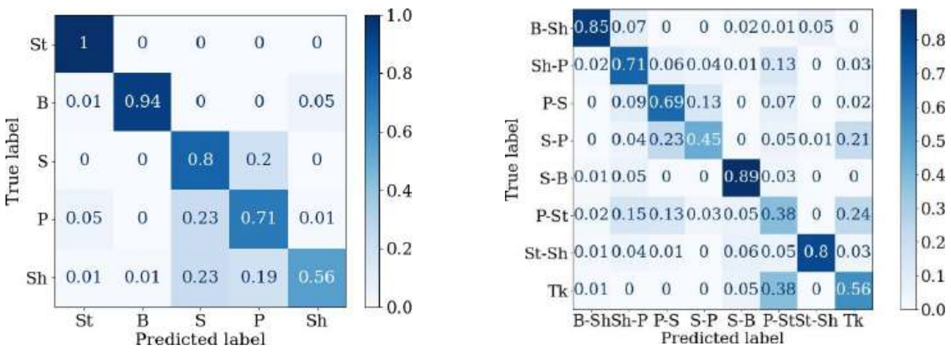
Table 8.3: Prediction of activities performed in one Rakat by a subject (P4) by GILE

Actual	Standing	Bowing	Short-standing	Prostrating	Sitting	Prostrating
Predicted	Standing	Bowing	Prostrating	Prostrating	Sitting	Prostrating
Remark	✓	✓	X	✓	✓	✓

hand, short-standing is mostly confused with prostrating or sitting and rarely with bowing. Let us take an example to elaborate on the problem. For one subject (P4), we find the expected activities and the activities of one Rakat predicted by GILE as shown in Table 8.3.

We can see from Table 8.3 that the short-standing is misclassified as prostrating. Therefore, the model will predict that this person has prostrated thrice, whereas, in reality, he has prostrated twice. Here, we have no clue to detect and fix this misclassification, as sometimes after the bowing phase, some people go to the short-standing phase and then immediately go to the prostrating phase without delaying a bit in the short-standing position. This whole movement, i.e., bowing to short-standing to prostrating is predicted as a single transition to the classifier, as there is barely any delay during those activities and barely any pause between those activities. A similar thing can happen while going prostration to another prostration, as some people do not sit and spend a bit of time in between consecutive prostrations, and they immediately go for another prostration after the first one. For these reasons, it becomes extremely difficult to detect this misclassification through incorporating our domain knowledge.

On the other hand, we explore another alternative approach (Approach-2), so that we can get enough context about the steady states. From Table 8.2, we find that the overall accuracy degrades in the cause of Approach-2. The individual confusion matrices of the steady states and transitional states demonstrate that the poor classification accuracy of the transitional states substantially contributes to the degraded accuracy. Figure 8.5 shows the confusion matrices of steady states and transitional states obtained by LR with Approach-2.



(a) Steady (b) Transition

Figure 8.5: Confusion matrices of (a) steady and (b) transitions obtained by LR with Approach-2 in the hierarchical fashion

Such degraded accuracy for short-lived activity recognition is not new in the literature. We find in the existing HAR literature that, the traditional features of short-lived samples are unstable and cannot describe the actions effectively [76, 168]. The transitions in Salat are also short-lived activities and therefore, the ML classifiers are unable to recognize this large set of transitions in Salat correctly. Therefore, as the accuracy of the recognition of transitions is poor, they fail to provide us with reliable context information, which we could have used to improve the accuracy of the steady state misclassifications. This scenario eventually leads us to attempt designing an improved methodology for better prediction of activities in Salat through better error handling.

Performance Analyses of the Proposed Methodology

To overcome the limitations of the machine learning classifiers discussed above, in this study, we propose a new methodology for activity recognition in Salat involving semantic rules and DTW. Below we present the performance analyses of this improved methodology.

Results of State Recognition

In our proposed methodology, we first recognize steady and transitional states. We have already mentioned earlier in Section 7.4.2, that this can be done using two approaches, and here, we compare the results obtained using both of these approaches. Here, we segment the signal using a sliding window of length 1.2 seconds with 50% overlap and a moving average filter with a window size of 10. Similar to the earlier case, we use four classical machine learning classifiers i.e., RF, LR, NB, and J48 over the segmented data and find RF as the best-performing one.

In the Signal Magnitude Area (SMA) based approach, we use a high-pass Butterworth filter of order 3 with a cutoff frequency of 1Hz following the convention [172] of the literature. We

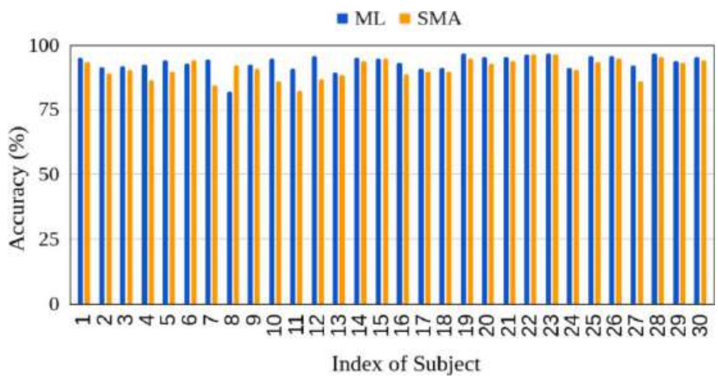


Figure 8.6: Performance comparison of two state recognition methods

adopt these parameters to obtain the linear acceleration component from the acceleration signals. Here, we set the threshold value to 1, i.e., if the SMA value of a segment is less than 1, then this segment is considered to be steady, otherwise, we consider the segment transitional. The two approaches give us almost similar results, which are depicted in Figure 8.6.

After predicting each segment, we merge contiguous blocks of similar states. Though we see in Figure 8.6 that the classification into steady and transition accuracies fluctuate around 90%, after merging the adjacent similar segments, the recognition accuracy goes around 99.5% for both. This improvement in accuracy indicates the feasibility of employing the classification of segments into steady and transition in both the proposed approaches. Another important aspect is, we find that this process is not very sensitive to the choices of parameters for segmentation. For example, we explore variations in these parameters, such as window size from 0.5s to 1.5s and moving average filter window length from 5 to 30, and find almost no change in the accuracy.

Results of Semantic Rule-based Classification for Steady State Recognition

Using the semantic rules, devised from the domain knowledge on the steady states, as already described in Section 7.4.2, we find that the bowing and standing activities can be recognized with 100% accuracy. The confusion matrix we obtain after applying the semantic rules is shown in Table 8.4. The table demonstrates achieving perfect accuracy while considering the bowing and standing states in isolation and the rest of the three states (short-standing, prostrating, and sitting) in combination.

Results of DTW-based Detection for Transitional State Recognition

Upon classifying the steady states using semantic rules, we apply DTW to recognize the transitions in Salat. We conduct this stage with (WT) and without (WOT) incorporating the

Table 8.4: Confusion matrix after applying semantic rules

	Standing	**Bowing**	**Short-standing + Prostrating + Sitting**
Standing	1	0	0
Bowing	0	1	0
Short-standing + Prostrating + Sitting	0	0	1

Table 8.5: Overall accuracy of classifying transitions using DTW with (WT) and without (WOT) applying the knowledge obtained using the semantic rules

Gender	**Accuracy (%)**	
	WT	**WOT**
Male	82.33	74.54
Female	84.25	82.14

knowledge of the recognition done using semantic rules, i.e., the knowledge of distinguishing bowing and standing states. Here, we find that the former approach, i.e., with the incorporation of the results of applying semantic rules, performs better. Table 8.5 presents the comparative results over WT and WOT approaches.

The reason for WT performing better is that it narrows down the search space and reduces the number of competing templates to consider for each unknown transition. To elaborate a bit more, after applying the semantic rules, we can recognize the bowing and standing steps with perfect accuracy. Incorporating this knowledge in our DTW stage means, for example, if we know that a steady state is bowing, then the next transition should be either B-Sh or B-Sh-P. Therefore, we can only match with these two types of templates to classify the unknown transition. However, if we have no knowledge about the steady states, then to classify the transition after bowing, we have to match with all possible ten types of template sets. This increases the chance of misclassification, as more possible transitions are there.

As WT performs better, we adopt this in our methodology. Another important point is, as we have already shown in Section 7.4.2, that the transition patterns vary substantially between men and women. Even though, we explore combining both male and female patterns together and perform classification using DTW. Here, due to the differences in the templates for males and females, we find the accuracy degrading significantly and dropping even below 50%. Therefore, we maintain separate template databases for males and females and carry on this classification separately. Table 8.6 presents the confusion matrices in this regard.

The confusion matrices show us that the majority of the transitions are classified accurately. Here, Null means the extra activities performed during Salat that are to be ignored. We set the DTW distance threshold to 500 for Null activities. This implies that, if the lowest DTW distance from an unknown transition to the template sets exceeds 500, then this is a Null activity. From the confusion matrices, we find that all the Null activities of both Male and Female datasets

Table 8.6: Confusion matrices of the classification of transitions for (a) male and (b) female datasets

	St-B	B-Sh	Sh-P	P-S	S-P	P-S-P	P-St	St-Sh	B-Sh-P	Tk	Null
St-B	79	0	0	0	0	1	0	0	0	0	0
B-Sh	0	60	0	0	1	1	0	0	1	0	0
Sh-P	0	0	38	3	8	0	3	0	1	1	0
P-S	0	0	0	69	9	1	3	0	0	0	0
S-P	0	0	0	0	52	0	0	1	0	0	0
P-S-P	0	0	1	6	1	8	1	0	0	0	0
P-St	0	0	0	1	1	0	54	0	0	0	0
St-Sh	0	0	0	0	0	0	0	4	0	0	0
B-Sh-P	0	2	0	0	0	0	0	0	12	0	0
Tk	0	0	0	0	0	1	16	1	0	37	0
Null	0	0	0	0	0	0	0	0	0	0	4

(a) Male

	St-B	B-Sh	Sh-P	P-S	S-P	P-S-P	P-St	St-Sh	B-Sh-P	Tk	Null
St-B	60	0	0	0	0	0	0	9	0	0	0
B-Sh	0	45	0	0	0	0	0	0	0	0	0
Sh-P	0	0	37	0	2	0	0	0	0	1	0
P-S	0	0	0	62	1	1	0	0	0	0	0
S-P	0	0	0	4	41	0	0	0	0	0	0
P-S-P	0	0	0	2	5	5	0	0	0	0	0
P-St	0	0	0	1	0	0	45	0	0	0	0
St-Sh	0	0	0	0	0	0	0	8	0	0	0
B-Sh-P	0	10	0	0	0	0	0	0	4	0	0
Tk	0	0	0	0	0	1	13	0	0	34	0
Null	0	0	0	0	0	0	0	0	0	0	2

(b) Female

got classified correctly. However, we find Tk is mostly confused with P-St for both male and female datasets. However, P-St is not confused with Tk, which eases fixing this confusion in the postprocessing stage. Similarly, we see P-S-P is sometimes confused with S-P or P-S. However, as mentioned in Section 7.4.2, such confusion can be fixed using the duration of the transitions.

Final Results after Postprocessing

From the confusion matrices in Table 8.6, we see that most of the instances lie on the diagonal except for a few cases showing potential misclassifications. These misclassifications can be fixed by considering the previous transitions, steady states, and durations as stated in Section 7.4.2. For example, if we find a prediction of P-St at some point whereas the immediate previous transition does not involve sitting, then we can assume that this P-St should be Tk. On the other hand, in

Table 8.7: Final accuracy, precision, recall and F1-Score of each activity

Activity	Precision	Recall	F1-score	Accuracy
Takbeer	0.95	1	0.97	0.95
Standing	1	1	1	1
Bowing	1	1	1	1
Short-standing	1	1	1	1
Prostrating	1	0.99	0.99	0.99
Sitting	0.96	1	0.98	0.96

terms of duration, there is a significant difference between B-Sh-P to B-Sh. As B-Sh-P means a person going from bowing to short-standing to prostrating, this whole transition generally takes more time (Mean = 5.1 sec, SD = 0.17 sec) than B-Sh, i.e., bowing to short-standing (Mean = 2.2 sec, SD = 0.25 sec). Similarly, we can differentiate P-S-P (Mean = 6.1 sec, SD = 0.21 sec) from P-S (Mean = 1.9 sec, SD = 0.11 sec) and S-P (Mean = 1.7 sec, SD = 0.15 sec) by comparing the duration of a transition. Thus, by incorporating this domain knowledge, we can detect misclassifications as well as resolve confusions resulting from any misclassification. This, in turn, improves the predictions of the previous stages and yields higher prediction accuracy. The final precision, recall, and F1-score of each activity of Salat after applying such post-processing is given in Table 8.7.

After the postprocessing, out of 728 activities, only seven activities are misclassified. Here, two sittings are classified as prostrating. Besides, five Takbeer activities are missed as they are performed right after going up from sitting to standing without any pause. Due to the absence of pause, the sitting-to-standing and Takbeer are considered one activity and predicted as sitting-to-standing. Thus, the overall final accuracy obtained by our proposed methodology becomes 99.03% with 100% Rakat count accuracy. However, one very important observation here is that, for the people who are slow and steady in performing the activities of Salat, the predictions are generally accurate and require little postprocessing.

Chapter 9

Discussion

This study, for the first time in the literature, establishes a proof that recognizing activities in Salat using a smartwatch is not only possible but also can be done with enhanced accuracy than other existing methods. By adopting a completely novel methodology, we can achieve near-perfect accuracies in recognizing all steps of Salat. In this chapter, we discuss some important aspects of our study, such as its acceptability to users, methodological advancement, scaled-up experimentation, etc. Before presenting these discussions, we first briefly elaborate on how we answer our research questions set earlier in this study.

Outcomes of the Exploration of Our Research Questions

In this section, we will shed light on the outcomes of the exploration of our research questions. In this regard, we focus on the three research questions already set in Section 5.1.

Mistakes in Salat - Prevalence and Frequency among People (RQ1)

To find the answer to RQ1, i.e., to assess the frequency or recurrence of various types of mistakes during Salat, we conduct an online survey. We ask the participants about five common mistakes in Salat and want to know how frequently they make these mistakes in their prayers.

Through a mixed-method analysis over the responses of the participants, we find that more than one-third of our participants make at least one mistake frequently. Thus, we can safely assume that mistakes in Salat are common. However, the most common mistake reported by our participants is forgetting the count of Rakat. Detailed on survey conduction and findings are presented in Section 6. These findings answer our first research question, i.e., RQ1.

Technological Assistance in Salat - Requirement and Acceptability (RQ2)

To find the answer to RQ2, in our survey, we ask the participants regarding their requirements and willingness to explore technological assistance to improve their Salat. The participants could also express the reason behind their answers qualitatively.

The mixed-method analysis over the survey responses reveals that the majority of our participants (above 70%), irrespective of differences in demographic factors, regularity in prayer, etc., are willing to explore technological assistance to improve their Salat. Here, an important finding is that the participants demand convenient technological devices. Details of these findings are presented in Section 6. These findings answer our second research question, i.e., RQ2.

Leveraging a Convenient Device for HAR in Salat With Improved Performance (RQ3)

To answer RQ3, i.e., to check the feasibility of leveraging a convenient device for HAR in Salat, in our study, we explore recognizing activities in Salat using a smartwatch. In this regard, we explore the conventional methodology using machine learning classifiers as the baseline methodology. Furthermore, we propose an improved methodology using semantic rules and DTW.

Rigorous experimentation reveals that our proposed methodology outperforms the performance of the baseline methodology as well as that of the earlier studies. We achieve a near-perfect accuracy (99.3%), which establishes the proof that activity recognition in Salat using a smartwatch is not only feasible but can also be done with improved accuracy. The findings of our exploration and experimentation in this regard are presented in detail in Section 7 and 8.3. These findings answer our third (or the last) research question, i.e., RQ3.

Acceptability of Technological Assistance in Salat by Real Users

Before approaching the gaps in the literature regarding activity recognition in Salat, we explore the real users to assess the necessity and acceptability of technological assistance among them. To the best of our knowledge, no earlier HAR studies focusing on Salat carry on any such exploratory study in this direction. However, in the literature, we find some other studies [184,185] doing this practice of assessing the acceptability of their HAR solutions which focus on different domains other than Salat. The study in [184], the proposed HAR system supporting elderly and diseased people. To assess the potential acceptability of the HAR system proposed in this study, a survey was conducted to assess their general technical affinity so that the general acceptability of the system can be envisaged. Besides, the research objective of [185] is to explore the acceptability of an in-home ambient sensor for activity recognition and assessment of post-stroke cases. The study involved twenty individuals with chronic stroke and conducted semi-structured interviews with each study participant for a detailed analysis of the acceptability of their solution.

On the contrary, our main objective behind conducting a survey is to assess the frequency of mistakes in Salat among people and whether they are willing to explore any technological assistance to improve their prayers. The survey helps us achieve deeper insight into these aspects. We perform both quantitative and qualitative analyses of the survey results to get a comprehensive view of the collected responses.

From the demography of the participants of our survey, presented in Table 6.1, it is clear that there is a diversity among the participants in terms of age, gender, country, occupation, etc. Though the majority of our participants are from Bangladesh, we achieve responses from 15 different countries. Even if we consider religious demography, we find that the participants vary greatly in their regularity in prayer, mistakes, attitude towards technological assistance in religious activities, etc. Thus, we can claim that this survey does not represent any specific community, but, rather represents the Muslim population in general.

One of our primary findings from the survey is that mistakes in Salat happen in reality. In fact, the survey uncovered that mistakes are pretty common, as we find one-third of our participants report that they make at least one mistake on a regular basis. We do not find any significant relationship between the frequency of making mistakes and with the regularity in prayer or other demographic factors. Besides, irrespective of the regularity in their prayers, we find most of the participants expressing their eagerness to improve the quality and quantity of their prayers. However, making frequent mistakes and willingness to improve prayer do not present sufficient motivation for designing technology for this purpose, if people are not ready to accept any technological assistance for their religious activities. As religious worships are between people and the Almighty, it is very important to know whether they would allow any technology to assist them in their worship. Therefore, we ask them whether they would welcome their devices such as smartphones, smartwatches, etc., to help them in improving their prayers. In response, the majority of the participants express their willingness in availing of technological assistance. In fact, many of them express their excitement over this idea and praise it highly.

We also assess the degree to which they are willing to pray while wearing a wearable and find the majority responding in the affirmative. They prefer smartwatches over smartphones primarily because of the convenience the smartwatches offer and their proximity. Along with statistical analyses, we perform thematic analyses on the qualitative answers, which provide us with deeper insights into the thoughts, concerns, etc., of the participants regarding such technological assistance. Thus, digging into the real world through this survey, we establish the requirement as well as assess the acceptability of the technological assistance in Salat.

A New Dataset for Salat Activity Recognition

One very important contribution of our study is the preparation of a smartwatch dataset consisting of Salat activities. To date, there is no such dataset present in the literature. Our dataset contains a large number of samples (3,50,762) collected from 30 subjects including both men and women.

It is of utmost importance for any HAR dataset to have a substantial number of subjects to obtain reliable results as stated in [186]. Our dataset includes data from 30 subjects that appear to be sufficient enough, as other benchmark datasets cover the number of subjects equal to us [158, 159] or less than that of ours [182]. Furthermore, ideally, a dataset should reflect the variability of real-world activities, and be flexible enough to emulate different experimental setups [120]. In a similar way, our dataset covers a diverse demography of subjects, as shown in Table 7.1, and accumulated real-world activities in different experimental setups.

In our collected dataset, the prayer patterns of men and women differ, as they all belong to a particular school of thought (Hanafi) and the school of thought prescribes so [183]. However, the postures of the subjects, even the same gender, vary substantially. The variation gets evident from the change in LOSO accuracy achieved with machine learning classifiers as shown in Section 8.2.1. Besides, while collecting data, some of our subjects make various types of mistakes in their prayers. For example, one elderly male subject (P1) make mistake in Rakat counting and prayed six Rakat instead of four. Another subject (P3) bows twice and prays five Rakat. While another young female subject (P20) makes three prostrations. We keep them in our dataset as they would help check the performance of a model in the presence of these mistakes.

The preparation of such a dataset takes much time and manual effort for both collecting raw data from the subjects and subsequent labeling. As our dataset contains a substantial number of samples from a considerable number of subjects [187], having diversity in demography, we believe that, the dataset will serve as a basis for testing different HAR approaches in the future. Moreover, Salat being a complex activity, our prepared Salat dataset can also be used for the evaluation of any complex activity recognition model.

Methodological Advancement for Activity Recognition in Salat

In this study, we propose a new methodology to recognize the activities in Salat. None of the existing studies for activity recognition in Salat works with smartwatch-collected data which entails a different set of challenges as discussed in [79]. The HAR studies for Salat, which adopt sensor-based approaches, all work with smartphones. Besides, none of their methodologies matches ours, as almost all of them use only a classical pattern recognition pipeline covering the steps of labeling, denoising, segmentation, training classical machine learning classifiers, and finally evaluating the model through cross-validation accuracy.

In our study, we first attempt the classical pattern recognition pipeline for Salat, however, with the smartwatch-collected dataset. Here, similar to the other studies, we find impressive cross-validation accuracy, however, the accuracy drops in the case of LOSO. To overcome the drop, we leverage the notion of DTW. Though one of the earlier studies on Salat [35] uses DTW, the purpose of the study is very straightforward. The study only attempts to detect prayer and non-prayer, i.e., whether the signal pattern is a prayer pattern or not. The study does not make any attempt to recognize individual activities or steps in Salat. It develops a pattern to represent one Rakah prayer and develops some thresholds to compare and predict whether a test pattern is a prayer or not. Moreover, in the case of other existing studies [7, 47, 67, 188, 189], the recognition of activities or methodologies under consideration are either not applicable for Salat [7] or are much more complex compared to ours one [6, 11].

Besides, though we find several HAR studies in the literature using mostly DTW [26, 110, 111], our study establishes that the mere adoption of DTW is not sufficient to develop a model for recognizing all steps of Salat. The model definitely needs some sort of post-processing mechanism to correct the misclassifications in the process of preparing the final output. This post-processing step is not found in any existing studies on Salat. To summarize, the process of state recognition, application of semantic rules, leveraging DTW, and post-processing - all these in combination is not seen in the literature yet. This pipeline is developed exclusively considering the nature of activities in Salat and its various steps, analyzing people's postures while performing these activities, and so on. Our study proves that all the steps of this pipeline complement each other to build a robust model for recognizing activities in Salat.

Recognition of a Complex Activity with Near-Perfect Accuracy

First of all, in the literature, we find that complex activities are not only less explored but also challenging to recognize [27, 66]. Salat, being a complex activity, is not an exception here, as recognizing each individual step in Salat is undoubtedly a challenging task. However, our proposed methodology achieves a near-perfect accuracy (99.3%) in recognizing individual steps or activities in Salat, which outperforms all the existing studies found to date to the best of our knowledge. Table 9.1 presents a comparison of the performances and other aspects of our model as well as related existing studies [2, 3, 33, 36, 39].

It is worth mentioning that the accuracies of the other existing studies mentioned in the table are taken from the respective studies as they were reported. All these studies adopt different settings for their experimentations with Salat. However, it is important to note that none of these existing studies reports LOSO accuracy, and therefore, the accuracies reported by these existing studies should be an overestimate - at least from the perspective of the LOSO scenario. Nonetheless, for a more comprehensive performance comparison, assessing all the methods in the same experimental setting could be explored in the future.

Table 9.1: Comparison over the performances and other aspects of our approach and other related studies

Author (Year)	Number of Subjects	Device (Placement)	Natural Usage	Applicable for all	Validation protocol	Accuracy (%)
Ghannam et al., [3] (2016)	-	Smartphone	X (Upper-back)	X	Separate training and testing data	91.0
Eskaf et al., [39] (2016)	10	Smartphone	✓ (Shirt's pocket)	X	Cross-validation	94.6
Obaid et al., [33] (2018)	20	Smartphone	✓ (Pocket)	X	Cross-validation	93.0
Ahmad et al., [2] (2019)	10	Smartphone	X (Upper-arm)	X	Cross-validation	97.5
Topu et al., [36] (2021)	8	Smartphone	✓ (Pocket)	X	Cross-validation	93.9
Our study	30	Smartwatch	✓ (Wrist)	✓	LOSO	99.3

Salat differs in its characteristics from other activities studied in the literature to a great extent, and our method is specifically devised for Salat keeping its specific characteristics in consideration. Here, first of all, our proposed approach of state recognition separates out the steady and transitional steps, and thus, narrows down our search space. Afterwards, we apply the notion of semantic rules, as the activities in Salat present specific semantics in their actions. The significance of applying semantic rules is proved by the accuracy obtained with and without applying these as presented in Table 8.5. It is worth mentioning that, the machine learning approaches demand extracting features from each window, and this incurs substantial computational cost. Instead, we use DTW, more specifically a variant of DTW, called fast DTW [174]. Thus, our proposed method reduces the computational complexity compared to existing machine learning approaches demanding feature extraction.

Robust Performance Analysis

With the classical machine learning approaches, we obtain a maximum of 95% cross-validation accuracy. With DL, this is even better (96.1%). However, we investigate that when we perform LOSO instead of cross-validation, we find an accuracy of less than 70% for some subjects. This means that, for some of the subjects, these models will be able to recognize activities in Salat correctly, while at the same time, there are high chances to fail for some other subjects. The cross-validation accuracies we achieve are unable to reveal these limitations of the classical machine learning approaches as well as DL-based approaches. However, the findings from LOSO motivate us to come up with an improved methodology to provide better and more uniform performance in a user-independent manner.

In the case of cross-validation, we shuffle the data and then divide it into a few folds. For each fold, the system is trained with all data, except for the data from that particular fold. The performance of the system is then tested on the fold that was kept out for training. Finally, performance is averaged across the different folds. Accuracy is typically higher in such validation, as activities performed by human beings have a strong subjective characteristic that is related to different factors such as age, gender, weight, height, etc., [190]. Therefore, the same activity performed by different persons might substantially vary, while a person tends to do the same activity almost in the same way every time. Accordingly, in the case of cross-validation, as the data of all the subjects are shuffled together, it is highly likely that the training folds already carry some data of the subjects whose data are present in the testing fold. Therefore, as the model has already learned those data, it can perform well on the unseen data of the same subjects.

However, this is not what happens in real-life deployment, as the accuracy of activity recognition can fluctuate when applied to data collected from new unseen participants, indicating a lack of generality across different persons [116]. Therefore, if we rely only on the cross-validation accuracy, we may miss out the latent weaknesses of our model. Nonetheless, our study proves a similar concern raised in the study [114], demonstrating that cross-validation accuracy, no matter how impressive, is not sufficient alone to understand the correct behavior of a HAR model in reality. The LOSO type of validation is highly important as it closely mimics the real scenarios. Therefore, we use LOSO as the validation protocol in addition to cross-validation to evaluate the classifiers' ability to recognize the activities of an unacquainted subject. This gives us a better understanding of the performance of our proposed methodology in the real world.

Scaled up Experimentation with Larger Number of Subjects

A noteworthy strength of the experimentation and demonstration of the validity of the results in our study is that our experimentation involves a considerable number of subjects covering diverse demography. Among the earlier HAR studies focusing on Salat, we find that the maximum number of subjects involved is only 20 [33]. Most of the existing studies collect data from fewer subjects, mostly 10 [2, 3, 39]. Some studies even work with fewer such as 8 subjects [36]. On the contrary, we experiment with a much larger number of subjects (30). Moreover, the diverse demography presented in Table 7.1, ensures that enough variation over the subjects has been captured The diversity and variation in turn facilitate demonstrating the validity of our results.

Fine-grained Recognition

Our model is capable of predicting each step of Salat except Taslim. Taslim marks the end of the prayer while sitting in the same position, turning only the head first to right and then to left saying a specific supplication [28, 29]. As this only includes the movement of the head, its impact on the smartwatch (placed on the wrist) is not significant enough. Therefore, the signal captured by the smartwatch does not exhibit substantial change during this step. This is why, this step cannot be recognized. All other steps, irrespective of whether it is steady or transitional, can be recognized by our model. To be precise, we recognize Takbeer, which is not recognized by any of the earlier studies to the best of our knowledge. This made our system capable of recognizing Witr Salat, Salat of all schools of thought where Takbeer is also offered in between the steps of Salat and the beginning of Salat.

Furthermore, the earlier sensor-based studies cannot differentiate between the postures of standing and short-standing states [3, 33, 39]. The reason behind this happening is the fact that these two stages differ only in the placement of hands, and a smartphone kept in a pocket or tied to the body is unable to capture this difference. However, a smartwatch resting on the wrist is capable of differentiating between the postures of standing and short-standing. Thus, we can successfully detect the short-standing and standing in an individual manner. Besides, while collecting data, we find many people spending not enough time in the short-standing phase or in the sitting phase in between the two prostrations. This is not considered in any of the earlier studies, and therefore, there is a chance that the existing methodology might fail to provide accurate output in such cases. However, our proposed methodology takes this into consideration, and no matter how quickly these steps are performed, can be recognized by our model. Thus, our model is capable of predicting the complete sequence of activities performed in Salat.

Tolerance to Extra Activities

We make our system robust by making it tolerant to extra activities that are not part of Salat, however, found to be often done by people [115]. Unfortunately, the task of recognizing these activities is ignored in all earlier HAR studies focusing on Salat. However, as our data collection device smartwatch rests on the wrist, and the extra activities are mostly done by a hand, they result in some extra transitions in the signal. Therefore, the actual state can get masked by the extra activities done by a hand. As in our proposed methodology, we match the transitional activities in Salat with the templates stored in our template database, considering these transitions introduced by the extra activities could result in wrong predictions. Therefore, detecting the extra activities and ignoring those is very important for our model. In this regard, we leverage DTW to eliminate the extra activities in the post-processing stage. This scenario presents the necessity of post-processing in the process of producing a complete and final prediction.

Usage of a Convenient Wearable for HAR in Salat

To recognize activities in Salat, in our study, we leverage a smartwatch as the data collection tool. To the best of our knowledge, we are the first to utilize a smartwatch in such a way for the purpose of HAR in Salat. In contrast, all the earlier sensor-based HAR studies focusing Salat used smartphones as the data collection tool.

As for their respective usages, both smartphones and smartwatches are convenient for their own purpose. Besides, smartphones appeared in the market earlier than smartwatches, and therefore, are more pervasive than smartwatches. We also find an evidence of this fact from our survey as stated in Section 6.5. However, both of these devices have different sets of advantages of their own, which make one of them more suitable for some scenarios than the other one. For example, in our case, i.e., the case of activity recognition in Salat, the data collection device is expected to be easy-to-use, convenient to pray with, non-distracting, and should be applicable to all. Smartphones, though more pervasive than smartwatches in today's world, pose some limitations from these perspectives when used as a data collection tool in HAR in Salat.

First of all, the placement of smartphones poses a problem. The placements proposed by some of the earlier studies in [2, 3], as shown in Figure 4.1, are not convenient at all. Notably, the most convenient placement for a smartphone would be to keep it in a pocket, which is also suggested in other earlier studies [33, 36, 39]. However, this is not applicable for all and in all cases either, as not all types of garments have pockets. Even if they have, it is not confirmed that the pockets would be of the same size or at the same location. On the other hand, smartwatches are worn on the wrist in general. Besides, the size and placement of the smartwatches pose almost no inconvenience to a worshipper to pray with in general. Moreover, the watch screen is in the line of sight of the user, and therefore, easy to receive any feedback provided on the screen. Therefore, when we are considering activity recognition in Salat, the smartwatch stands out in terms of convenience and practicality than the smartphones. Our findings from the survey also support this understanding. As per our survey findings, many participants explicitly expressed that, if they have to pray with a device, then they prefer a smartwatch over a smartphone. Considering all these, we can claim that, in the case of HAR in Salat, smartwatches are more convenient than smartphones.

Contribution to the HAR Literature

In our study, we propose a new methodology for recognizing a complex activity Salat and the methodology yields a near-perfect accuracy. Our study contributes to the HAR literature in several ways. First of all, the pipeline of our proposed methodology, i.e., the delicate combination of semantic rules, DTW, and custom context-based postprocessing, is the first of its kind in the HAR literature to the best of our knowledge. As we can recognize a complex activity following this new methodology with near-perfect accuracy, the methodology can be exploited in the future for recognizing other complex activities such as sports training, different types of Yoga, different types of exercises, military training, etc. More specifically, to recognize the activities that involve the sequential execution of a set of simple activities, our proposed methodology might be a good option to be explored. For example, the study in [191] recognizes six different types of exercises for frozen-shoulder rehabilitants. Each of these exercises involves specific postures and they are repeated multiple times. In such a case, we can explore our methodology to detect these exercises and count the repetitions.

On the other hand, our prepared dataset can also be valuable in the HAR literature. The dataset can serve as a benchmark dataset for complex activities. The dataset contains a good number of steady and transitional activities, and therefore, future researchers can design and test different methodologies for complex activity recognition leveraging this dataset. Nonetheless, future methodologies can also be experimented with the dataset for comparative analysis over their performances.

Besides, for the first time, our survey in this study reveals the eagerness of people to be helped in their worship. This might encourage future researchers to detect activities in other types of worship to assist people. For example, recognizing Tawaf, Sai, stoning, and other worships performed by the Hajj pilgrims [192], and providing their accurate counts to the pilgrims through activity recognition, might be some other prominent research areas worth exploring in the future.

Scope of Our Study - Recognition, Potential Extensions,and Beyond

The scope of this study is limited to the recognition of the activities performed in Salat. More specifically, this study recognizes the steady states of Salat as listed in Table 1.1 along with Takbeer. After the activities performed in Salat by a worshipper get recognized by the methodology proposed in this study, the sequence of recognized activities can be used later to assess the completeness and correctness of the prayer. The assessment can be done by an application or by the worshipper himself.

Besides, it is worth mentioning that this study covers recognizing the activities in Salat, and identifying different mistakes in Salat is its logical consequence. This happens as mistakes related to the count of various activities in Salat can be captured based on the activities recognized. Examples of such mistakes include forgetting to perform a specific activity (e.g., a Rakah), performing a specific activity more than the prescribed times, etc.

Even though activities recognized by this study can be utilized for identifying different types of mistakes, there also exist other types of mistakes that cannot be identified in this way. Examples of such mistakes include wrong recitation, wrong direction of facing, wrong postures such as placing hands on the chest instead of the belly, etc. For identifying these types of mistakes, other suitable approaches such as speech recognition, direction identification, etc., need to be performed and blended.

Chapter 10

Avenues for Future Work

In this chapter, considering the promising outcomes of our study, we further report some scopes of future work. The future scopes cover improvement of recognition methodology, an extension of coverage to different users and prayers, real-world deployment, exploring low-resource alternatives, and focusing on other application domains. Below we elaborate on each of these potential future scopes in detail.

Improvement of the Recognition Methodology

We plan to improve our activity recognition methodology in several ways. First of all, we plan to make the labeling process semi-supervised. As we manually label the data, this task takes a substantial amount of time and effort. In fact, this serves as a bottleneck to building a larger dataset. Therefore, it would be valuable if we can make the labeling process semi-supervised as done in many previous HAR studies [193, 194] to expedite the labeling process. We would also like to explore various alternatives to the currently-implemented pattern matching technique using DTW.

Besides, in Table 9.1, we compare the accuracy we achieve by applying our methodology on our collected dataset with that of the reported accuracy of the earlier studies. However, in the experiments done by the earlier studies [2, 3, 33], the experimental settings, subjects, etc., are not the same as ours. Therefore, the comparison would be more comprehensive if that could be done in the same experimental settings, which is left as our future work.

Extension of Coverage of Our Study

Future work can improve this study by including more subjects to validate the effectiveness of the presented approach at a large scale. Besides, we could not collect data from people of different schools of thought or from people, who, due to some disability or so, cannot perform

Salat in the conventional way (for example, people who perform the complete prayer while sitting). Besides, different unconventional prayers such as Salatut Tasbih, Taraweeh, etc., are yet to be covered in our study. Addressing these issues can be another potential direction for future work. The model needs to be tested with those prayer patterns and we need to make necessary adjustments in the model where needed.

Need for A Longitudinal Study

In this study, we collect data once from every subject and no longitudinal study is undertaken. However, the same person might pray differently in different situations and this difference is not captured in our dataset. Moreover, the demography of the person, the place of prayer, and the surroundings might also influence his Salat. To find out what factors influence the performance of prayer and mistake frequency and how they do so and at the same time to have a deeper observation into the patterns of mistakes in Salat, a longitudinal study can be undertaken observing the prayers of the same person over a considerable period of time.

Real-world Deployment

As we achieve satisfactory results for activity recognition in Salat, we wish to take this to the next step, which is real-life deployment generating relevant alerts and reports. To do so, we plan to develop an app with features outlined in this study that would provide details of the activities performed in Salat, i.e., their sequence, time taken to perform each activity, etc. This can be extended to work in real-time so that the activity recognition takes place online and the count of Rakat or other activities can be displayed on the watch screen as per the preference of a worshipper as well as the applicable rules in Islam. In this way, whenever a worshiper gets confused, he can check his confusion with minimal effort. Moreover, when he makes a mistake, such as missing a prostration or so, the device can generate an alert in a suitable way.

Additionally, our methodology can also be used to analyze the continuous stream of accelerometer data and detect prayer patterns, i.e., whether a person is praying or not, as done in [39]. Therefore, apps built on our proposed methodology can also act as a prayer tracker if a person wears the smartwatch during praying. Therefore, an auto-adjusting alarm feature can be developed, which can keep reminding a person about offering prayer before a sufficient time interval of that prayer. Upon detection of the prayer, the app can also send the device to silent mode so that no disturbance occurs from the device during performing the prayer. Such an app can reveal various important aspects about a person's prayer such as which prayer the person misses most, which prayer he offers late, which prayer he rushes through, which mistakes he makes more often, which activities in the prayer he spends less time, etc. A person, upon receiving this information, can take steps for improving his own prayer.

In fact, based on these future explorations, another feature worth investigating could be giving people feedback about their prayers either instantly or upon analyzing their prayers for a certain period. For example, if a person is found to miss the Fajr prayer frequently, he can be given suggestions regarding how to wake up for Fajr or the app can warn the person when he is staying up late at night. Similarly, if a person is found to hurry in his prayer or any specific activity, he can be reminded of the importance of praying in a slow and steady manner in Islam with appropriate references. Furthermore, when a Muslim travels covering s specific distance, he has to shorten his prayer, i.e., pray less number of Rakat for some prayers [28, 29]. The app can be made to keep track of these things too and remind the person about this. Thus, such an app has great potential to help Muslims to improve the quality and quantity of their prayers. Therefore, our future goal is to develop such an app and deploy this app to different types of users and ameliorate the quality of the app based on the feedback from the users after their real usage.

Utilizing Low-Resource Alternatives

In our study, the wearable we use is a high-resource one. Therefore, another possible future work could be to work next with low-resource and low-cost wearables. These days fitness bands are also becoming very pervasive [195] and they can be a good low-cost and low-resource alternative to smartwatches. In addition to that, a desirable extension of this work would be to explore the applicability of other smart devices such as smart glasses, smart earbuds, etc., [196] for the purpose of activity recognition in Salat.

Handling Variability in The Sensed Signals of Devicesfrom Different Brands

In our study, we used a single smartwatch for data collection. In reality, the signals captured by wearables of different brands or different models might vary across devices. This happens as each of them is equipped with different sensors, algorithms, etc. [197]. Handling this variability of signals is necessary to ensure the robustness of the system. Therefore, experimenting HAR in Salat with smartwatches of different models and different brands would make the study more comprehensive.

Enlarging Application Domains

Our work can serve as a guideline to recognize various activities in different application domains such as different types of Yoga, sports, firearms training, etc. On the other hand, our survey reveals the eagerness of people to be helped in their worship. Thus, future attempts can be made to detect activities in other types of worship to assist people. For example, recognizing Tawaf, Sa'i, stoning, and other worships performed by the Hajj pilgrims [192], and providing their accurate counts to the pilgrims through activity recognition, can be a potential direction of future work. Especially, in the context of Hajj, smartwatches would be a convenient and suitable option due to the heavy crowd there as well as the adoption of bands by the Hajj authority [198].

Chapter 11

Conclusion

In today's world, HAR solutions are being leveraged widely for assisting people in numerous fields for solving diversified problems [19–21]. Salat, being the most fundamental worship of the Muslim community as well as a complex activity by definition [6, 11], has got the attention of HAR researchers too over the last decade. However, several limitations still exist in these HAR studies focusing on Salat, and the literature is yet to provide a convenient robust solution for activity recognition in Salat.

To this extent, in this study, we approach to address the gaps in the literature focusing on activity recognition in Salat. However, before that, we perform an exploratory study by conducting a survey to find out the requirement and acceptability of such HAR solutions to help people in their Salat. By performing qualitative and quantitative analyses over the survey responses, we find out that mistakes are common in prayers. Besides, people in general, are willing to explore such HAR solutions provided that they are convenient to assist them in their prayers.

Subsequently, we propose an activity recognition methodology for Salat leveraging a smartwatch considering its convenience and feasibility in real-life deployment for daily use. We collect smartwatch-sensed data from 30 subjects while performing Salat and prepare a dataset having a large number of samples (3,50,762). Utilizing the dataset, we analyze and propose a new HAR methodology. In our proposed methodology, we blend together usages of machine learning algorithms, semantic rules, DTW, and custom post-processing in a delicate manner.

Rigorous experimentation reveals that our proposed methodology outperforms all the previous studies, as we achieve near-perfect accuracy (99.3%) in recognizing the activities in Salat. Here, we perform user-independent accuracy analysis so that we can analyze the actual performance of our methodology in the real world. Besides, we recognize Takbeer and differentiate between the postures of standing and short-standing to recognize them independently along with other activities in Salat and provide a complete prediction about the steps of Salat performed by a worshipper. Additionally, we consider the fact that people while praying, makes some extra activities using their hands. Such extra activities which should be overlooked, as they do not nullify Salat [115]. In our methodology, we incorporate strategies to overlook the extra activities, which increases the robustness of our proposed methodology.

Thus, our experimental results demonstrate the significant potential of our proposed methodology to accurately and robustly recognize activities in Salat using a single smartwatch. Being motivated by the potential, we plan to develop an app in the future for providing people with accurate details about their prayers so that they can observe and assess their Salat for further improvement. We also plan to scale up our study with more subjects. Our future goals also cover exploring different types of Salat and Salat performed following different schools of thought. Nonetheless, in the future, we plan to go beyond Salat and cover other religious tasks such as Tawaf, Sa'i, etc., [192] for activity recognition using a smartwatch potentially exploiting our proposed methodology or its extensions.

References

[1] Dreamstime, "Muslim man praying position step guide instructions symbol, islam religious activity icon set in flat illustration vector isolated." [Online; accessed March 14, 2023].

[2] N. Ahmad, L. Han, K. Iqbal, R. Ahmad, M. A. Abid, and N. Iqbal, "Sarm: salah activities recognition model based on smartphone," *Electronics*, vol. 8, no. 8, p. 881, 2019.

[3] R. Al-Ghannam and H. Al-Dossari, "Prayer activity monitoring and recognition using acceleration features with mobile phone," *Arabian Journal for Science and Engineering*, vol. 41, no. 12, pp. 4967–4979, 2016.

[4] Tilawat Khashiah, "Emotional recitation by mishary rashid al afasy - surah hud," 2016. [Online; accessed March 12, 2023].

[5] ShopZ.com.bd, "Samsung galaxy watch active 2 (44mm) black." https://www.shopz.com.bd/product/samsung-galaxy-watch-active-2-44mm-black/.

[6] L. Liu, Y. Peng, S. Wang, M. Liu, and Z. Huang, "Complex activity recognition using time series pattern dictionary learned from ubiquitous sensors," *Information Sciences*, vol. 340, pp. 41–57, 2016.

[7] S. Saguna, A. Zaslavsky, and D. Chakraborty, "Complex activity recognition using context-driven activity theory and activity signatures," *ACM Trans. Comput.-Hum. Interact.*, vol. 20, dec 2013.

[8] A. Jordao, A. C. Nazare Jr, J. Sena, and W. R. Schwartz, "Human activity recognition based on wearable sensor data: A standardization of the state-of-the-art," *arXiv preprint arXiv:1806.05226*, 2018.

[9] J. Ben-Arie, Z. Wang, P. Pandit, and S. Rajaram, "Human activity recognition using multidimensional indexing," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 8, pp. 1091–1104, 2002.

[10] H. C. Tan and L. C. De Silva, "Human activity recognition by head movement using elman network and neuro-markovian hybrids," in *proc. of Image and Vision Computing New Zealand*, pp. 320–326, 2003.

[11] Y. Liu, L. Nie, L. Liu, and D. S. Rosenblum, "From action to activity: Sensor-based activity recognition," *Neurocomputing*, vol. 181, pp. 108–115, 2016. Big Data Driven Intelligent Transportation Systems.

[12] A. O. Ige and M. H. Mohd Noor, "A survey on unsupervised learning for wearable sensor-based activity recognition," *Applied Soft Computing*, vol. 127, p. 109363, 2022.

[13] A. Ferrari, D. Micucci, M. Mobilio, and P. Napoletano, "Trends in human activity recognition using smartphones," *Journal of Reliable Intelligent Environments*, vol. 7, no. 3, pp. 189–213, 2021.

[14] A. Chelli and M. Pätzold, "A machine learning approach for fall detection and daily living activity recognition," *IEEE Access*, vol. 7, pp. 38670–38687, 2019.

[15] C. Rougier, J. Meunier, A. St-Arnaud, and J. Rousseau, "Robust video surveillance for fall detection based on human shape deformation," *IEEE Transactions on circuits and systems for video Technology*, vol. 21, no. 5, pp. 611–622, 2011.

[16] L. Chen and C. D. Nugent, *Human activity recognition and behaviour analysis*. Springer, 2019.

[17] M. Hynes, H. Wang, E. McCarrick, and L. Kilmartin, "Accurate monitoring of human physical activity levels for medical diagnosis and monitoring using off-the-shelf cellular handsets," *Personal and Ubiquitous Computing*, vol. 15, pp. 667–678, 2011.

[18] T. Iso and K. Yamazaki, "Gait analyzer based on a cell phone with a single three-axis accelerometer," in *Proceedings of the 8th conference on Human-computer interaction with mobile devices and services*, pp. 141–144, 2006.

[19] C. M. Ranieri, S. MacLeod, M. Dragone, P. A. Vargas, and R. A. F. Romero, "Activity recognition for ambient assisted living with videos, inertial units and ambient sensors," *Sensors*, vol. 21, no. 3, p. 768, 2021.

[20] A. Taha, H. H. Zayed, M. Khalifa, and E.-S. M. El-Horbaty, "Human activity recognition for surveillance applications," in *Proceedings of the 7th International Conference on Information Technology*, pp. 577–586, 2015.

[21] A. Avci, S. Bosch, M. Marin-Perianu, R. Marin-Perianu, and P. Havinga, "Activity recognition using inertial sensing for healthcare, wellbeing and sports applications: A

survey," in *23th International conference on architecture of computing systems 2010*, pp. 1–10, VDE, 2010.

[22] L. Minh Dang, K. Min, H. Wang, M. Jalil Piran, C. Hee Lee, and H. Moon, "Sensor-based and vision-based human activity recognition: A comprehensive survey," *Pattern Recognition*, vol. 108, p. 107561, 2020.

[23] 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, no. 1, pp. 2059–2085, 2015.

[24] H.-T. Cheng, F.-T. Sun, M. Griss, P. Davis, J. Li, and D. You, "Nuactiv: Recognizing unseen new activities using semantic attribute-based learning," in *Proceeding of the 11th annual international conference on Mobile systems, applications, and services*, pp. 361–374, 2013.

[25] D. Minnen, T. Starner, J. A. Ward, P. Lukowicz, and G. Troster, "Recognizing and discovering human actions from on-body sensor data," in *2005 IEEE International Conference on Multimedia and Expo*, pp. 1545–1548, IEEE, 2005.

[26] C. Pham and P. Olivier, "Slice&dice: Recognizing food preparation activities using embedded accelerometers," in *Ambient Intelligence: European Conference, Ami 2009, Salzburg, Austria, November 18-21, 2009. Proceedings*, pp. 34–43, Springer, 2009.

[27] S. Dernbach, B. Das, N. C. Krishnan, B. L. Thomas, and D. J. Cook, "Simple and complex activity recognition through smart phones," in *2012 eighth international conference on intelligent environments*, pp. 214–221, IEEE, 2012.

[28] M. H. Katz, *Prayer in Islamic thought and practice*. No. 6, Cambridge University Press, 2013.

[29] Y. Y. Haddad and J. I. Smith, *The Oxford Handbook of American Islam*. Oxford Handbooks, 2014.

[30] N. A. Jaafar, N. A. Ismail, and Y. A. Yusoff, "An investigation of motion tracking for solat movement with dual sensor approach," *ARPN Journal of Engineering and Applied Sciences*, vol. 10, no. 23, pp. 17981–17986, 2015.

[31] A. Koubâa, A. Ammar, B. Benjdira, A. Al-Hadid, B. Kawaf, S. A. Al-Yahri, A. Babiker, K. Assaf, and M. B. Ras, "Activity monitoring of islamic prayer (salat) postures using deep learning," in *2020 6th Conference on Data Science and Machine Learning Applications (CDMA)*, pp. 106–111, IEEE, 2020.

[32] M. M. Rahman, R. A. A. Alharazi, and M. K. I. B. Z. Badri, "Monitoring and alarming activity of islamic prayer (salat) posture using image processing," in *2021 8th International*

Conference on Computer and Communication Engineering (ICCCE), pp. 238–243, IEEE, 2021.

[33] O. Alobaid and K. Rasheed, "Prayer activity recognition using an accelerometer sensor," in *Proceedings on the International Conference on Artificial Intelligence (ICAI)*, pp. 271–277, The Steering Committee of The World Congress in Computer Science, Computer . . . , 2018.

[34] A. Muaremi, J. Seiter, F. Gravenhorst, G. Tröster, A. Bexheti, and B. Arnrich, "Monitor pilgrims: prayer activity recognition using wearable sensors," in *Proceedings of the 8th International Conference on Body Area Networks*, pp. 161–164, 2013.

[35] M. Ali, M. Shafi, U. Farooq, *et al.*, "Salat activity recognition using smartphone triaxial accelerometer," in *2018 5th International Multi-Topic ICT Conference (IMTIC)*, pp. 1–7, IEEE, 2018.

[36] T. A. Topu, M. M. Rahman, M. S. Hossain, and A. Al Marouf, "Prayer activity recognition using smartphone," in *2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, pp. 1–6, IEEE, 2021.

[37] I. Q. . Answer, "Importance of prayer," 2001. `https://islamqa.info/en/answers/12305/importance-of-prayer`, Last accessed on 2023-4-8.

[38] P. R. Center, "The world's muslims: Unity and diversity," 2012. `https://www.pewresearch.org/religion/2012/08/09/the-worlds-muslims-unity-and-diversity-2-religious-commitment/`.

[39] K. Eskaf, W. M. Aly, and A. Aly, "Aggregated activity recognition using smart devices," in *2016 3rd International Conference on Soft Computing & Machine Intelligence (ISCMI)*, pp. 214–218, IEEE, 2016.

[40] G. Bieber, M. Haescher, and M. Vahl, "Sensor requirements for activity recognition on smart watches," in *Proceedings of the 6th International Conference on PErvasive Technologies Related to Assistive Environments*, pp. 1–6, 2013.

[41] L. Porzi, S. Messelodi, C. M. Modena, and E. Ricci, "A smart watch-based gesture recognition system for assisting people with visual impairments," in *Proceedings of the 3rd ACM international workshop on Interactive multimedia on mobile & portable devices*, pp. 19–24, 2013.

[42] G. Bieber, T. Kirste, and B. Urban, "Ambient interaction by smart watches," in *Proceedings of the 5th international conference on pervasive technologies related to assistive environments*, pp. 1–6, 2012.

[43] E. Glowacki, Y. Zhu, E. Hunt, K. Magsamen-Conrad, and J. Bernhardt, "Facilitators and barriers to smartwatch use among individuals with chronic diseases: A qualitative study," *University of Texas, Austin. Accessed November*, vol. 11, p. 2018, 2016.

[44] G. M. Weiss, J. L. Timko, C. M. Gallagher, K. Yoneda, and A. J. Schreiber, "Smartwatch-based activity recognition: A machine learning approach," in *2016 IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI)*, pp. 426–429, IEEE, 2016.

[45] S. Bhattacharya and N. D. Lane, "From smart to deep: Robust activity recognition on smartwatches using deep learning," in *2016 IEEE International conference on pervasive computing and communication workshops (PerCom Workshops)*, pp. 1–6, IEEE, 2016.

[46] "Samsung Galaxy Watch Active 2 review: the best smartwatch for Android." https://www.theguardian.com/technology/2020/may/13/samsung-galaxy-watch-active-2-review-best-smartwatch-for-android. Last accessed on: March 12, 2023.

[47] L. Peng, L. Chen, Z. Ye, and Y. Zhang, "Aroma: A deep multi-task learning based simple and complex human activity recognition method using wearable sensors," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 2, jul 2018.

[48] K. Chen, D. Zhang, L. Yao, B. Guo, Z. Yu, and Y. Liu, "Deep learning for sensor-based human activity recognition: Overview, challenges, and opportunities," *ACM Computing Surveys (CSUR)*, vol. 54, no. 4, pp. 1–40, 2021.

[49] C. J. Caspersen, K. E. Powell, and G. M. Christenson, "Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research.," *Public health reports*, vol. 100, no. 2, p. 126, 1985.

[50] L. Minh Dang, K. Min, H. Wang, M. Jalil Piran, C. Hee Lee, and H. Moon, "Sensor-based and vision-based human activity recognition: A comprehensive survey," *Pattern Recognition*, vol. 108, p. 107561, 2020.

[51] Q. Li, R. Gravina, and G. Fortino, "Posture and gesture analysis supporting emotional activity recognition," in *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pp. 2742–2747, IEEE, 2018.

[52] M. Derawi and P. Bours, "Gait and activity recognition using commercial phones," *computers & security*, vol. 39, pp. 137–144, 2013.

[53] R. Mohamed, T. Perumal, M. N. Sulaiman, and N. Mustapha, "Multi resident complex activity recognition in smart home: A literature review," *Int. J. Smart Home*, vol. 11, no. 6, pp. 21–32, 2017.

[54] V. Osmani, S. Balasubramaniam, and D. Botvich, "Human activity recognition in pervasive health-care: Supporting efficient remote collaboration," *Journal of network and computer applications*, vol. 31, no. 4, pp. 628–655, 2008.

[55] M. Al-Tusi, "A concise description of islamic law and legal opinions," 2008.

[56] M. J. Mughniyya, "Prayer (salat), according to the five islamic schools of law," 1997. https://www.al-islam.org/prayer-salat-according-five-islamic-schools-law-muhammad-jawad-mughniyya, Last accessed on 2023-2-27.

[57] A.-H. Murad, "Understanding the four madhhabs," 1997. http://www.masud.co.uk/ISLAM/ahm/newmadhh.htm, Last accessed on 2023-2-27.

[58] "Mistakes in prayer," 2014. https://islamqa.info/en/answers/173637/mistakes-in-the-prayer, Last accessed on 2023-2-27.

[59] J. R. Kwapisz, G. M. Weiss, and S. A. Moore, "Activity recognition using cell phone accelerometers," *ACM SigKDD Explorations Newsletter*, vol. 12, no. 2, pp. 74–82, 2011.

[60] M. J. Mathie, B. G. Celler, N. H. Lovell, and A. C. Coster, "Classification of basic daily movements using a triaxial accelerometer," *Medical and Biological Engineering and Computing*, vol. 42, no. 5, pp. 679–687, 2004.

[61] A. Salarian, H. Russmann, F. J. Vingerhoets, P. R. Burkhard, and K. Aminian, "Ambulatory monitoring of physical activities in patients with parkinson's disease," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 12, pp. 2296–2299, 2007.

[62] M. M. Hassan, M. Z. Uddin, A. Mohamed, and A. Almogren, "A robust human activity recognition system using smartphone sensors and deep learning," *Future Generation Computer Systems*, vol. 81, pp. 307–313, 2018.

[63] F. Attal, S. Mohammed, M. Dedabrishvili, F. Chamroukhi, L. Oukhellou, and Y. Amirat, "Physical human activity recognition using wearable sensors," *Sensors*, vol. 15, no. 12, pp. 31314–31338, 2015.

[64] M. Zhang and A. A. Sawchuk, "Usc-had: A daily activity dataset for ubiquitous activity recognition using wearable sensors," in *Proceedings of the 2012 ACM conference on ubiquitous computing*, pp. 1036–1043, 2012.

[65] A. Z. M. Faridee, S. R. Ramamurthy, H. S. Hossain, and N. Roy, "Happyfeet: Recognizing and assessing dance on the floor," in *Proceedings of the 19th International Workshop on Mobile Computing Systems & Applications*, pp. 49–54, 2018.

[66] F. R. Sayem, M. M. Sheikh, and M. A. R. Ahad, “Feature-based method for nurse care complex activity recognition from accelerometer sensor,” in *Adjunct Proceedings of the 2021 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2021 ACM International Symposium on Wearable Computers*, pp. 446–451, 2021.

[67] P. Bharti, D. De, S. Chellappan, and S. K. Das, “Human: Complex activity recognition with multi-modal multi-positional body sensing,” *IEEE Transactions on Mobile Computing*, vol. 18, no. 4, pp. 857–870, 2019.

[68] A. Afiq, M. Zakariya, M. Saad, A. Nurfarzana, M. H. M. Khir, A. Fadzil, A. Jale, W. Gunawan, Z. Izuddin, and M. Faizari, “A review on classifying abnormal behavior in crowd scene,” *Journal of Visual Communication and Image Representation*, vol. 58, pp. 285–303, 2019.

[69] U. A. Akansha, M. Shailendra, and N. Singh, “Analytical review on video-based human activity recognition,” in *2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom)*, pp. 3839–3844, IEEE, 2016.

[70] J. Shao, K. Kang, C. Change Loy, and X. Wang, “Deeply learned attributes for crowded scene understanding,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4657–4666, 2015.

[71] J. P. Vital, D. R. Faria, G. Dias, M. S. Couceiro, F. Coutinho, and N. M. Ferreira, “Combining discriminative spatiotemporal features for daily life activity recognition using wearable motion sensing suit,” *Pattern Analysis and Applications*, vol. 20, pp. 1179–1194, 2017.

[72] S. Jadooki, D. Mohamad, T. Saba, A. S. Almazyad, and A. Rehman, “Fused features mining for depth-based hand gesture recognition to classify blind human communication,” *Neural Computing and Applications*, vol. 28, pp. 3285–3294, 2017.

[73] T. Liu, Z. Chen, H. Liu, Z. Zhang, and Y. Chen, “Multi-modal hand gesture designing in multi-screen touchable teaching system for human-computer interaction,” in *Proceedings of the 2nd International Conference on Advances in Image Processing*, ICAIP ’18, (New York, NY, USA), p. 198–202, Association for Computing Machinery, 2018.

[74] O. D. Lara, A. J. Pérez, M. A. Labrador, and J. D. Posada, “Centinela: A human activity recognition system based on acceleration and vital sign data,” *Pervasive and mobile computing*, vol. 8, no. 5, pp. 717–729, 2012.

[75] Y.-S. Lee and S.-B. Cho, "Activity recognition with android phone using mixture-of-experts co-trained with labeled and unlabeled data," *Neurocomputing*, vol. 126, pp. 106–115, 2014.

[76] J.-L. Reyes-Ortiz, L. Oneto, A. Samà, X. Parra, and D. Anguita, "Transition-aware human activity recognition using smartphones," *Neurocomputing*, vol. 171, pp. 754–767, 2016.

[77] Z. Yan, V. Subbaraju, D. Chakraborty, A. Misra, and K. Aberer, "Energy-efficient continuous activity recognition on mobile phones: An activity-adaptive approach," in *2012 16th international symposium on wearable computers*, pp. 17–24, Ieee, 2012.

[78] M. Köse, Ö. D. İncel, and C. Ersoy, "Performance evaluation of classification methods for online activity recognition on smart phones," in *2012 20th Signal Processing and Communications Applications Conference (SIU)*, pp. 1–4, IEEE, 2012.

[79] O. Ogbanufe and N. Gerhart, "Watch it! factors driving continued feature use of the smartwatch," *International Journal of Human–Computer Interaction*, vol. 34, no. 11, pp. 999–1014, 2018.

[80] G. M. Weiss, K. Yoneda, and T. Hayajneh, "Smartphone and smartwatch-based biometrics using activities of daily living," *IEEE Access*, vol. 7, pp. 133190–133202, 2019.

[81] R.-A. Voicu, C. Dobre, L. Bajenaru, and R.-I. Ciobanu, "Human physical activity recognition using smartphone sensors," *Sensors*, vol. 19, no. 3, p. 458, 2019.

[82] C. A. Ronao and S.-B. Cho, "Human activity recognition with smartphone sensors using deep learning neural networks," *Expert systems with applications*, vol. 59, pp. 235–244, 2016.

[83] H. Wang, T. T.-T. Lai, and R. Roy Choudhury, "Mole: Motion leaks through smartwatch sensors," in *Proceedings of the 21st annual international conference on mobile computing and networking*, pp. 155–166, 2015.

[84] S. Balli, E. A. Sağbaş, and M. Peker, "Human activity recognition from smart watch sensor data using a hybrid of principal component analysis and random forest algorithm," *Measurement and Control*, vol. 52, no. 1-2, pp. 37–45, 2019.

[85] J. R. Quinlan, *C4. 5: programs for machine learning*. Elsevier, 2014.

[86] L. Bao and S. S. Intille, "Activity recognition from user-annotated acceleration data," in *Pervasive Computing: Second International Conference, PERVASIVE 2004, Linz/Vienna, Austria, April 21-23, 2004. Proceedings 2*, pp. 1–17, Springer, 2004.

[87] N. Ravi, N. Dandekar, P. Mysore, and M. L. Littman, “Activity recognition from accelerometer data,” in *Aaai*, vol. 5, pp. 1541–1546, Pittsburgh, PA, 2005.

[88] H. Mart´ın, A. M. Bernardos, J. Iglesias, and J. R. Casar, “Activity logging using lightweight classification techniques in mobile devices,” *Personal and ubiquitous computing*, vol. 17, pp. 675–695, 2013.

[89] L. Xu, W. Yang, Y. Cao, and Q. Li, “Human activity recognition based on random forests,” in *2017 13th international conference on natural computation, fuzzy systems and knowledge discovery (ICNC-FSKD)*, pp. 548–553, IEEE, 2017.

[90] A. Wang, H. Chen, C. Zheng, L. Zhao, J. Liu, and L. Wang, “Evaluation of random forest for complex human activity recognition using wearable sensors,” in *2020 International Conference on Networking and Network Applications (NaNA)*, pp. 310–315, IEEE, 2020.

[91] X. Yun, J. Calusdian, E. R. Bachmann, and R. B. McGhee, “Estimation of human foot motion during normal walking using inertial and magnetic sensor measurements,” *IEEE transactions on Instrumentation and Measurement*, vol. 61, no. 7, pp. 2059–2072, 2012.

[92] D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. L. Reyes-Ortiz, “Human activity recognition on smartphones using a multiclass hardware-friendly support vector machine,” in *Ambient Assisted Living and Home Care: 4th International Workshop, IWAAL 2012, Vitoria-Gasteiz, Spain, December 3-5, 2012. Proceedings 4*, pp. 216–223, Springer, 2012.

[93] N. Ahmed, J. I. Rafiq, and M. R. Islam, “Enhanced human activity recognition based on smartphone sensor data using hybrid feature selection model,” *Sensors*, vol. 20, no. 1,p. 317, 2020.

[94] P. Siirtola and J. Röning, “Recognizing human activities user-independently on smartphones based on accelerometer data,” *IJIMAI*, vol. 1, no. 5, pp. 38–45, 2012.

[95] T. Brezmes, J.-L. Gorricho, and J. Cotrina, “Activity recognition from accelerometer data on a mobile phone,” in *Distributed Computing, Artificial Intelligence, Bioinformatics, Soft Computing, and Ambient Assisted Living: 10th International Work-Conference on Artificial Neural Networks, IWANN 2009 Workshops, Salamanca, Spain, June 10-12, 2009. Proceedings, Part II 10*, pp. 796–799, Springer, 2009.

[96] O. D. Lara and M. A. Labrador, “A mobile platform for real-time human activity recognition,” in *2012 IEEE consumer communications and networking conference (CCNC)*, pp. 667–671, IEEE, 2012.

[97] Y. Kim and T. Moon, “Human detection and activity classification based on micro-doppler signatures using deep convolutional neural networks,” *IEEE geoscience and remote sensing letters*, vol. 13, no. 1, pp. 8–12, 2015.

[98] J. Donahue, L. Anne Hendricks, S. Guadarrama, M. Rohrbach, S. Venugopalan, K. Saenko, and T. Darrell, "Long-term recurrent convolutional networks for visual recognition and description," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2625–2634, 2015.

[99] S. Ha and S. Choi, "Convolutional neural networks for human activity recognition using multiple accelerometer and gyroscope sensors," in *2016 international joint conference on neural networks (IJCNN)*, pp. 381–388, IEEE, 2016.

[100] S.-M. Lee, S. M. Yoon, and H. Cho, "Human activity recognition from accelerometer data using convolutional neural network," in *2017 ieee international conference on big data and smart computing (bigcomp)*, pp. 131–134, IEEE, 2017.

[101] S. Mekruksavanich and A. Jitpattanakul, "Rnn-based deep learning for physical activity recognition using smartwatch sensors: A case study of simple and complex activity recognition," *Mathematical Biosciences and Engineering*, vol. 19, no. 6, pp. 5671–5698, 2022.

[102] M. Lv, W. Xu, and T. Chen, "A hybrid deep convolutional and recurrent neural network for complex activity recognition using multimodal sensors," *Neurocomputing*, vol. 362, pp. 33–40, 2019.

[103] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[104] Y. Guan and T. Plötz, "Ensembles of deep lstm learners for activity recognition using wearables," *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies*, vol. 1, no. 2, pp. 1–28, 2017.

[105] Y. Zhao, R. Yang, G. Chevalier, X. Xu, and Z. Zhang, "Deep residual bidir-lstm for human activity recognition using wearable sensors," *Mathematical Problems in Engineering*, vol. 2018, pp. 1–13, 2018.

[106] N. H. Friday, M. A. Al-garadi, G. Mujtaba, U. R. Alo, and A. Waqas, "Deep learning fusion conceptual frameworks for complex human activity recognition using mobile and wearable sensors," in *2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)*, pp. 1–7, IEEE, 2018.

[107] M. Lv, W. Xu, and T. Chen, "A hybrid deep convolutional and recurrent neural network for complex activity recognition using multimodal sensors," *Neurocomputing*, vol. 362, pp. 33–40, 2019.

[108] S. Seto, W. Zhang, and Y. Zhou, "Multivariate time series classification using dynamic time warping template selection for human activity recognition," in *2015 IEEE symposium series on computational intelligence*, pp. 1399–1406, IEEE, 2015.

[109] C. O'Rourke and M. Madden, "Activity recognition based on accelerometer data using dynamic time warping with ensembles," in *Proceedings of AICS-2011, the 22 nd Conference on Artificial Intelligence and Cognitive Science*, Citeseer, 2011.

[110] F. Nurwanto, I. Ardiyanto, and S. Wibirama, "Light sport exercise detection based on smartwatch and smartphone using k-nearest neighbor and dynamic time warping algorithm," in *2016 8th International Conference on Information Technology and Electrical Engineering (ICITEE)*, pp. 1–5, IEEE, 2016.

[111] R. Varatharajan, G. Manogaran, M. K. Priyan, and R. Sundarasekar, "Wearable sensor devices for early detection of alzheimer disease using dynamic time warping algorithm," *Cluster Computing*, vol. 21, pp. 681–690, 2018.

[112] Y. Li, D. Xue, E. Forrister, G. Lee, B. Garner, and Y. Kim, "Human activity classification based on dynamic time warping of an on-body creeping wave signal," *IEEE Transactions on Antennas and Propagation*, vol. 64, no. 11, pp. 4901–4905, 2016.

[113] M. El-Hoseiny and E. Shaban, "Muslim prayer actions recognition," in *2009 Second International Conference on Computer and Electrical Engineering*, vol. 1, pp. 460–465, IEEE, 2009.

[114] A. Dehghani, T. Glatard, and E. Shihab, "Subject cross validation in human activity recognition," *arXiv preprint arXiv:1904.02666*, 2019.

[115] .-. Fataawa 'Ulama' al-Balad al Haraam, "Moving whilst praying," 2011. https://islamqa.info/en/answers/12683/moving-whilst-praying, Last accessed on 2023-2-27.

[116] H. Qian, S. J. Pan, and C. Miao, "Latent independent excitation for generalizable sensor-based cross-person activity recognition," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, pp. 11921–11929, 2021.

[117] AskIslamPedia, "Maqam e ibrahim." https://www.askislampedia.com/en/wiki/-/wiki/English_wiki/Maqam+E+Ibrahim, Last accessed on 2023-4-2.

[118] A. P. of Pakistan. https://www.app.com.pk/photos-section/prime-minister-muhammad-shehbaz-sharif-offering-prayer-in-riaz-ul-jannah-masjid-al-nabwi-saw/, Last accessed on 2023-4-2.

[119] R. San-Segundo, H. Blunck, J. Moreno-Pimentel, A. Stisen, and M. Gil-Mart´ın, "Robust human activity recognition using smartwatches and smartphones," *Engineering Applications of Artificial Intelligence*, vol. 72, pp. 190–202, 2018.

[120] R. Chavarriaga, H. Sagha, A. Calatroni, S. T. Digumarti, G. Tröster, J. del R. Millán, and D. Roggen, "The opportunity challenge: A benchmark database for on-body sensor-based activity recognition," *Pattern Recognition Letters*, vol. 34, no. 15, pp. 2033–2042, 2013. Smart Approaches for Human Action Recognition.

[121] S. I. Salim, N. A. Al-Nabhan, M. Rahaman, N. Islam, T. R. Toha, J. Noor, A. Quaium, A. Mostak, M. Hossain, M. M. Mushfiq, *et al.*, "Human-survey interaction (hsi): A study on integrity of human data collectors in a mass-scale hajj pilgrimage survey," *IEEE Access*, vol. 9, pp. 112528–112551, 2021.

[122] M. C. Monteiro de Barros, F. C. Leão, H. Vallada Filho, G. Lucchetti, A. Moreira-Almeida, and M. F. Prieto Peres, "Prevalence of spiritual and religious experiences in the general population: A brazilian nationwide study," *Transcultural Psychiatry*, p. 13634615221088701, 2022.

[123] I. Subchi, Z. Zulkifli, R. Latifa, and S. Sa'diyah, "Religious moderation in indonesian muslims," *Religions*, vol. 13, no. 5, p. 451, 2022.

[124] R. R. Griffiths, E. S. Hurwitz, A. K. Davis, M. W. Johnson, and R. Jesse, "Survey of subjective" god encounter experiences": Comparisons among naturally occurring experiences and those occasioned by the classic psychedelics psilocybin, lsd, ayahuasca, or dmt," *PloS one*, vol. 14, no. 4, p. e0214377, 2019.

[125] A. Weigold, I. K. Weigold, and E. J. Russell, "Examination of the equivalence of self-report survey-based paper-and-pencil and internet data collection methods.," *Psychological methods*, vol. 18, no. 1, p. 53, 2013.

[126] R. Crutzen and A. S. Göritz, "Social desirability and self-reported health risk behaviors in web-based research: three longitudinal studies," *BMC public health*, vol. 10, no. 1, pp. 1–10, 2010.

[127] C. J. Hopwood, E. W. Good, and L. C. Morey, "Validity of the dsm–5 levels of personality functioning scale–self report," *Journal of Personality Assessment*, vol. 100, no. 6, pp. 650–659, 2018.

[128] L. Mok and A. Anderson, "The complementary nature of perceived and actual time spent online in measuring digital well-being," *Proceedings of the ACM on human-computer interaction*, vol. 5, no. CSCW1, pp. 1–27, 2021.

[129] J. Junger-Tas and I. H. Marshall, "The self-report methodology in crime research," *Crime and justice*, vol. 25, pp. 291–367, 1999.

[130] H. F. Levin-Aspenson and D. Watson, "Mode of administration effects in psychopathology assessment: Analyses of gender, age, and education differences in self-rated versus interview-based depression.," *Psychological Assessment*, vol. 30, no. 3, p. 287, 2018.

[131] C. Wacharamanotham, L. Eisenring, S. Haroz, and F. Echtler, "Transparency of chi research artifacts: Results of a self-reported survey," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1–14, 2020.

[132] F. Shahid, W. Rahman, A. B. Islam, N. Paul, N. Khan, M. S. Rahman, M. M. Haque, and A. A. Al Islam, "Two tell-tale perspectives of ptsd: neurobiological abnormalities and bayesian regulatory network of the underlying disorder in a refugee context," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 3, no. 3, pp. 1–45, 2019.

[133] F. Shahid, W. Rahman, M. S. Rahman, S. A. Purabi, A. Seddiqa, M. Mostakim, F. Feroz, T. R. Soron, F. Hossain, N. Khan, *et al.*, "Leveraging free-hand sketches for potential screening of ptsd," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 4, no. 3, pp. 1–22, 2020.

[134] R. Likert, "The likert-type scale," *Archives of Psychology*, vol. 140, no. 55, pp. 1–55, 1932.

[135] P. Bhardwaj *et al.*, "Types of sampling in research," *Journal of the Practice of Cardiovascular Sciences*, vol. 5, no. 3, p. 157, 2019.

[136] M. Hibberts, R. Burke Johnson, and K. Hudson, "Common survey sampling techniques," *Handbook of survey methodology for the social sciences*, pp. 53–74, 2012.

[137] C. H. Goulden *et al.*, "Methods of statistical analysis.," *Methods of statistical analysis.*, no. 2nd ed, 1952.

[138] P. E. McKnight and J. Najab, "Mann-whitney u test," *The Corsini encyclopedia of psychology*, pp. 1–1, 2010.

[139] W. Kruskal, "Kruskal and wallis' test," *Journal of American Statisticians Association*, pp. 583–618, 1952.

[140] M. J. Eynon, C. O'Donnell, and L. Williams, "Gaining qualitative insight into the subjective experiences of adherers to an exercise referral scheme: A thematic analysis," *Journal of Health Psychology*, vol. 23, no. 11, pp. 1476–1487, 2018.

[141] V. Clarke, V. Braun, and N. Hayfield, "Thematic analysis," *Qualitative psychology: A practical guide to research methods*, vol. 3, pp. 222–248, 2015.

[142] D. Karantonis, M. Narayanan, M. Mathie, N. Lovell, and B. Celler, "Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring," *IEEE Transactions on Information Technology in Biomedicine*, vol. 10, no. 1, pp. 156–167, 2006.

[143] J. Sharp, "The moving average filter," 1977.

[144] A. M. Khan, Y.-K. Lee, S.-Y. Lee, and T.-S. Kim, "Human activity recognition via an accelerometer-enabled-smartphone using kernel discriminant analysis," in *2010 5th international conference on future information technology*, pp. 1–6, IEEE, 2010.

[145] A. M. Khan, Y.-K. Lee, S. Y. Lee, and T.-S. Kim, "A triaxial accelerometer-based physical-activity recognition via augmented-signal features and a hierarchical recognizer," *IEEE transactions on information technology in biomedicine*, vol. 14, no. 5, pp. 1166–1172, 2010.

[146] B. Quigley, M. Donnelly, G. Moore, and L. Galway, "A comparative analysis of windowing approaches in dense sensing environments," in *Proceedings*, vol. 2, p. 1245, MDPI, 2018.

[147] O. Banos, J.-M. Galvez, M. Damas, H. Pomares, and I. Rojas, "Window size impact in human activity recognition," *Sensors*, vol. 14, no. 4, pp. 6474–6499, 2014.

[148] L. Cao, Y. Wang, B. Zhang, Q. Jin, and A. V. Vasilakos, "Gchar: An efficient group-based context—aware human activity recognition on smartphone," *Journal of Parallel and Distributed Computing*, vol. 118, pp. 67–80, 2018.

[149] N. A. Capela, E. D. Lemaire, and N. Baddour, "Feature selection for wearable smartphone-based human activity recognition with able bodied, elderly, and stroke patients," *PloS one*, vol. 10, no. 4, p. e0124414, 2015.

[150] J. Wang, Y. Chen, S. Hao, X. Peng, and L. Hu, "Deep learning for sensor-based activity recognition: A survey," *Pattern Recognition Letters*, vol. 119, pp. 3–11, 2019. Deep Learning for Pattern Recognition.

[151] A. Ferrari, D. Micucci, M. Mobilio, and P. Napoletano, "Hand-crafted features vs residual networks for human activities recognition using accelerometer," in *2019 IEEE 23rd International Symposium on Consumer Technologies (ISCT)*, pp. 153–156, IEEE, 2019.

[152] C. Hu, Y. Chen, L. Hu, and X. Peng, "A novel random forests based class incremental learning method for activity recognition," *Pattern Recognition*, vol. 78, pp. 277–290, 2018.

[153] W. I. D. Mining, “Data mining: Concepts and techniques,” *Morgan Kaufinann*, vol. 10, pp. 559–569, 2006.

[154] T. Hastie, R. Tibshirani, J. H. Friedman, and J. H. Friedman, *The elements of statistical learning: data mining, inference, and prediction*, vol. 2. Springer, 2009.

[155] X. Yang, A. Dinh, and L. Chen, “Implementation of a wearerable real-time system for physical activity recognition based on naive bayes classifier,” in *2010 International Conference on Bioinformatics and Biomedical Technology*, pp. 101–105, IEEE, 2010.

[156] A. R. Jiménez and F. Seco, “Multi-event naive bayes classifier for activity recognition in the ucami cup,” in *Proceedings*, vol. 60, MDPI, 2018.

[157] D. P. Kingma and M. Welling, “Auto-encoding variational bayes,” *arXiv preprint arXiv:1312.6114*, 2013.

[158] D. Micucci, M. Mobilio, and P. Napoletano, “Unimib shar: A dataset for human activity recognition using acceleration data from smartphones,” *Applied Sciences*, vol. 7, no. 10, p. 1101, 2017.

[159] D. Anguita, A. Ghio, L. Oneto, X. Parra, J. L. Reyes-Ortiz, *et al.*, “A public domain dataset for human activity recognition using smartphones.,” in *Esann*, vol. 3, p. 3, 2013.

[160] N. A. Capela, E. D. Lemaire, and N. Baddour, “Improving classification of sit, stand, and lie in a smartphone human activity recognition system,” in *2015 IEEE International Symposium on Medical Measurements and Applications (MeMeA) Proceedings*, pp. 473–478, IEEE, 2015.

[161] M.-W. Lee, A. M. Khan, and T.-S. Kim, “A single tri-axial accelerometer-based real-time personal life log system capable of human activity recognition and exercise information generation,” *Personal and Ubiquitous Computing*, vol. 15, pp. 887–898, 2011.

[162] J.-Y. Yang, J.-S. Wang, and Y.-P. Chen, “Using acceleration measurements for activity recognition: An effective learning algorithm for constructing neural classifiers,” *Pattern recognition letters*, vol. 29, no. 16, pp. 2213–2220, 2008.

[163] S. Nazir, M. H. Yousaf, and S. A. Velastin, “Inter and intra class correlation analysis (iicca) for human action recognition in realistic scenarios,” in *8th International Conference of Pattern Recognition Systems (ICPRS 2017)*, pp. 1–6, IET, 2017.

[164] M. Ziaeefard and R. Bergevin, “Semantic human activity recognition: A literature review,” *Pattern Recognition*, vol. 48, no. 8, pp. 2329–2345, 2015.

[165] B. Su, Q. Tang, J. Jiang, M. Sheng, A. A. Yahya, and G. Wang, "A novel method for short-time human activity recognition based on improved template matching technique," in *Proceedings of the 15th ACM SIGGRAPH Conference on Virtual-Reality Continuum and Its Applications in Industry-Volume 1*, pp. 233–242, 2016.

[166] G. Bieber, P. Koldrack, C. Sablowski, C. Peter, and B. Urban, "Mobile physical activity recognition of stand-up and sit-down transitions for user behavior analysis," in *Proceedings of the 3rd International Conference on PErvasive Technologies Related to Assistive Environments*, pp. 1–5, 2010.

[167] V. Ahanathapillai, J. D. Amor, M. Tadeusiak, and C. J. James, "Wrist-worn accelerometer to detect postural transitions and walking patterns," in *XIII Mediterranean Conference on Medical and Biological Engineering and Computing 2013* (L. M. Roa Romero, ed.), (Cham), pp. 1515–1518, Springer International Publishing, 2014.

[168] B. Su, Q. Tang, J. Jiang, M. Sheng, A. A. Yahya, and G. Wang, "A novel method for short-time human activity recognition based on improved template matching technique," in *Proceedings of the 15th ACM SIGGRAPH Conference on Virtual-Reality Continuum and Its Applications in Industry-Volume 1*, pp. 233–242, 2016.

[169] J. Margarito, R. Helaoui, A. M. Bianchi, F. Sartor, and A. G. Bonomi, "User-independent recognition of sports activities from a single wrist-worn accelerometer: A template-matching-based approach," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 4, pp. 788–796, 2015.

[170] T. K. Vintsyuk, "Speech discrimination by dynamic programming," *Cybernetics*, vol. 4, no. 1, pp. 52–57, 1968.

[171] D. M. Karantonis, M. R. Narayanan, M. Mathie, N. H. Lovell, and B. G. Celler, "Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring," *IEEE transactions on information technology in biomedicine*, vol. 10, no. 1, pp. 156–167, 2006.

[172] M. Straczkiewicz, P. James, and J.-P. Onnela, "A systematic review of smartphone-based human activity recognition for health research," *arXiv preprint arXiv:1910.03970*, 2019.

[173] J.-H. Li, L. Tian, H. Wang, Y. An, K. Wang, and L. Yu, "Segmentation and recognition of basic and transitional activities for continuous physical human activity," *IEEE access*, vol. 7, pp. 42565–42576, 2019.

[174] S. Salvador and P. Chan, "Toward accurate dynamic time warping in linear time and space," *Intelligent Data Analysis*, vol. 11, no. 5, pp. 561–580, 2007.

[175] Y. Permanasari, E. H. Harahap, and E. P. Ali, “Speech recognition using dynamic time warping (dtw),” in *Journal of physics: Conference series*, vol. 1366, p. 012091, IOP Publishing, 2019.

[176] A. V. Haridas, R. Marimuthu, and V. G. Sivakumar, “A critical review and analysis on techniques of speech recognition: The road ahead,” *International Journal of Knowledge-Based and Intelligent Engineering Systems*, vol. 22, no. 1, pp. 39–57, 2018.

[177] S. K. Saksamudre, P. Shrishrimal, and R. Deshmukh, “A review on different approaches for speech recognition system,” *International Journal of Computer Applications*, vol. 115, no. 22, 2015.

[178] A. Mohod and D. S. Datar, “An efficient and elastic approach for partial shape matching using dtw,” *Advent Technology*, vol. 1, no. 4, 2013.

[179] M. J. Mathie, B. G. Celler, N. H. Lovell, and A. C. Coster, “Classification of basic daily movements using a triaxial accelerometer,” *Medical and Biological Engineering and Computing*, vol. 42, pp. 679–687, 2004.

[180] D. Riboni and C. Bettini, “Cosar: hybrid reasoning for context-aware activity recognition,” *Personal and Ubiquitous Computing*, vol. 15, pp. 271–289, 2011.

[181] D. Berrar, “Cross-validation.,” 2019.

[182] M. Zhang and A. A. Sawchuk, “Usc-had: A daily activity dataset for ubiquitous activity recognition using wearable sensors,” in *Proceedings of the 2012 ACM conference on ubiquitous computing*, pp. 1036–1043, 2012.

[183] B. Darul Ifta, “Difference between men and women prayers.” https://islamqa.org/hanafi/daruliftaa-birmingham/19580/difference-between-men-and-women-prayers/, Last accessed on 2023-3-8.

[184] L. Schrader, A. Vargas Toro, S. Konietzny, S. Rüping, B. Schäpers, M. Steinböck, C. Krewer, F. Müller, J. Güttler, and T. Bock, “Advanced sensing and human activity recognition in early intervention and rehabilitation of elderly people,” *Journal of Population Ageing*, vol. 13, pp. 139–165, 2020.

[185] R. Proffitt, “Acceptability of in-home ambient sensors for activity recognition and assessment post-stroke,” *Archives of Physical Medicine and Rehabilitation*, vol. 103, no. 12, pp. e53–e54, 2022.

[186] J. W. Lockhart and G. M. Weiss, “Limitations with activity recognition methodology & data sets,” in *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*, pp. 747–756, 2014.

[187] K. Caine, “Local standards for sample size at chi,” in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI ’16, (New York, NY, USA), p. 981–992, Association for Computing Machinery, 2016.

[188] S. Mekruksavanich and A. Jitpattanakul, “Deep convolutional neural network with rnns for complex activity recognition using wrist-worn wearable sensor data,” *Electronics*, vol. 10, no. 14, p. 1685, 2021.

[189] C. Xu, D. Chai, J. He, X. Zhang, and S. Duan, “Innohar: A deep neural network for complex human activity recognition,” *Ieee Access*, vol. 7, pp. 9893–9902, 2019.

[190] A. Ferrari, D. Micucci, M. Mobilio, and P. Napoletano, “On the personalization of classification models for human activity recognition,” *IEEE Access*, vol. 8, pp. 32066–32079, 2020.

[191] H.-C. Lin, S.-Y. Chiang, K. Lee, and Y.-C. Kan, “An activity recognition model using inertial sensor nodes in a wireless sensor network for frozen shoulder rehabilitation exercises,” *Sensors*, vol. 15, no. 1, pp. 2181–2204, 2015.

[192] M. N. Mohamed, *Hajj & Umrah from A to Z*. IslamKotob, 1996.

[193] P. Bota, J. Silva, D. Folgado, and H. Gamboa, “A semi-automatic annotation approach for human activity recognition,” *Sensors*, vol. 19, no. 3, p. 501, 2019.

[194] M. Stikic, D. Larlus, and B. Schiele, “Multi-graph based semi-supervised learning for activity recognition,” in *2009 international symposium on wearable computers*, pp. 85–92, IEEE, 2009.

[195] A. Islam, R. Aravind, T. Blascheck, A. Bezerianos, and P. Isenberg, “Preferencesand effectiveness of sleep data visualizations for smartwatches and fitness bands,” in *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, pp. 1–17, 2022.

[196] A. Godfrey, V. Hetherington, H. Shum, P. Bonato, N. Lovell, and S. Stuart, “From a to z: Wearable technology explained,” *Maturitas*, vol. 113, pp. 40–47, 2018.

[197] A. Henriksen, M. Haugen Mikalsen, A. Z. Woldaregay, M. Muzny, G. Hartvigsen, L. A. Hopstock, and S. Grimsgaard, “Using fitness trackers and smartwatches to measure physical activity in research: analysis of consumer wrist-worn wearables,” *Journal of medical Internet research*, vol. 20, no. 3, p. e110, 2018.

[198] A. A. Awsat, “Saudi arabia launches smart bracelet for this year’s hajj,” 2021. https://english.aawsat.com/home/article/3081476/saudi-

arabia-launches-smart-bracelet-year%E2%80%99s-hajj, Last accessed on 2023-3-25.

Generated using Postgraduate Thesis LaTeX Template, Version 1.03. Department of Computer Science and Engineering, Bangladesh University of Engineering and Technology, Dhaka, Bangladesh.

This thesis was generated on Friday 5th May, 2023 at 4:49pm.