IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 28 July 2023, accepted 23 August 2023, date of publication 1 September 2023, date of current version 12 September 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3311261

APPLIED RESEARCH

Leveraging a Smartwatch for Activity Recognition in Salat

ISHRAT JAHAN[1], NAJLA ABDULRAHMAN AL-NABHAN[2], JANNATUN NOOR[1,3], MASFIQUR RAHAMAN[1], AND A. B. M. ALIM AL ISLAM[1]

[1]Next-Generation Computing (NeC) Research Group, Department of Computer Science and Engineering, Bangladesh University of Engineering and Technology, Dhaka 1000, Bangladesh
[2]Department of Computer Science, King Saud University, Riyadh 11421, Saudi Arabia
[3]Computing for Sustainability and Social Good (C2SG) Research Group, School of Data and Sciences, BRAC University, Dhaka 1212, Bangladesh

Corresponding author: Ishrat Jahan (1017052011@grad.cse.buet.ac.bd)

This work was supported by the Deputyship for Research and Innovation, Ministry of Education, Saudi Arabia, under Project DRI-KSU-762.

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Ethics Committee, Bangladesh University of Engineering and Technology.

ABSTRACT Salat, the most important worship of Muslims and the second pillar of Islam, is an integral part of the Muslim community. Being a complex activity, Salat involves a series of steady and transitional activities to be performed in a specific sequence. On top of that, Salat has variations based on time, priority, school of thought, etc., making activity recognition in Salat more challenging. Existing research studies related to recognizing individual activities in Salat either demand capturing images by a camera or carrying a smartphone (sometimes in inconvenient places) while praying. 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. To address these gaps, in this study, we first assess the requirement and acceptability of technological solutions for activity recognition in Salat by conducting an exploratory study. Upon establishing the requirement, we propose an activity recognition methodology using a smartwatch to recognize different activities in Salat. We prepare a Salat activity dataset using a smartwatch and propose a new methodology using semantic rules and Dynamic Time Warping (DTW) that achieves a near-perfect accuracy (99.3%) in recognizing activities in Salat. Besides, our proposed methodology offers fine-grained recognition of the individual activities in Salat and is robust enough to overlook the extra transitional activities a person performs while praying, which does not nullify Salat. Therefore, this research is expected to lead to a comprehensive solution for monitoring Salat.

INDEX TERMS Complex activity recognition, DTW, Salat, smartwatch.

I. 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 [1]. Over the years, different sensing technologies are used in HAR studies such as body-worn sensors, cameras, wearable sensors, etc., to improve recognition performance. In recent times, wearable sensor devices have been widely used in HAR studies. 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., [2]. 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., [3], [4], as they are equipped with all the essential sensors necessary for activity recognition. Practical deployments of HAR have been done in fall detection [5], [6], behavioral monitoring [7], gait analysis [8],

The associate editor coordinating the review of this manuscript and approving it for publication was Dian Tjondronegoro.

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

[9], Ambient Assisted Living (AAL) [10], surveillance systems [11], sports coaching [12], and many more [2]. Thus activities under HAR research have become ever spreading.

Activities that have been covered by HAR research so far are huge in number, and diverse in nature from walking, running, biking, etc. [13], to kitchen activities [14], construction activities [15] and whatnot. Moreover, they vary greatly in complexity. However, simple activities are explored extensively whereas complex activity recognition is yet a less-explored area [16]. In this study, we focus on a particular human activity, which is Muslim prayer or Islamic prayer activity known as Salat.

A. SALAT - A COMPLEX RELIGIOUS ACTIVITY UNDER RECOGNITION

Second is the most fundamental daily worship of Muslims and the second pillar of Islam among the five pillars [17], [18]. 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. According to a survey [19] conducted by the Pew Research Center across 39 countries and territories, a majority of Muslims in the Middle East and North Africa, Southeast Asia, South Asia, and sub-Saharan Africa pray several times a day. Additionally, in 12 countries, three-quarters or more of the participants report performing all five prayers daily. This widespread practice of Salat reflects its profound significance, particularly among religious Muslims.

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 [17], [18]. The activities in Salat belong to both static/steady and transitional activities. Table 1 presents a list of these activities. The * marked transitions, though not regular, can happen from time to time, for example, when a person is praying too fast. Moreover, Salat exhibits variations based on time, priority, school of thought, etc. Accordingly, Salat, by its very nature, falls under the category of complex activities. Considering the importance of Salat among Muslims and its complexity, some HAR studies focused activity recognition in Salat [20], [21], [22].

B. 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. Among these studies, in many cases, the placements of smartphones lack convenience and are impractical for daily use [21], [22] or not applicable for all [23], [24], [25]. Salat, being a very frequent worship, the device to be used for its activity recognition, should be practical for daily use. 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.

TABLE 1. Different activities in Salat (The * marked ones are not standard transitions but can also be found in different scenarios.)

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) Bowing to short-standing to prostrating (B-Sh-P)* Prostrating to sitting to prostrating (P-S-P)* Standing to short-standing (St-Sh)*

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 [26]. The primary advantage is that people are accustomed to wearing watches and they can be worn for a long time, i.e., at home, at night, and during any kind of activity [26]. The wrist placement of smartwatches presents one of the least intrusive placements [27] for monitoring activities. Additionally, the battery life of smartwatches is more durable than smartphones [28], whereas they can combine almost all features of smartphones for data collection and even continuous monitoring [29]. Furthermore, the notifications are more easily observed than that through smartphones, due to their proximity to the user's line of sight [29]. Moreover, many studies in the literature made successful use of smartwatches for various activity recognition tasks [30], [31]. Therefore, considering all these, we find it worth attempting to recognize activities in Salat using smartwatches.

C. MOTIVATIONS BEHIND OUR STUDY

The vacancies in the literature regarding HAR in Salat motivate us to perform this study. First of all, the absence of usage of any natural sensing device for HAR in Salat in the literature [21], [25], [32], [33], [34] is our primary motivation. Next, recognition of Takbeer and differentiation between the postures of standing and short-standing could not be done previously due to using a smartphone. The size of the datasets used in the earlier studies is also a matter of concern as the majority cover only 10 subjects [21], [22], [25] or even less [24].

Moreover, none of the existing studies reported perfect or near-perfect accuracy, which is important for such sensitive usage. In fact, the adoption of cross-validation alone to measure the performance of the models of almost all the studies [22], [23], [24], raise serious concern as to whether they truly reflect the models' realistic performances [35]. Additionally, the accuracy of HAR solution for Salat might hamper if a person performs any extra activity in Salat, that does not nullify prayer [36] as shown in Figure 1. This issue is also not considered yet.

FIGURE 1. Examples of extra activities performed during Salat [37].

(a) Masjid-al-Haram, Mecca [38] (b) Masjid-an-Nabawi, Medina [39]

FIGURE 2. Muslims praying at different places during Hajj and Umrah times.

On top of all these, a HAR solution capable of recognizing activities in Salat can help Muslims assess the correctness and completeness of their prayers, find out mistakes, pinpoint the areas to focus more on, track progress, etc. All these can help a worshipper improve his prayer quality. Besides, in the case of individual Salat in crowds (for example, in Makkah or Madinah during Hajj/Umrah as shown in Figure 2, such solutions can assist in determining the mistake and removing the confusion and prevent the worshipper to stay some more time in the crowd. Therefore, the importance of such studies is immense. Being motivated by all these, in this study, we explore a new approach to activity recognition in Salat using a smartwatch.

D. PROBLEM FORMULATION AND RESEARCH CHALLENGES

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?

RQ2 Do people need technological assistance for improving their prayer? How willing are they to accept such technological assistance?

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?

In the process of answering the research questions, we envision some research challenges entailing our study. First of all, religion, being a sensitive and private topic, we need to explore the acceptability of such solutions among people as well as the permissibility from scholars. Next, we have to deal with the higher variability in the sensed signals of the smartwatch due to its placement on the wrist [40]. Besides, to ensure robustness, we have to handle the variations of Salat and extra activities that do not nullify prayer. Moreover, as Salat is a very important worship for Muslims, achieving near-perfect accuracy has no alternative. In doing so, detecting and fixing the wrong predictions of the classifier using domain knowledge in a post-processing step perhaps has no alternative which presents a noteworthy challenge.

E. 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 conducting an online survey. afterwards, to develop a new methodology for HAR in Salat using a smartwatch, we first prepare our own dataset using Samsung Galaxy Watch Active 2 [41]. Then, we preprocess the raw data and separate the steady and transitional states using two alternative approaches - 1) using traditional machine learning classifiers such as Random Forest (RF), etc., and 2) using Signal Magnitude Area (SMA) [42]. Subsequently, we detect some of the steady states using semantic rules derived from domain knowledge and use Dynamic Time Warping (DTW) for the transitional states. Finally, we apply some postprocessing on the predictions to further enhance the accuracy. In this process, we achieve up to 99.3% overall accuracy. 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 convenient and non-distracting wearable 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.
- 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.

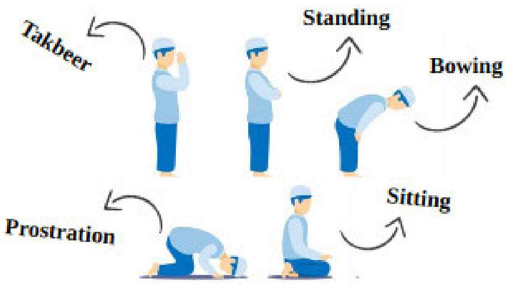


FIGURE 3. **Basic steps of Salat [43].**

F. ORGANIZATION OF THE STUDY

Our paper is further segmented into different sections. Section II contains some backgrounds and preliminaries. Next, Section III is about literature reviews of recent papers related to our study. After that, in Section IV, we present our exploratory study i.e., survey and interviews to establish the acceptability and permissibility of such assistance for Salat. Next, our proposed methodology is presented in Section V. Later, Section VI contains experimental evaluation and findings. Furthermore, in Section VII, we present the discussion and comparative analysis of our proposed methodology with the existing literature. Finally, the conclusion and future prospects of this research are stated in Section VIII respectively.

II. BACKGROUNDS AND PRELIMINARIES

In this Section, we describe some concepts used in our study covering Salat, its variations, and potential mistakes in Salat.

A. ACTIVITY UNDER STUDY: SALAT

Salat is 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) [17], [18]. 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" consisting 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 3 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 [16], [44]. 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 recite the first chapter from the Qur'an. In some prayers, this is followed by the recitation of (part of) some other chapter. Then a person bends down keeping his hands on the knees as shown in Figure 3 which is called Ruku or bowing. After that, the person goes up and stands straight with hands on both sides. 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. In each Rakah, a person has to prostrate twice and sits down for a short time in between. 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 the left reciting a specific supplication.

1) 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 [17], [18], [45]. For example, one category is obligatory or Fard, which is compulsory for Muslims to offer within a prescribed time frame. Other categories are Sunnah, Nafil, Wajib, etc., varying in priority.

Secondly, the number of Rakah to offer varies based on the time of the prayer [17], [18], [45]. 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. Additionally, another prayer, called Witr, which needs to be prayed after the Ishaa, differs from other prayers as one has to give Takbeer after the recitation is done in the last Rakah 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 [46], [47]. Besides, male and female prayer vary in posture in some schools of thought, whereas others claim prayers of both genders to be the same [17], [47], [48]. Even if two persons belong to the same school of thought, prayer postures can still vary. For example, while raising hands for Takbeer, some raise their hands up until their ear, some raise up until their neck, and so on. Differences also happen due to the lack of knowledge of the standard ways of performing Salat. Postures vary due to differences in capability too. For example, if someone cannot offer prayer in standing due to any disability, he is allowed to offer the whole prayer through sitting.

2) 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, differences within the units, or lack of knowledge, people tend to make mistakes [17], [49]. Moreover, when 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.

III. RELATED WORK

We situate our research work in a body of related studies exploring the recognition of various human activities. In the following, we go through some relevant topics and shed some light on the literature.

A. 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.

1) 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 attempted to be recognized using different techniques and sensors [50], [51]. Many research studies focused on Activities in Daily Life (ADL) such as walking, running, jogging, etc., [52], [53], [54]. Other studies focused on transitional activities too [55], [56]. The study in [57] attempted exercise activities whereas kitchen activities were recognized by the study in [14]. On the other hand, the study in [58] recognized ten different dance micro steps. Another study [15] explored construction activities using the accelerometer data. Moreover, recently fall detection has also gained much interest because of its vast application in healthcare [5], [6]. These studies present a few examples of the wide variety of simple activities covered by the HAR researchers.

2) COMPLEX ACTIVITIES

Complex activities compile a series of multiple actions [16], [44] in a concurrent, interleaved, or overlapping manner. Examples include playing a game, cooking, cleaning, buying, etc. They demonstrate realistic representations of people's daily lives [16].

The study in [50], attempted to recognize simple activities as well as some complex activities such as cooking, cleaning, etc., through a smartphone and found that the performance of recognizing complex activities appears to be poor (50%) than simple ones (93%). Another research study [59] presented a machine learning approach to correctly classify highly-correlated and imbalanced nursing activities. Besides, the study in [44] proposed an algorithm capable of mining temporal patterns from low-level actions to represent high-level human activities. Another study [60], proposed a model to recognize and classify complex at-home activities through wearable sensing leveraging selective multi-modal sensor suites from wearable devices. Additionally, the study in [16], built a dictionary of time series patterns, called shapelets, to represent atomic activities and used shapelet-based models to recognize sequential, concurrent, and generic complex activities. However, to the best of our knowledge, complex activities are still less explored and remains challenging enough to recognize compared to simple activities [16], [44].

B. DATA COLLECTION IN HAR

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.

1) 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 [61], fraud detection [62], extraction of information from videos [61], video assistance and surveillance [61], and public security [63], etc. There also exist applications in surgical operations [64] and classifying static signs of the sign language [65]. However, due to the requirement of image processing, this approach is computationally more expensive. Besides, privacy and security issues entail this approach, and accuracy is greatly affected by occlusion, change in illumination or background, etc., [44].

2) SENSOR-BASED DATA COLLECTION

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

Smartphones have been extensively studied for recognizing different physical activities in recent years [4] due to their wide availability and equipment 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 [66], 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.). Besides, common human activities [52], transportation activities [67], gait analysis [8], [9] have also been explored extensively using smartphones. Some studies, in this regard, focused on energy efficiency [68] while others [69], [70] focused on online recognition of activities.

With the recent emergence of smartwatches, HAR research arguably achieved a new dimension. Studies conducted with both smartwatches and smartphones found that smartwatches can achieve superior performance for a wider range of activities than smartphones [30]. Besides, as explored in a smartwatch-based study [31], 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 [71] explored obtaining keyboard usage information of a laptop using sensors of the Samsung Galaxy Live smartwatch. In addition to that, the study in [72], classified eight different daily human activities with a Moto 360 smartwatch, using PCA and Random Forest.

C. 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.

1) 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. The most widely-used classifiers in HAR are decision-tree classifiers such as J48 Decision tree [73], Random Forest (RF) [74], etc. They have been used in numerous HAR studies [75], [76], [77] yielding impressive accuracy. Many studies [78], [79] used Naïve-Bayes [80] , while others [81], [82], [83] used Support Vector Machine (SVM). Besides, Instance-based Learning (IBL) algorithms, especially KNN is used in a number of HAR studies [84], [85], [85], [86].

Some studies [66] also investigated the potential of ensemble classifiers which are generally computationally more expensive. On the other hand, several studies [52], [81] compared different classification techniques using a combination of time and frequency domain features.

2) 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. Convolutional Neural Networks (CNN) [87] and Recurrent Neural Networks (RNN) [88] are the two most popular deep learning models in this regard. Different research studies [89], [90] 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 which is adopted by many HAR studies for the purpose of complex activity recognition [91], [92]. However, recently the LSTMs [93], with their capability of memorizing and modeling the long-term dependency in the supplied data, have taken a dominant role in the HAR domain [94], [95]. Nevertheless, in recent times, hybrid deep learning models combining both CNNs and RNNs are also explored for activity recognition tasks [92], [96].

3) HAR USING DYNAMIC TIME WARPING (DTW)

Dynamic Time Warping (DTW) [97], though extensively applied in speech recognition, has also been proven effective in HAR research. Many studies bypassed the complex feature engineering step of ML classifiers and used template selection approaches such as DTW instead for HAR. For example, the study in [98], modified DTW to improve computational efficiency and similarity measure accuracy. Another study [99] proposes a new ensemble classifier based on DTW and uses combined information from multiple time-series sensors to map them with corresponding activities. On the other hand, the study in [100], classifies light sport exercise activities using the accelerometer sensor on a smartphone and smartwatch that is placed on the left hand of the user using KNN and DTW. Another research applied DTW to process different shapes of foot movements captured using wearable sensors [101]. Besides, the study in [102], performs HAR for six different human activities by exploring time-phased data and the signal magnitude of an on-body creeping wave.

D. HAR RESEARCH RELATED TO SALAT

In the literature, we find a handful of HAR research studies that focus on recognizing the activities in Salat. The first study targeting automatic recognition of prayer movements was conducted in 2009 [103] using a computer vision-based approach. Later, the study in [20], investigated motion tracking for Salat activity recognition leveraging two Kinect devices and using Hidden Markov Model.

The study in [32] used deep learning for the first time for HAR in Salat and built an image dataset for the basic Salat positions. Besides, the most recent computer vision-based study on Salat [104], proposed an assistive intelligent framework to evaluate the correctness of prayer postures. However, this approach is computationally expensive and requires a good video recording facility, and sufficient lighting with no occlusion and background change. This might also cause distraction and pose serious privacy concerns (especially for Muslim women who observe modesty) which makes this approach less suitable in this regard.

The first sensor-based approach for Salat activity recognition was proposed in 2016 [21] using smartphone accelerometer. However, the suggested placement of the smartphone (at the upper back of the user), manual cleaning of data, etc., are some other limitations of this study. Another study [25] used 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 which is Salat while placing the smartphone in the shirt's pocket. On the other hand, the study in [33] developed a pattern for the whole prayer and used DTW to decide

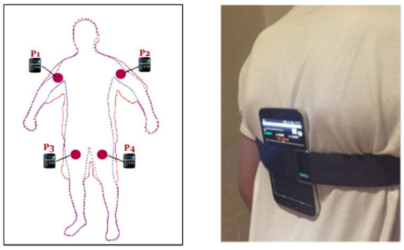


FIGURE 4. Suggested smartphone placement positions in [21] and [22].

whether the whole activity is a prayer pattern or not. 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 [33] nor [34] attempt to recognize different steps and activities within the prayer. A similar study was conducted in [23] using mobile accelerometer data and performance comparison was done among three feature extraction approaches and eight machine learning classifiers placing the smartphone on the hip area. Another research study [22], utilized a smartphone's accelerometer to help Alzheimer's patients in their prayers using a bunch of machine learning classifiers. They found the upper arm placement of the smartphone yielding better accuracy which is not a convenient position at all. Lastly, the study in [24], recognized six steps in prayer using three smartphone sensors placing the smartphone in the pant's pocket. Therefore, we see that the placement of smartphones is inconvenient or impractical in these studies (Figure 4). Even in the case of placing it in pockets, a concern is that not all garments have pockets or pockets of a specific size or a specific position. Therefore, the natural usage of sensing devices is missing in these studies.

IV. 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 is completely anonymous and the participants voluntarily fill it out. We also conduct interviews with some of the participants who fill out the survey to dig more about the mistakes. Besides, we interview some Islamic scholars to confirm the permissibility of such solutions in Islam. The following subsections contain details on each of these phases.

A. 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 [105], [106], [107], [108], judging computer efficacy [109], assessing social desirability [109], [110], measuring personality [109], [111], exploring digital well-being [112], criminology [113], psychopathology assessment [114], assessing openness in research content sharing [115], investigating psychological disorders [116], [117], etc., have already utilized self-reporting based surveys. Moreover, the research communities on HCI [115], CSCW [112], and ubiquitous computing [116], [117] often leverage self-reporting-based surveys. Accordingly, in our study, we utilize a self-reporting-based survey.

B. ETHICAL APPROVAL AND INFORMED CONSENT

We confirm that prior to conducting the survey, we have obtained ethical approval from our university's Institutional Review Board (IRB) in accordance with established guidelines and regulations for research involving human participants. The survey is conducted online through a secure platform allowing us to present the participants with comprehensive information on the purpose of this research study at the outset. We explicitly assure them about the confidentiality and anonymity of their responses. Additionally, we ensure the voluntary nature of their involvement and convey to them that their valuable time and effort spent in this regard would contribute towards better technology planning and design through this research study. Only those participants who are comfortable with these assurances voluntarily choose to take part in the survey. Thus, the informed consent process is an integral part of our study design.

C. OVERVIEW OF OUR QUESTIONNAIRE

We collect demographic information at the beginning of the survey. 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 confusion during prayers. We take responses for these questions in a 5-point Likert scale [118].

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. 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. The questionnaire items of our survey is presented in Table 2.

D. SURVEY PARTICIPANT DEMOGRAPHY

The sampling strategies used in our survey are convenience sampling [119], [120], referral sampling [119], [120], and list-based sampling [119], [120]. 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

TABLE 2. Table of survey questionnaire items.

Section Name	Question	Response format
Demography	1. What is your year of birth?	Multiple choice
	2. What is your gender?	Multiple choice
	3. What is the highest level of education or degree you have achieved?	Multiple choice
	4. What is the country of your nationality? (In case of having multiple nationalities, please mention the country of your first nationality)	Multiple choice
	5. In which country are you living now?	Multiple choice
	6. What is your occupation?	Multiple choice
	7. What is your present religious or spiritual identity? (Religious identities are presented in alphabetic order)	Multiple choice
Your Usage and Ownership of Technological Device(s)	8. Do you have experience(s) of using the following devices? [Smartphone]	Binary choice
	9. Do you have experience(s) of using the following devices? [Smartwatch]	Binary choice
	10. Do you currently own the following devices? [Smartphone]	Binary choice
	11. Do you currently own the following devices? [Smartwatch]	Binary choice
Regarding Your Prayer	12. I hardly miss my prayers - to what extent do you agree with this statement?	5-point Likert
	13. I am willing to improve my prayer quality and quantity - to what extent do you agree with this statement?	5-point Likert
Your Frequency of Confusions during Prayer	14. How often do you forget or get confused about the following aspects of your prayer? [Count of Rakat (Units)]	5-point Likert
	15. How often do you forget or get confused about the following aspects of your prayer? [Count of Ruku (Bowing)]	5-point Likert
	16. How often do you forget or get confused about the following aspects of your prayer? [Count of Sujud (Prostration)]	5-point Likert
	17. How often do you forget or get confused about the following aspects of your prayer? [Sitting for Tashahhud]	5-point Likert
	18. How often do you forget or get confused about the following aspects of your prayer? [Reciting another Surah after Surah Fatihah]	5-point Likert
Your Opinion on Tech Assistance for Improving Prayer	19. Would you welcome if your devices such as smartphone, smartwatch, etc., assist you to improve your prayer quality?	5-point Likert
	20. Can you please let us know the reason(s) behind your response to the last question?	Open-ended
	21. If you need to pray to wear a smartwatch or fitness band while praying to avail the above-mentioned features, then how willing will you be to do that? (Please assume that you will have the wearable with you)	5-point Likert
	22. If you are not willing to use any wearable as mentioned in the previous question, then can you please tell us the reason behind it?	Open-ended
Concluding Section	23. Please feel free to share with us any idea(s)/suggestion(s)/ critique(s) about this survey.	Open-ended

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.

We get responses from 126 participants in total who are from 15 different countries. Among them, two participants reported that they are not religious or spiritual. Therefore, we had to 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. In Islam, Salat, along with all other religious obligations, and accountability in general, begin at puberty [121], [122], [123]. Therefore, it appears safe to go without having this specific population from our survey. As per occupation, our participants cover students, teachers, IT professionals, engineers, homemakers, etc. Table 3 presents the demography of our participants.

E. SURVEY DATA ANALYSIS AND FINDINGS

We analyze the quantitative survey data 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 the Chi-squared test [124]. Additionally, we use the Mann-Whitney U test [125] and Kruskal-Wallis test [126] to compare whether there is any statistically significant difference in the dependent variable for the independent groups. For better understanding, the dependent variables of our analyses along with their corresponding independent variables are presented in Table 4.

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. However, 43% of the survey participants have experience using a smartwatch, and currently, 31% of them own one.

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 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 5 depicts these findings.

TABLE 3. 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

TABLE 4. Table of dependent variables with corresponding independent variables used in statistical analysis (DF = Demographic factors, mean_mist_freq = Mean mistake frequency, regu_in_pr = Regularity in prayer, will_to_imp_pr = Willingness to improve prayer, welcm_tech_salat = Welcome technological assistance in Salat, will_to_wear = Willingness to wear a wearable).

Dependent variable	Independent variable
mean_mist_freq	DF, regu_in_pr, will_to_imp_pr
welcm_tech_salat	DF, regu_in_pr, will_to_imp_pr, mean_mist_freq
will_to_wear	DF, regu_in_pr, will_to_imp_pr, mean_mist_freq

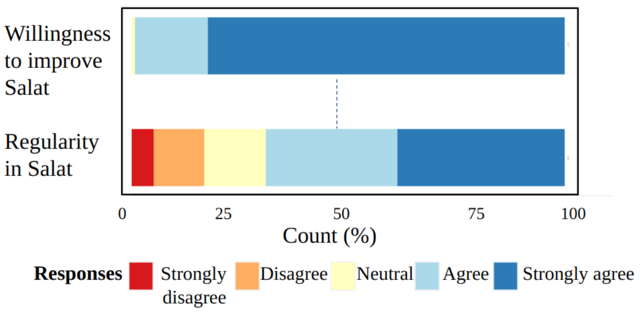


FIGURE 5. Summary of participants' regularity in prayer and willingness to improve prayer.

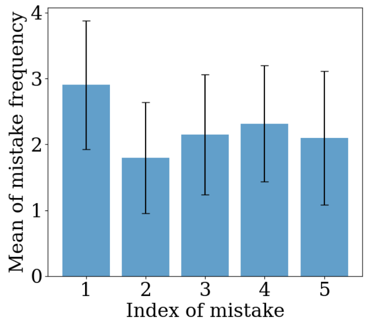


FIGURE 6. Summary of the frequency of different mistakes in Salat (1 = forgetting count of Rakat, 2 = forgetting count of bowing, 3 = forgetting count of prostration, 4 = forgetting to sit for Tashahhud, and 5 = forgetting to recite another Surah after Surah Fatiha).

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 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 further assure that the choice of the levels of the Likert scale is supported by the existing literature [127], [128].

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 the 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.

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 7a 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 of 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 7b presents the summary of the responses of the participants to this question.

In addition to that, we have a few qualitative questions. The free-text responses to these questions help us to have a deeper understanding of the users' perspectives and underlying

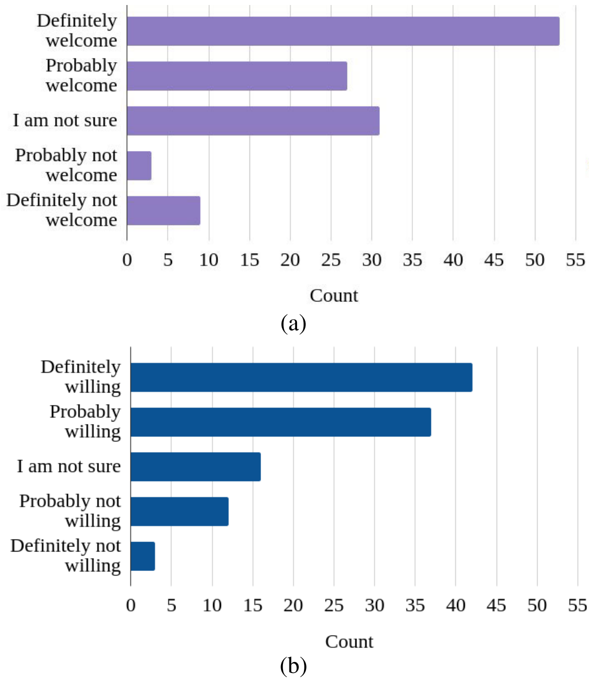


FIGURE 7. Summary of the responses to the questions regarding (a) eagerness to explore technological assistance in Salat and (b) willingness to pray to wear a convenient wearable (smartwatch, fitness band, etc).

reasonings behind their opinion. In line with the quantitative analysis, the qualitative responses also reflect the keenness of the participants to explore technologies to improve Salat. As per their responses, the primary reason behind their interest is the fact that technological devices have already become an indispensable part of today's life, and therefore, if they are capable of providing any good regarding Salat, according to them, people should embrace it. Many express their hope that success in this work would benefit the Muslims greatly. In this regard, one 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 a regular basis. So, it will help me to track my improvement easily."* (P7)

However, few of the participants, think that the Islamic guidance regarding the mistakes in Salat is enough. Some participants also express their confusion regarding the permissibility of such solutions. Besides, quite a few participants emphasize that technological solutions should not cause disturbance during i.e., no alert or so inside prayer. For example, one participant expresses his concern as follows.

> *"I am interested, but I would want to make sure it is permissible to use a device to assist my prayer."* (P87)
>
> *"If the devices can assist properly, no problem; but in times of praying, I don't want the natural environment to be harmed."* (P29)

TABLE 5. Background of the scholars interviewed.

	Education	Profession
Scholar-1	Teacher at a central mosque	Dawrae Hadith
Scholar-2	Teacher at an Islamic school (Jamia[1])	Dawrae Hadith, Ifta, M.Sc. in Islamic Studies
Scholar-3	Teacher at an Islamic school (Jamia[1])	Dawrae Hadith, Ifta

F. FINDINGS FROM INTERVIEW

To learn more about the mistakes, we take interviews of 11 people who have already participated in our survey. We ask them whether there are other mistakes apart from the five they were asked about in the survey, that they make regularly or have seen others making. Also whether they do any random movements or random mistakes in Salat or have seen others doing so. In response, three of our participants report that they sometimes forget to make the extra Takbeer followed by a specific supplication in Witr Salat. Some of the interviewees mention that they forget some of the supplications to be recited at specific points of Salat sometimes. Finally, some report that they make mistakes while performing some of the unconventional prayers such as Eid prayer, funeral prayer, etc., as the patterns of those prayers vary from regular prayers. No other mistakes are reported by any of them.

G. EXPLORING PERMISSIBILITY OF TECHNOLOGICAL ASSISTANCE IN SALAT

We explore the permissibility of adopting technological assistance in Salat in Islam to address the concern raised by our survey participants. Subsequently, we engage with three esteemed scholars of Bangladesh, interviewing them to delve into our topic in detail. Their short educational and professional background is presented in Table 5. We first provide them with an overview of our study and obtain their consent to utilize their opinions as references in our research. Next, we mention the potential applications of our study and asked them about the ruling of availing such technological assistance in Salat. All of them permit the usage of any technological assistance outside of prayer i.e., getting details of the prayer upon completion or getting feedback to improve prayer quality, etc. Similarly, all of them express their concern regarding availing any assistance such as an alert for mistakes, etc., while in prayer. Below we quote the opinions of our scholars.

> *"There is no harm in using such technologies after Salat to check details or get feedback. I have also confirmed with two of my senior Muftis (scholars). However, having real-time alerts from such devices regarding mistakes is something that needs research from our end, and in my opinion, it is better not to use such assistance."* (Scholar-1)
>
> *"The idea is revolutionary! You know what, Muslims are struggling most with their Iman (faith) in this era of technology. What you guys are doing,*

[1] Jamia is equivalent to a university.

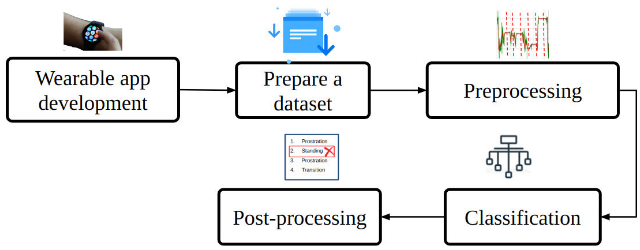


FIGURE 8. Pipeline of our proposed methodology.

is the best way to do ihsan (favor) to your fellow Muslim brothers. I really welcome the idea of monitoring prayer using a watch very much. However, you can generate alerts regarding mistakes inside Salat for beginners such as reverts or so but not for others. I pray that Allah makes you successful in your endeavors." (Scholar-2)

"*You cannot have any sort of assistance while a person is praying. Anything else is fine... If a person gets to know about a mistake through your device after completing his prayer, he can take corrective measures or repeat the prayer and there is no harm in it. And if your device monitors the prayer of an individual and provides necessary feedback to improve prayer then that is good. I mean, out of Salat, you can take any sort of help but not inside prayer.* " (Scholar-3)

V. PROPOSED METHODOLOGY

This Section 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 8 depicts the pipeline of our methodology from a high level.

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 predicting the final activity recognition results. We elaborate on each of these steps of operations in the following sections.

A. 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 [41] as shown in Figure 9. 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 9. Samsung galaxy watch Active 2 [129] - smartwatch.

This app is designed to record the sensor values from the accelerometer, gyroscope, and magnetometer pertaining to all three axes, with the timestamp, in files. Unfortunately, our app cannot record the magnetometer data. From our investigation, it seems that our watch model that runs on Tizen might not support the magnetic sensor [130], despite its inclusion in the official documentation. In our study, 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 [131].

B. 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 human subjects individually. We use convenience or opportunity sampling strategy for recruiting the subjects [119], [120]. All subjects have agreed to the usage of the recorded data for scientific and research purposes. 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.

We capture and record video of the whole procedure using a timestamp camera app to facilitate the task of ground-truth labeling. We explicitly take consent from the participants regarding the video capturing. Due to this video capturing, some participants express their discomfort in praying formally as they find it difficult to concentrate while being captured and therefore, want to mimic the prayer movements skipping the recitations. We take into consideration this discomfort and keep it open for all participants to either pray formally or mimic the prayer movements skipping the recitation part and instead remaining still in those positions. In total, 14 subjects (9 female and 5 male) out of 30 choose to mimic the prayer movements while others perform the full formal prayer as they do usually. This allows us to capture a broader range of prayer behaviors as we find that among our participants who prayed the full prayer, almost all of them prayed at a slower pace whereas in reality prayer speed varies based on many factors. The inclusion of the subjects who mimic prayer movements, helps us to get these variations in our dataset.

Figure 10 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.

TABLE 6. 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%)

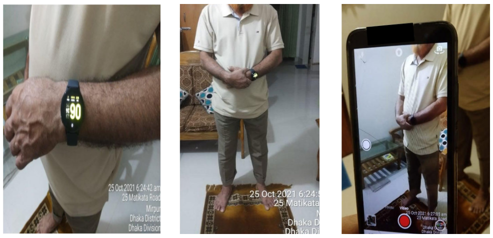


FIGURE 10. Snapshots captured during our data collection.

1) PARTICIPANT DEMOGRAPHY

All 30 subjects, from whom we collect data, are from Bangladesh and currently living here. The number of male and female participants are 17 and 13 respectively. Our study includes participants across a wide age range, spanning from 15 to 67 years old, thus ensuring diversity in the age demographic. The demography of the subjects is given in Table 6.

C. 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.

1) 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 [132]. Many primary research studies on activity recognition use this notion for the purpose of denoising and smoothening [21], [133], [134]. Figure 11 shows the denoising process using a moving average filter.

2) 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.

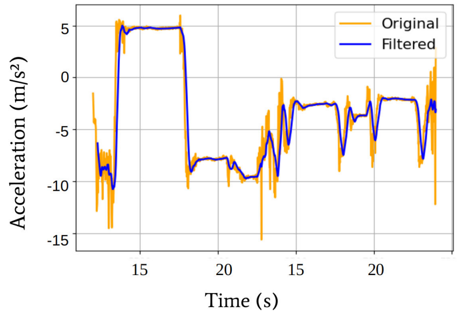


FIGURE 11. Comparison of the original (raw) and filtered data captured during our data collection.

Algorithm 1 Algorithm for Preprocessing Raw Signal

Require: D accelerometer data of x, y, and z axes with labels, n total number of the collected data
Ensure: 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 $end \neq 0$ **then**
      $start \leftarrow end - samplesPerWindow \times overlap$
    else
      $end \leftarrow start + samplesPerWindow$
    end if
    $s \leftarrow D[start : end]$
    $S.append(s)$
  end while

3) 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 [135]. There are three basic types of windowing used in HAR - activity-defined window segmentation, event-defined window segmentation, and sliding window segmentation [135]. 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 windows, 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 [136]. 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 [137]. 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 [138], [139]. In summary, the purpose of preprocessing is to get rid of unwanted noises and break the continuous data into fixed-sized windows to effectively capture the salient characteristics of the raw data. Algorithm 1 depicts our approach to preprocessing the signal. At the end of this stage, the signal is denoised and segmented into equal-sized chunks for further processing.

D. 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 [3], [140]. 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. 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.

1) 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.

a: 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* [141], which presents an important aspect of developing HAR systems [3]. The use of features rather than raw data generally enhances classification accuracy as reported in the literature [141]. 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. Features extracted for classical machine learning classifiers.

Domain	Features
Time	Mean, Max, Median, Standard deviation, Variance, Skewness, Kurtosis, RMS, Inter-quartile range, AUC (Area under the curve), MAD (Mean absolute deviation), Number of peaks, Absolute energy, Peak-to-peak distance, Zero-crossing rate, and Entropy
Frequency (FFT)	Mean, Median, Standard deviation, Inter-quartile range, and Power
Others	Pairwise correlation between all 3 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 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.

b: 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 [74], J48 decision tree [73], Naive Bayes [80], and Logistic regression [142]. We select these classifiers considering their high accuracies in the existing HAR studies [75], [76], [78], [79].

c: 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 [137]. 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 of 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 [143].

The model adopted from the study in [143] learns to automatically disentangle domain-agnostic and domain-specific features. Here, the domain stands for a specific person's data. To effectively disentangle these two latent spaces, the study claims to develop a novel Independent Excitation mechanism. As stated in this study [143], the novelty lies in the fact

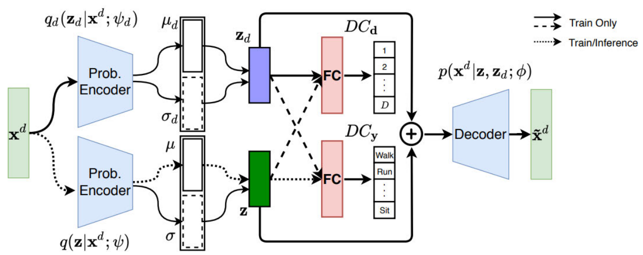


FIGURE 12. Workflow diagram of the deep learning model (GILE) [143] used in our experiment.

that this mechanism, inspired by the independent optical excitation of distinct neural populations [144], minimizes the correlations between domain-agnostic and domain-specific features. As a result, the model can generalize better to new unseen target domains. In this manner, the model proves itself to be robust for cross-person or user-independent HAR. For better visualization, the workflow diagram of the model is presented in Figure 12. This model undergoes experimentation with three benchmark datasets [145], [146], [147] 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.

d: 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. 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 [148]. 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, rather group them together into a single class 'Transition'. In the second approach (Approach-2), we recognize all the steady and transitional activities mentioned in Table 1 as different individual classes. A summary of the two approaches are presented in Table 8.

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 [42], [134], [149]. 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 [42], [134], [149]. 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.

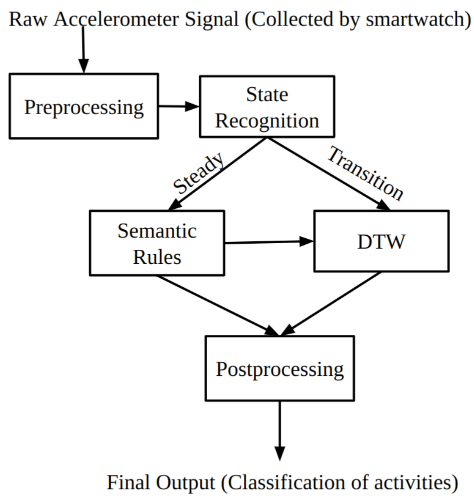


FIGURE 13. Pipeline of our proposed hierarchical methodology.

2) CLASSIFICATION WITH IMPROVED METHODOLOGY USING SEMANTIC RULES AND DTW

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 section. 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 [150]. 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 [150], [151]. On the other hand, sometimes different action classes have similar patterns, which is related to inter-class variation [150]. 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 of 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, [148]. To recognize the transitions with better accuracy, a viable alternative to ML-based approaches can be template-matching [2]. 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 [152]. We find many prior research studies adopting template matching [14], [100], [152], [153], [154], especially for transitions [153], [154]. 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 [152]. In fact, the study in [155] observes a higher generalization ability of the template-based methods compared to several ML classifiers

TABLE 8. Different approaches of classification using machine learning.

	Description	# of classes	Classes
Approach-1	All transitions are grouped into a single class 'Transition'	7	5 steady states (St, B, Sh, P, S), Takbeer (Tk), and Transition (T)
Approach-2	All transitions are recognized individually	12	5 steady states (St, B, Sh, P, S), Takbeer (Tk), and all transitional states (St-B, B-Sh, Sh-P, P-S, S-P, P-St)

while classifying activity data collected with a wrist-worn accelerometer. Therefore, in our proposed methodology, we adopt Dynamic Time Warping (DTW) [156], 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 [155]. 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 13 presents the pipeline of our proposed improved methodology integrating all these stages.

a: 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 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 [42], [131], [134], [149]. 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) [42], [131]. 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 1 [131].

$$SMA = \frac{1}{t}\left(\int_0^t |x(t)|dt + \int_0^t |y(t)|dt + \int_0^t |z(t)|dt\right) \quad (1)$$

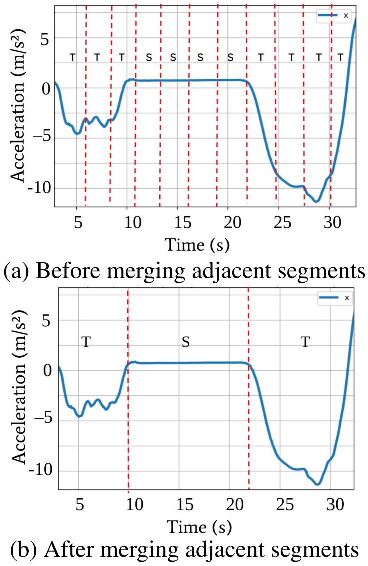


FIGURE 14. Impact of merging adjacent segments based on prediction labels (S = Steady and T = Transition).

Here, $x(t)$, $y(t)$, and $z(t)$ refer to the body components of the x, y, and z-axes of the 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 [157]. 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 [131]. 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 [158]. Figure 14 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.

FIGURE 15. The three axes of a smartwatch.

(a) Standing (b) Bowing

FIGURE 16. The placements of hands in (a) standing and (b) bowing positions.

b: 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 of an activity [151]. 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 [151]. 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 15. 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 16. 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 15, 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 17a and Figure 17c. Here, Figure 17 shows the summary of the real values collected from 30 subjects. Besides, as the acceleration towards the 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 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.

Algorithm 2 Algorithm for Recognizing Steady States Through Semantic Rules

Require: S Signal segments labeled as Steady or Transition by the state recognizers
Ensure: L new labels array
  $count \leftarrow 0$
  $L \leftarrow []$
  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**
        $L[count] \leftarrow Bowing$
        $ps \leftarrow$ previous Steady segment of s
        if $(ps.x - ps.y).mean > 0$ **then**
          $L[count - 2] \leftarrow Standing$
        else
          $L[count - 2] \leftarrow ShortStanding$
        end if
      else
        $L[count] \leftarrow Unknown$
      end if
    else
      $L[count] \leftarrow Transition$
    end if
    $count = count + 1$
  end for

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 the 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. Among the five steady states, we have semantic rules to recognize two steady states. Therefore, the other three steady states, which we could not be recognized in this stage, are marked as 'Unknown'. They are going to be recognized in the later stages of our methodology. The corresponding algorithm is presented in Algorithm 2.

A special case to be considered here is that some worshippers tend to swing their hands while in the short-standing position instead of standing still. In our dataset, we also record this phenomenon as we found two of our male subjects (P5 and P8) doing so. We present the corresponding signal of two worshippers where the first one swings his hands while in a short-standing position and the other one does not and prays in a standard way in Figure 18. This particular deviation of posture from the standard way does not affect our recognition of bowing and standing using semantic rules. Figure 18 depicts how in both cases the previous stage, i.e., state recognition stage segments and labels the signal. We see that in the non-ideal case, the short-standing state, due to hand movement, could not be separated or marked as a steady state. Everything else remains the same in both cases. More importantly, we can verify from the figures that our assumption about the values and relationships of the axes discussed above in the standing and bowing phases hold in both scenarios. Therefore, this non-standard posture in short-standing does not pose any anomaly in this stage and our intended task of recognizing bowing and standing using semantic rules can still be carried on seamlessly.

c: CLASSIFICATION OF TRANSITIONS USING DYNAMIC TIME WARPING (DTW)

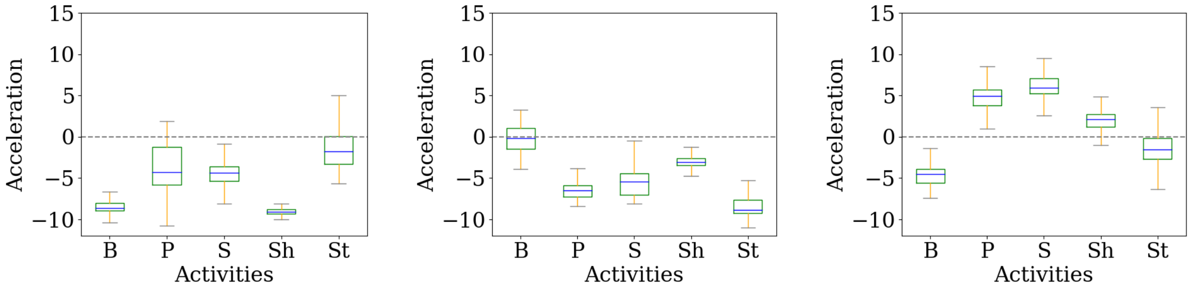
In this stage of our proposed methodology, we aim to classify the transitions that occur between the steady states. This classification, in turn, helps us to infer the steady states. We use Dynamic Time Warping (DTW) [156], more specifically a variant of DTW called FastDTW [97], for classifying the transitions. Dynamic time warping (DTW) [156] 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 19 presents the matching technique of DTW in contrast to Euclidean matching.

A well-known application of DTW has been in automatic speech recognition, to cope with different speaking speeds [159], [160], [161]. Besides, it is used in partial shape-matching applications [162]. However, we also find this algorithm being adopted in HAR research [14], [100], [101], [155], 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.

For the classification of transitional activities, we first create a template database with templates for each of the transitions. We begin the template database construction process by extracting the signal portions corresponding to transitional activities for all subjects. Subsequently, for each transition class, we retain patterns that exhibit the highest dissimilarity. As for the similar ones, we retain only one or two representative templates of that cluster by subjective filtering and discard the others. This process helps to capture the nuances and variations in transition signals among different subjects and covers a broad range of transition patterns maintaining a balanced representation of transitional activities in the template database.

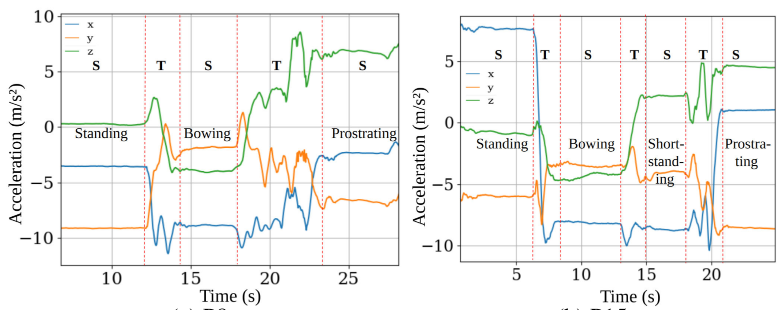
To demonstrate the process, we present an example in Figure 20. Among the 3 templates presented in this figure for the class Sh-P, we select one subjectively from the first two due to their high similarity and keep the third one as it is much more dissimilar than the other two. Among the three irregular transitions mentioned in Table 1, B-Sh-P and P-S-P are included as we find in our dataset that, many subjects do not spend much time in the short-standing phase before bowing and in the sitting phase in between two prostrations. Besides, some worshippers tend to swing their hands while in the short-standing phase as shown in Figure 18. Therefore, the short-standing in B-Sh-P and the sitting in P-S-P can not be recognized individually as steady states for not having enough time spent or unusual hand movement. 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 21 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



(a) x-axis (b) y-axis (c) z-axis

FIGURE 17. 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).



(a) P8 (b) P15

FIGURE 18. Comparison of the signal of 1-unit prayer of two subjects P8 and P15 where (a) P8 swings his hand in the short-standing position and (b) P15 does not and prays in a standard way. The red dashed lines indicate the signal segmentation and predicted labels (S = Steady, T = Transition) by the state recognition stage.

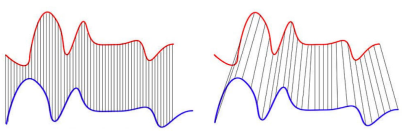

(a) Euclidiean match (b) DTW match

FIGURE 19. A comparison of Euclidean and DTW matching.

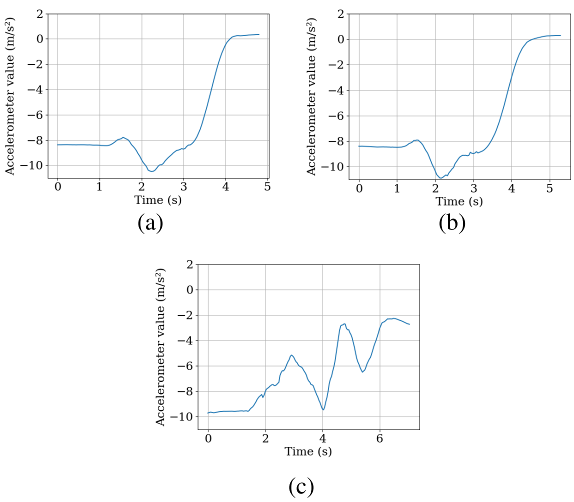
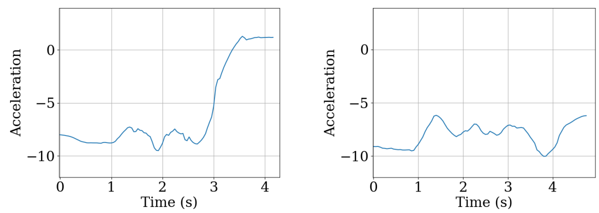


FIGURE 20. Template selection.



(a) Male (b) Female

FIGURE 21. Comparison of two representative templates of Sh-P from (a) Male and (b) Female template databases.

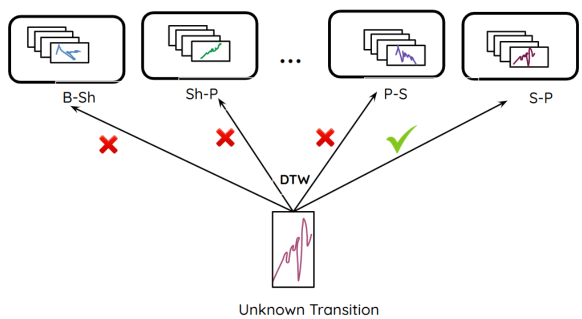


FIGURE 22. Classification of an unknown transition using DTW.

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 21. 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.

$$d(X, T_k) = \frac{1}{N} \sum_{i=1}^{i=N} DTW(X, T_i^k) \tag{2}$$

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.

$$d = \min_{k=1..s} d(X, T_k) \tag{3}$$

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. By 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 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.

d: 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 by making necessary corrections. This type of postprocessing is also found in the existing HAR literature [53], [163]. We describe our postprocessing techniques below in detail.

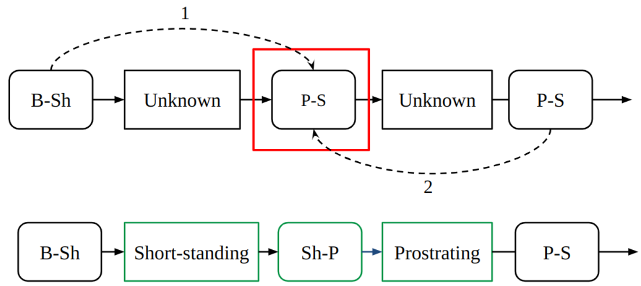


FIGURE 23. An example of fixing misclassified transition using neighboring transitions.

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 23, 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.

Similarly, some misclassified transitions can be fixed by considering the duration of 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

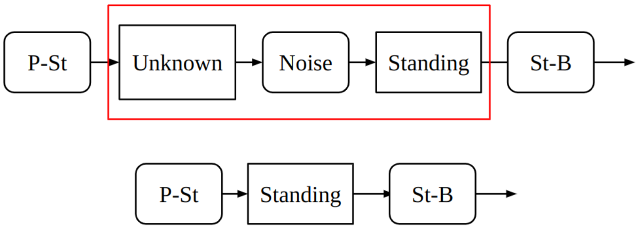


FIGURE 24. An example of ignoring extra movements.

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.

Lastly, Null activities are also removed in this stage. The Null activities or extra movements, 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 24 presents such a scenario of detecting and eliminating extra movements.

The full algorithm of postprocessing is presented in Algorithm 3. Here, we first remove the Null activities or extra movements. Then, we apply our postprocessing logic stated earlier and fix the predictions accordingly. Based on our observations, we find Takbeer to be mostly confused with P-St. Therefore, we have handled it exclusively. Besides, we detect transitions with abnormal durations and fix them accordingly. Lastly, 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.

E. VALIDATION PROTOCOL AND EVALUATION METRICS

We use both k-fold cross-validation [164] and Leave-One-Subject-Out (LOSO) [1] 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 [1]. Definitions of the evaluation metrics are given below.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \tag{4}$$

Algorithm 3 Algorithm for Postprocessing

Require: S Signal segments with labels predicted by state recognizer, semantic rules, and DTW
Ensure: Final prediction labels

```
for each segment s in S do
    if s is a Transition and label of s is Noise then
        Combine s with the previous and the next steady state of s
        if any of the steady states' label is known then
            Update label of the combined segment according to the known steady state
        else
            Label the combined segment as Unknown
        end if
    end if
end for
for each segment s in S do
    if s is a Transition then
        if label of s in L is P-St and previous steady state of s is not Prostrating or Sitting then
            update label of s as Takbeer
        end if
        if s is a Transition and label of s is B-Sh-P and duration of s < threshold duration of B-Sh-P then
            update label of s as B-Sh
        end if
        if s is a Transition and label of s is P-S or S-P and duration of s > threshold duration of P-S or S-P then
            update label of s as P-S-P
        end if
    end if
end for
for each segment s in S do
    if s is a Transition then
        if label of s is not consistent with the previous transitional segment's label then
            update label according to the previous and next transitional state
        end if
        if the previous steady segment's label is Unknown then
            update label of the previous steady segment according to the label of s and its previous transitional segment's label
        end if
    end if
end for
```

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{5}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{6}$$

$$\text{F1-score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{7}$$

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.

VI. EXPERIMENTAL EVALUATION

In this Section, we present the experimental evaluation of the baseline methodology as well as our proposed methodology presented in Section V, 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.

A. ADOPTION OF APPROPRIATE SENSOR

The literature suggests evaluating each sensor individually offline to determine its contribution to the overall process [165]. Therefore, we do some initial experimentation with both accelerometer and gyroscope data collected by our app to assess their contribution. Following our baseline methodology using machine learning, the cross-validation accuracy we get for detecting the steady states using an accelerometer is pretty impressive (above 90%). Details of experimentation using an accelerometer are provided in the following subsections. In contrast, the gyroscope does not detect any change in angular velocity when a person is still and the value of the gyroscope in all 3 axes is around 0 in all steady states of Salat. For this reason, unlike accelerometers, the recognition of steady states is not possible with the gyroscope. Therefore, we aim to recognize the transitional activities using gyroscope data following our baseline methodology, which yields only 44.63% accuracy which is too poor. Consequently, considering the lower accuracy and the inability of the gyroscope data to contribute significantly to recognizing steady states, we decide not to include the gyroscope data in our further experimentation.

B. 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 25 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.

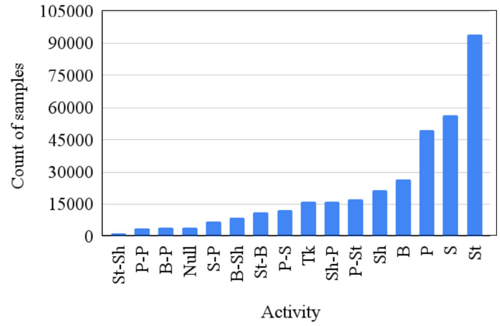


FIGURE 25. Distribution of activities in our dataset.

TABLE 9. 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

As mentioned in the study of [51], in the context of human activity recognition, the diversity 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 9. We expect that the diversity in each of these four demographic factors covers a wide range of populations.

C. RESULTS AND FINDINGS

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

1) 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 [74], J48 decision tree [73], Naive Bayes [80], and Logistic regression [142]. 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 V-D1.d 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 10 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 [143], which is the deep learning model we used in

TABLE 10. 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

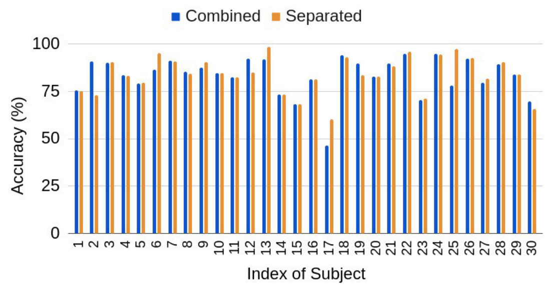
Classifier	Single classifier				Hierarchical			
	Cross-validation	LOSO			Cross-validation	LOSO		
		Avg	Max	Min		Avg	Max	Min
RF	91.22	74.99	92.91	54.32	92.49	78.93	94.80	55.38
LR	85.12	75.77	90.01	65.63	87.48	77.5	90.31	66.78
NB	75.24	66.12	43.20	30.08	76.32	43.90	68.44	31.14
J48	86.54	73.35	58.39	58.09	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

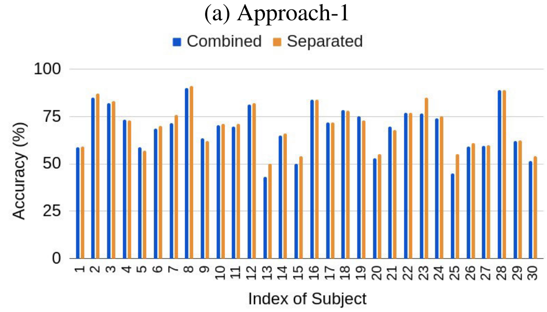
our baseline methodology, elaborated in Section V-D1.c, the sliding window length is 1.28 sec with 50% overlap. For both types of classifiers, these values give us the best performances tuned up through our experiments.

From the results shown in Table 10, 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 a hierarchical fashion. Besides, accuracy is much higher in Approach-1. This is expected, as in Approach-2, we classify the transitions and 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 a hierarchical fashion. The reason behind this finding is 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 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 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 [48]) 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 26 shows a comparison in this regard for both approaches.



(a) Approach-1



(b) Approach-2

FIGURE 26. LOSO accuracy of GILE when datasets from male and female subjects are considered in combined and separated manners of Approach-1 and Approach-2 in the hierarchical fashion.

2) LIMITATIONS OF MACHINE LEARNING CLASSIFIERS

It is clear from Table 10, 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 both approaches in the hierarchical fashion in Figure 27. It is evident from Figure 27 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 28.

From the confusion matrices in Figure 28, we find that bowing and standing activities are less confused by the

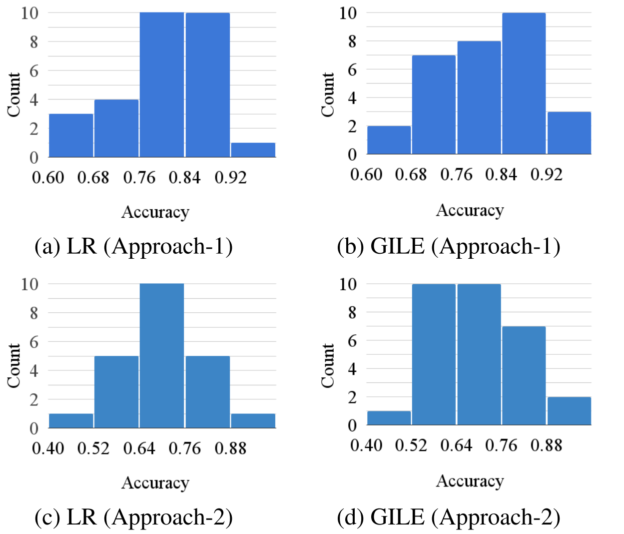


FIGURE 27. Histogram of LOSO accuracies obtained by (a) LR and (b) GILE for Approach-1 and Approach-2 in the hierarchical fashion.

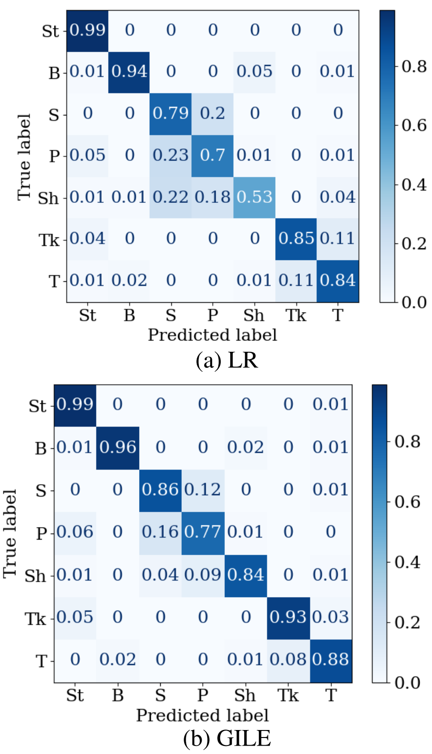


FIGURE 28. Confusion matrices obtained by (a) LR and (b) GILE with Approach-1 in the hierarchical fashion (St = Standing, B = Bowing, S = Sitting, P = Prostrating, Sh = Short-standing, Tk = Takbeer, and T = Transition.)

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 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

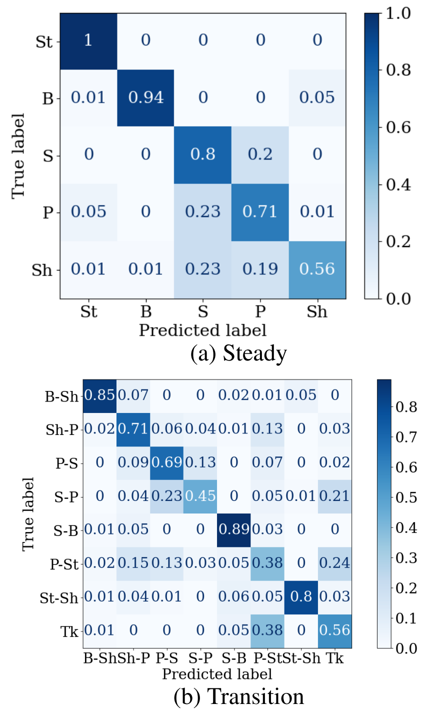


FIGURE 29. Confusion matrices of (a) steady and (b) transitions obtained by LR with Approach-2 in the hierarchical fashion.

the activities of one Rakat predicted by GILE as shown in Table 11.

We can see from Table 11 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 by 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 10, 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 29 shows the confusion matrices of steady states and transitional states obtained by LR with Approach-2.

Such degraded accuracy for short-lived activity recognition is not new in the literature. We find in the existing HAR

TABLE 11. 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	✓	✓	✓

literature that, the traditional features of short-lived samples are unstable and cannot describe the actions effectively [70], [152]. 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 mis-classifications. This scenario eventually leads us to attempt to design an improved methodology for better prediction of activities in Salat through better error handling.

D. PERFORMANCE ANALYSES OF THE PROPOSED METHODOLOGY

Below we present the performance analyses of our improved methodology.

1) RESULTS OF STATE RECOGNITION

In our proposed methodology, we first recognize steady and transitional states. We have already mentioned earlier in Section V-D2.a, 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 [157] of the literature. We adopt these parameters to obtain the linear acceleration component from the acceleration signals. Subsequently, we focus on identifying the optimal threshold value for normalized SMA. We calculate the accuracy of state recognition across the entire dataset while adjusting the normalized SMA value within the range of 0.2 to 2. Through this iterative process, we discovered that the highest accuracy was achieved when the normalized SMA value was set to 1. Therefore, we set this as the optimal threshold for our analysis. Figure 30 depicts our findings of tuning normalized SMA value. This implies that if the normalized 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 31.

After predicting each segment, we merge contiguous blocks of similar states. Though we see in Figure 31 that the classification into steady and transition accuracies fluctuate around 90%, after merging the adjacent similar segments,

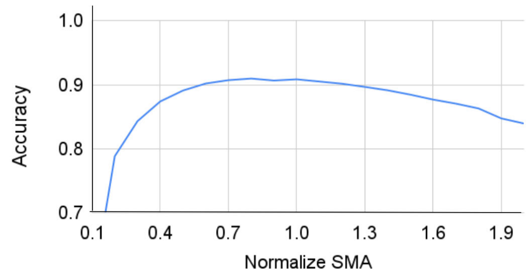


FIGURE 30. Accuracy variation with different normalized SMA values.

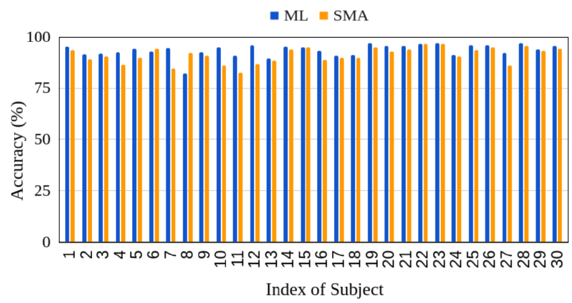


FIGURE 31. Performance comparison of two state recognition methods.

TABLE 12. Accuracy of state recognition using ML varying sliding window size and moving average filter window size.

Window Size	Moving average filter window size	Accuracy (%)
0.5	5	0.8999
	15	0.9012
	30	0.8891
1	5	0.8931
	15	0.9044
	30	0.8832
1.5	5	0.8691
	15	0.9061
	30	0.8789

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. Table 12 summarizes our findings regarding the state recognition accuracy achieved through varying sliding window sizes and moving average filter window sizes.

2) RESULTS OF SEMANTIC RULE-BASED CLASSIFICATION FOR STEADY STATE RECOGNITION

We apply the semantic rules that we derived from our domain knowledge, on the five steady states - standing, bowing, short-standing, prostrating, and sitting. We find that using these

TABLE 13. 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

rules, bowing and standing can be accurately recognized with a remarkable 100% accuracy. As we cannot differentiate among the other three states using these rules, therefore, they are marked as 'Unknown' in this stage which is later classified using DTW in the later stage. Thus, the semantic rules serve as reliable indicators for the identification of these activities. Another important aspect is that the confidence gained from the accurate recognition of these two activities eventually helped optimize our subsequent phases also.

3) 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 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 13 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 V-D2.c, 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 14 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 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 V-D2.d, such confusion can be fixed using the duration of the transitions.

4) FINAL RESULTS AFTER POSTPROCESSING

From the confusion matrices in Table 14, 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 V-D2.d. 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 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 confusion 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 15.

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%.

a: RAKAT COUNTING ACCURACY

After the final prediction is made by our proposed methodology about each activity, we go through the predicted series of activities to count the number of Rakat. Ideally, as standing and bowing occur only once in a Rakah, therefore, the number of standing activity or bowing activities should be equal to the number of Rakat. In that sense, as our proposed methodology can recognize bowing and standing with 100% accuracy,

TABLE 14. Female.

	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

TABLE 15. 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

Algorithm 4 Algorithm for counting number of Rakat

Require: Ordered list of predicted steady-state activities S
Ensure: $rakatCount = 0$
  for each activity s in S **do**
    if s is Standing **then**
      $standingFound = True$
    else if (s is Sitting or s is Prostrating) and $standingFound = True$ **then**
      $rakatCount += 1$
      $standingFound = False$
    end if
  end for

therefore, the Rakat count accuracy is also 100% in ideal scenarios. However, we also consider the non-ideal cases i.e., where a person might make mistakes such as bowing multiple times in a single Rakah or not bowing at all, etc. Considering all these, we have devised an algorithm for Rakat counting which is presented in Algorithm 4. Basically, we check for a standing phase followed by a sitting phase to decide that a Rakat is found. That means, upon encountering a standing activity, we assume that a Rakat has just started and then we keep searching for sitting phase activities in the subsequent activities and once we find a sitting phase activity such as sitting or prostrating, we count that as a complete Rakat.

Therefore, it is clear that our Rakat counting accuracy is dependent on the accuracy of recognition of standing and sitting activities, i.e., sitting and prostrating. According to our final confusion matrix presented in Table 15, the recognition accuracy for the standing activity is 100%. Similarly, the sitting activities (prostrating and sitting) are recognized with near-perfect accuracy and occasionally get confused between sitting and prostrating but not with other activities. However, even if such confusion occurs, it does not impact our Rakat counting algorithm since both sitting and prostrating activities are treated equally in the algorithm. Consequently, our Rakat counting accuracy remains consistently high at 100%.

VII. DISCUSSION

In this Section, 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.

A. OUTCOMES OF THE EXPLORATION OF OUR RESEARCH QUESTIONS

In this Section, we will shed light on the outcomes of exploring our research questions set in Section I-D. We find the answer to RQ1 and RQ2 through our online survey. Through a mixed-method analysis of the responses of the participants, we find that more than one-third of our participants make at least one mistake frequently where forgetting the Rakat count is reported to be the most common mistake. This answers our RQ1. Next, we find that the majority of our participants (above 70%) exhibit a willingness to explore technological assistance to improve their Salat which answers RQ2.

And lastly, to answer RQ3, we explore recognizing activities in Salat using a smartwatch and develop a methodology in this regard. The impressive accuracy of our methodology proves that HAR in Salat using a smartwatch is feasible and surpasses the accuracy of other currently available methods.

B. ESTABLISHMENT OF REQUIREMENT, ACCEPTABILITY AND PERMISSIBILITY OF TECHNOLOGICAL ASSISTANCE IN SALAT

Before approaching the gaps in the literature regarding activity recognition in Salat, we survey the real users to assess the necessity, and acceptability of technological assistance among them. We do not find any research in the literature conducting any such study in this direction. However, in the literature, we find some other HAR studies [166], [167] doing this practice of assessing the acceptability of their HAR solutions which focus on different domains other than Salat. By conducting our survey, we have gained a more profound understanding of different facets relating to the need and willingness of the general population to embrace technological assistance in Salat. Considering the demography of the survey participants presented in Table 3, we can claim that this survey does not represent any specific community, but, rather represents the Muslim population in general. More importantly, our interviews with Islamic scholars clarify the concerns regarding the permissibility of such assistance and provide us with a firm ground to go further in this direction.

C. 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. This appears to be sufficient enough, as other benchmark datasets cover the number of subjects equal to us [146], [147] or less than that of ours [51]. Furthermore, ideally, a dataset should reflect the variability of real-world activities, and be flexible enough to emulate different experimental setups [145]. In a similar way, our dataset covers a diverse demography of subjects, as shown in Table 6. The preparation of such a dataset takes much time and manual effort for both collecting raw data and subsequent labeling. We believe that our dataset will serve as a basis for testing different HAR approaches for Salat and also for the evaluation of any complex activity recognition model in the future.

D. METHODOLOGICAL ADVANCEMENT FOR ACTIVITY RECOGNITION IN SALAT

In this study, we propose a new methodology to recognize the activities in Salat. None of the methodologies presented by the existing studies matches ours, as almost all of them use only a classical pattern recognition pipeline whereas we use the notion of DTW blended with semantic rules, postprocessing, etc. Though one of the earlier studies on Salat [33] uses DTW, it only attempts to detect whether the signal pattern is a prayer pattern or not. Moreover, in the case of other existing studies [60], [168], [169], [170], [171], the recognition of activities or methodologies under consideration are either not applicable for Salat [168] or are much more complex compared to ours one [16], [44]. Besides, though we find several HAR studies in the literature using DTW [14], [100], [101], 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.

E. RECOGNITION OF A COMPLEX ACTIVITY WITH NEAR-PERFECT ACCURACY

First of all, in the literature, we find that complex activities are comparatively less explored and challenging to recognize [50], [59]. Salat, being a complex activity, is not an exception here, as recognizing each individual step in Salat is undoubtedly a challenging task. Our method is specifically devised for Salat keeping its specific characteristics in consideration yielding a near-perfect accuracy (99.3%) in recognizing individual steps or activities in Salat. This outperforms all the existing studies found to date to the best of our knowledge. Table 16 presents a comparison of the performances and other aspects of our model as well as related existing studies [21], [22], [23], [24], [25].

F. 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 machine learning approaches. This happens as, 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. Thus, The LOSO type of validation is highly important, however, absent in

TABLE 16. 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., [21] (2016)	-	Smartphone	X (Upper-back)	X	Separate training and testing data	91.0
Eskaf et al., [25] (2016)	10	Smartphone	✓ (Shirt's pocket)	X	Cross-validation	94.6
Obaid et al., [23] (2018)	20	Smartphone	✓ (Pocket)	X	Cross-validation	93.0
Ahmad et al., [22] (2019)	10	Smartphone	X (Upper-arm)	X	Cross-validation	97.5
Topu et al., [24] (2021)	8	Smartphone	✓ (Pocket)	X	Cross-validation	93.9
Our study	30	Smartwatch	✓ (Wrist)	✓	LOSO	99.3

the existing studies of HAR in Salat. Therefore, in our case, the usage of LOSO gives us a better understanding of the performance of our proposed methodology in the real world. Besides, 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.

G. FINE-GRAINED RECOGNITION WITH TOLERANCE TO EXTRA ACTIVITIES

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 [17], [18]. As this only includes the movement of the head, its impact on the smartwatch (placed on the wrist) is not significant enough to recognize. All other steps, irrespective of whether it is steady or transitional, can be recognized by our model. We recognize Takbeer and short-standing which is not recognized by the earlier studies. 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. Unlike the earlier studies, our proposed methodology takes this into consideration, and no matter how quickly these steps are performed, can be recognized by our model.

Besides, we are the first to 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 [36]. However, as the smartwatch rests on the wrist, and the extra activities are mostly done by hand, they result in some extra transitions in the signal which can cause wrong predictions. In this regard, we leverage DTW to eliminate the extra activities in the post-processing stage. This also presents the necessity of post-processing in the process of producing a complete and final prediction.

H. USAGE OF A CONVENIENT WEARABLE FOR HAR IN SALAT

To the best of our knowledge, we are the first to utilize a smartwatch for HAR in Salat. As for their respective usage, both smartphones and smartwatches are convenient for their own purposes. However, in 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. Our findings from the survey also support this understanding. Considering all these, we can claim that, our endeavor to use smartwatches for HAR in Salat is a valuable addition to the literature.

I. 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 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.

Besides, it is worth mentioning that this study covers recognizing the activities in Salat, and identifying different mistakes in Salat is its logical consequence such as mistakes include forgetting to perform a specific activity (e.g., a Rakah), performing a specific activity more than the prescribed times, etc. This happens as mistakes related to the count of various activities in Salat can be captured based on the activities recognized. However, there also exist other types of mistakes that cannot be identified in this way such as wrong recitation, wrong direction of facing, wrong postures such as placing hands on the chest instead of the belly, etc, which are beyond the scope of this study.

One important point is that our proposed methodology, in its present form, is not intended for real-time recognition of Salat activities. It is because, we take the whole prayer signal and then separate the steady and transitional states, and apply various logics and algorithms to recognize the states. Again, this can be extended to work in real-time too. Nonetheless, the applications of our proposed methodology are not confined only to providing real-time feedback to the worshippers. Rather, the most useful application could be helping worshippers monitor their prayers, getting valuable insights regarding their prayers that can help them in improving their prayers, etc.

Accordingly, this research opens up possibilities for the development of innovative apps that can provide a comprehensive analysis of one's prayer such as the counts of performed activities, durations of prayer and each activity in prayer, sequence of activities, number of Rakat, etc. In this way, an app can be developed to highlight areas for improvement in the prayer as well as track progress and offer feedback in accordance with Islamic teachings. Users could gain a deeper understanding of their prayer habits, including missed prayers, delayed prayers, rushed prayers, common mistakes in the prayer, and time spent in individual activities of the prayer. Armed with this information, users can take proactive steps toward improving the quality and consistency of their prayer. These insights can be delivered to the worshiper instantly or after monitoring their prayer patterns over a specific duration. Besides, the worshiper can also be reminded about his mistakes upon completing the prayer so that s/he can take corrective actions, which the scholars generally permit as stated in Section IV-G.

In addition to general worshipers, such assistance could be particularly valuable for beginners or new Muslims (converted), who often face challenges in accessing proper guidance, especially residing in remote areas. Obviously, they can learn how to pray from online resources, however, through such technology, they can monitor their prayers and get feedback for improvement in the absence of a guide.

Additionally, leveraging our proposed methodology, future solutions can be made to facilitate a distraction-free environment for prayers by automatically switching all devices to Silent mode upon detecting ongoing prayer activities. Furthermore, an auto-adjusting alarm feature can be integrated to keep notifying or warning the users for prayers. This can be especially helpful for Fajr or dawn prayer as many people find it difficult to wake up for the Fajr prayer.

Therefore, though Islam has already provided "Sahu Sijda" [17], [18] as a means of correction upon making specific mistakes, we assert that our study captures a broader aspect beyond correction of mistakes by aiming at helping Muslims in improving their prayer quality.

J. CONTRIBUTION TO THE HAR LITERATURE

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 explore. For example, the study in [172] 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.

VIII. CONCLUSION AND FUTURE WORKS

In today's world, HAR solutions are being leveraged widely for assisting people in numerous fields [10], [11], [12]. Salat, being the most fundamental worship of the Muslim community as well as a complex activity by definition [16], [44], 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 and 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 to find out the requirement, acceptability, and permissibility of such HAR solutions. Subsequently, we prepare a smartwatch-sensed dataset of offering Salat and propose a new HAR methodology blending machine learning algorithms, semantic rules, DTW, and custom post-processing in a delicate manner. Rigorous user-independent experimentation reveals that our proposed methodology outperforms all the previous studies, achieving a near-perfect accuracy (99.3%)

Being motivated by the potential of this study, we wish to take this to the next step, which is real-life deployment i.e., building an app with features outlined in Section VII-I. Additionally, apps built on our proposed methodology can also act as a prayer tracker i.e., detect whether a person has prayed or not. Besides, this can be extended to work in real-time as per the applicable rules in Islam.

Moreover, this study can be improved by including more subjects from people of different schools of thought or from people, who, due to some disability or so, cannot perform Salat in a conventional way, etc. In addition to that, another possible future work could be to work with low-resource and low-cost wearables as the one we use is a high-resource one. Lastly, our survey reveals the eagerness of people to be helped in their worship, and therefore, 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 [173], and providing their accurate counts to the pilgrims can be a potential direction of future work.

ACKNOWLEDGMENT

(Ishrat Jahan and Najla Abdulrahman Al-Nabhan contributed equally to this work.)

REFERENCES

[1] 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," 2018, *arXiv:1806.05226*.

[2] L. M. Dang, K. Min, H. Wang, M. J. Piran, C. H. Lee, and H. Moon, "Sensor-based and vision-based human activity recognition: A comprehensive survey," *Pattern Recognit.*, vol. 108, Dec. 2020, Art. no. 107561.

[3] A. O. Ige and M. H. M. Noor, "A survey on unsupervised learning for wearable sensor-based activity recognition," *Appl. Soft Comput.*, vol. 127, Sep. 2022, Art. no. 109363.

[4] A. Ferrari, D. Micucci, M. Mobilio, and P. Napoletano, "Trends in human activity recognition using smartphones," *J. Reliable Intell. Environ.*, vol. 7, no. 3, pp. 189–213, Sep. 2021.

[5] 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.

[6] C. Rougier, J. Meunier, A. St-Arnaud, and J. Rousseau, "Robust video surveillance for fall detection based on human shape deformation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 21, no. 5, pp. 611–622, May 2011.

[7] L. Chen and C. D. Nugent, *Human Activity Recognition and Behaviour Analysis*. Berlin, Germany: Springer, 2019.

[8] 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," *Pers. Ubiquitous Comput.*, vol. 15, no. 7, pp. 667–678, Oct. 2011.

[9] T. Iso and K. Yamazaki, "Gait analyzer based on a cell phone with a single three-axis accelerometer," in *Proc. 8th Conf. Human-Comput. Interact. Mobile Devices Services*, Sep. 2006, pp. 141–144.

[10] 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, Jan. 2021.

[11] A. Taha, H. H. Zayed, M. E. Khalifa, and E.-S.-M. El-Horbaty, "Human activity recognition for surveillance applications," in *Proc. 7th Int. Conf. Inf. Technol.*, May 2015, pp. 577–586.

[12] 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 *Proc. 23th Int. Conf. Archit. Comput. Syst.*, Feb. 2010, pp. 1–10.

[13] M. Shoaib, S. Bosch, O. Incel, H. Scholten, and P. Havinga, "A survey of online activity recognition using mobile phones," *Sensors*, vol. 15, no. 1, pp. 2059–2085, Jan. 2015.

[14] C. Pham and P. Olivier, "Slice&Dice: Recognizing food preparation activities using embedded accelerometers," in *Ambient Intelligence*. Berlin, Germany: Springer, 2009, pp. 34–43.

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

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

[17] M. H. Katz, *Prayer in Islamic Thought and Practice*, vol. 6. Cambridge, U.K.: Cambridge Univ. Press, 2013.

[18] Y. Y. Haddad and J. I. Smith, *The Oxford Handbook of American Islam*. Oxford, U.K.: Oxford Handbooks, 2014.

[19] Pew Research Center. (2012). *The World's Muslims: Unity and Diversity*. [Online]. Available: https://www.pewresearch.org/religion/2012/08/09/the-worlds-muslims-unity-and-diversity-2-religious-commitment/

[20] N. A. Jaafar, N. A. Ismail, and Y. A. Yusoff, "An investigation of motion tracking for solat movement with dual sensor approach," *ARPN J. Eng. Appl. Sci.*, vol. 10, no. 23, pp. 17981–17986, 2015.

[21] R. Al-Ghannam and H. Al-Dossari, "Prayer activity monitoring and recognition using acceleration features with mobile phone," *Arabian J. Sci. Eng.*, vol. 41, no. 12, pp. 4967–4979, Dec. 2016.

[22] 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, Aug. 2019.

[23] O. Alobaid and K. Rasheed, "Prayer activity recognition using an accelerometer sensor," in *Proc. Int. Conf. Artif. Intell. (ICAI)*, 2018, pp. 271–277.

[24] T. A. Topu, Md. M. Rahman, Md. S. Hossain, and A. Al Marouf, "Prayer activity recognition using smartphone," in *Proc. 12th Int. Conf. Comput. Commun. Netw. Technol. (ICCCNT)*, Jul. 2021, pp. 1–6.

[25] K. Eskaf, W. M. Aly, and A. Aly, "Aggregated activity recognition using smart devices," in *Proc. 3rd Int. Conf. Soft Comput. Mach. Intell. (ISCMI)*, Nov. 2016, pp. 214–218.

[26] G. Bieber, M. Haescher, and M. Vahl, "Sensor requirements for activity recognition on smart watches," in *Proc. 6th Int. Conf. Pervasive Technol. Rel. Assistive Environments*, May 2013, pp. 1–6.

[27] L. Porzi, S. Messelodi, C. M. Modena, and E. Ricci, "A smart watch-based gesture recognition system for assisting people with visual impairments," in *Proc. 3rd ACM Int. Workshop Interact. Multimedia Mobile Portable Devices*, Oct. 2013, pp. 19–24.

[28] G. Bieber, T. Kirste, and B. Urban, "Ambient interaction by smart watches," in *Proc. 5th Int. Conf. Pervasive Technol. Rel. Assistive Environments*, Jun. 2012, pp. 1–6.

[29] 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," presented at the Annu. Univ. Texas McCombs Healthcare Symp., Austin, TX, USA, 2016.

[30] G. M. Weiss, J. L. Timko, C. M. Gallagher, K. Yoneda, and A. J. Schreiber, "Smartwatch-based activity recognition: A machine learning approach," in *Proc. IEEE-EMBS Int. Conf. Biomed. Health Informat. (BHI)*, Feb. 2016, pp. 426–429.

[31] S. Bhattacharya and N. D. Lane, "From smart to deep: Robust activity recognition on smartwatches using deep learning," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun. Workshops (PerCom Workshops)*, Mar. 2016, pp. 1–6.

[32] A. Koubaa, A. Ammar, B. Benjdira, A. Al-Hadid, B. Kawaf, S. A. Al-Yahri, A. Babiker, K. Assaf, and M. Ba Ras, "Activity monitoring of Islamic prayer (Salat) postures using deep learning," in *Proc. 6th Conf. Data Sci. Mach. Learn. Appl. (CDMA)*, Mar. 2020, pp. 106–111.

[33] M. Ali, M. Shafi, U. Farooq, and A. u. Rehman Khan, "Salat activity recognition using smartphone triaxial accelerometer," in *Proc. 5th Int. Multi-Topic ICT Conf. (IMTIC)*, Apr. 2018, pp. 1–7.

[34] A. Muaremi, J. Seiter, F. Gravenhorst, A. Bexheti, B. Arnrich, and G. Troester, "Monitor pilgrims: Prayer activity recognition using wearable sensors," in *Proc. 8th Int. Conf. Body Area Netw.*, 2013, pp. 161–164.

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

[36] Fataawa 'Ulama' Al-Balad Al Haraam. (2011). *Moving Whilst Praying*. Accessed: Feb. 27, 2023. [Online]. Available: https://islamqa.info/en/answers/12683/moving-whilst-praying

[37] Tilawat Khashiah, *Emotional Recitation By Mishary Rashid Al Afasy—Surah Hud*, 2016.

[38] AskIslamPedia. *Maqam e Ibrahim*. Accessed: Apr. 2, 2023. [Online]. Available: https://www.askislampedia.com/en/wiki/-/wiki/English_wiki/Maqam+E+Ibrahim

[39] *Prime Minister Muhammad Shehbaz Sharif Offering Prayer in Riaz ul Jannah, Masjid Al-Nabwi SAW*, AP Pakistan. Accessed: Apr. 2, 2023. [Online]. Available: https://www.app.com.pk/photos-section/prime-minister-muhammad-shehbaz-sharif-offering-prayer-in-riaz-ul-jannah-masjid-al-nabwi-saw/

[40] R. San-Segundo, H. Blunck, J. Moreno-Pimentel, A. Stisen, and M. Gil-Martín, "Robust human activity recognition using smartwatches and smartphones," *Eng. Appl. Artif. Intell.*, vol. 72, pp. 190–202, Jun. 2018.

[41] *Samsung Galaxy Watch Active 2 Review: The Best Smartwatch for Android*. Accessed: Mar. 12, 2023. [Online]. Available: https://www.theguardian.com/technology/2020/may/13/samsung-galaxy-watch-active-2-review-best-smartwatch-for-android

[42] 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 Recognit. Lett.*, vol. 29, no. 16, pp. 2213–2220, Dec. 2008.

[43] *Muslim Man Praying Position Step Guide Instructions Symbol, Islam Religious Activity Icon Set in Flat Illustration Vector Isolated*, Dreamstime, Brentwood, TN, USA, 2023.

[44] Y. Liu, L. Nie, L. Liu, and D. S. Rosenblum, "From action to activity: Sensor-based activity recognition," *Neurocomputing*, vol. 181, pp. 108–115, Mar. 2016.

[45] M. Al-Tusi, *A Concise Description of Islamic Law and Legal Opinions*. U.K.: ICAS Press, 2008.

[46] M. J. Mughniyya. (1997). *Prayer (Salat), According to the Five Islamic Schools of Law*. Accessed: Jan. 27, 2023. [Online]. Available: https://www.al-islam.org/prayer-salat-according-five-islamic-schools-law-muhammad-jawad-mughniyya

[47] A.-H. Murad. (1997). *Understanding the Four Madhhabs*. Accessed: Jan. 27, 2023. [Online]. Available: http://www.masud.co.uk/ISLAM/ahm/newmadhh.htm

[48] B. D. Ifta. *Difference Between Men and Women Prayers*. Accessed: Mar. 8, 2023. [Online]. Available: https://islamqa.org/hanafi/daruliftaa-birmingham/19580/difference-between-men-and-women-prayers/

[49] (2014). *Mistakes in Prayer*. Accessed: Feb. 27, 2023. [Online]. Available: https://islamqa.info/en/answers/173637/mistakes-in-the-prayer

[50] S. Dernbach, B. Das, N. C. Krishnan, B. L. Thomas, and D. J. Cook, "Simple and complex activity recognition through smart phones," in *Proc. 8th Int. Conf. Intell. Environments*, Jun. 2012, pp. 214–221.

[51] M. Zhang and A. A. Sawchuk, "USC-HAD: A daily activity dataset for ubiquitous activity recognition using wearable sensors," in *Proc. ACM Conf. Ubiquitous Comput.*, Sep. 2012, pp. 1036–1043.

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

[53] M. J. Mathie, B. G. Celler, N. H. Lovell, and A. C. F. Coster, "Classification of basic daily movements using a triaxial accelerometer," *Med. Biol. Eng. Comput.*, vol. 42, no. 5, pp. 679–687, Sep. 2004.

[54] A. Salarian, H. Russmann, F. J. G. Vingerhoets, P. R. Burkhard, and K. Aminian, "Ambulatory monitoring of physical activities in patients with Parkinson's disease," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 12, pp. 2296–2299, Dec. 2007.

[55] M. M. Hassan, M. Z. Uddin, A. Mohamed, and A. Almogren, "A robust human activity recognition system using smartphone sensors and deep learning," *Future Gener. Comput. Syst.*, vol. 81, pp. 307–313, Apr. 2018.

[56] 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, Dec. 2015.

[57] 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 *Proc. 11th Annu. Int. Conf. Mobile Syst., Appl., Services*, Jun. 2013, pp. 361–374.

[58] A. Z. M. Faridee, S. R. Ramamurthy, H. M. S. Hossain, and N. Roy, "HappyFeet: Recognizing and assessing dance on the floor," in *Proc. 19th Int. Workshop Mobile Comput. Syst. Appl.*, Feb. 2018, pp. 49–54.

[59] F. R. Sayem, M. M. Sheikh, and M. A. R. Ahad, "Feature-based method for nurse care complex activity recognition from accelerometer sensor," in *Proc. Adjunct Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput. Proc. ACM Int. Symp. Wearable Comput.*, Sep. 2021, pp. 446–451.

[60] P. Bharti, D. De, S. Chellappan, and S. K. Das, "HuMAn: Complex activity recognition with multi-modal multi-positional body sensing," *IEEE Trans. Mobile Comput.*, vol. 18, no. 4, pp. 857–870, Apr. 2019.

[61] A. A. Afiq, M. A. Zakariya, M. N. Saad, A. A. Nurfarzana, M. H. M. Khir, A. F. Fadzil, A. Jale, W. Gunawan, Z. A. A. Izuddin, and M. Faizari, "A review on classifying abnormal behavior in crowd scene," *J. Vis. Commun. Image Represent.*, vol. 58, pp. 285–303, Jan. 2019.

[62] U. A. Akansha, M. Shailendra, and N. Singh, "Analytical review on video-based human activity recognition," in *Proc. 3rd Int. Conf. Comput. Sustain. Global Develop. (INDIACom)*, Mar. 2016, pp. 3839–3844.

[63] J. Shao, K. Kang, C. C. Loy, and X. Wang, "Deeply learned attributes for crowded scene understanding," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 4657–4666.

[64] J. P. M. Vital, D. R. Faria, G. Dias, M. S. Couceiro, F. Coutinho, and N. M. F. Ferreira, "Combining discriminative spatiotemporal features for daily life activity recognition using wearable motion sensing suit," *Pattern Anal. Appl.*, vol. 20, no. 4, pp. 1179–1194, Nov. 2017.

[65] 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 Comput. Appl.*, vol. 28, no. 11, pp. 3285–3294, 2017.

[66] Ó. 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," *Pervas. Mobile Comput.*, vol. 8, no. 5, pp. 717–729, Oct. 2012.

[67] 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, Feb. 2014.

[68] Z. Yan, V. Subbaraju, D. Chakraborty, A. Misra, and K. Aberer, "Energy-efficient continuous activity recognition on mobile phones: An activity-adaptive approach," in *Proc. 16th Int. Symp. Wearable Comput.*, Jun. 2012, pp. 17–24.

[69] M. Köse, Ö. D. Incel, and C. Ersoy, "Performance evaluation of classification methods for online activity recognition on smart phones," in *Proc. 20th Signal Process. Commun. Appl. Conf. (SIU)*, Apr. 2012, pp. 1–4.

[70] 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, Jan. 2016.

[71] H. Wang, T. T.-T. Lai, and R. Roy Choudhury, "MoLe: Motion leaks through smartwatch sensors," in *Proc. 21st Annu. Int. Conf. Mobile Comput. Netw.*, Sep. 2015, pp. 155–166.

[72] 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," *Meas. Control*, vol. 52, pp. 37–45, Jan. 2019.

[73] J. R. Quinlan, *C4. 5: Programs for Machine Learning*. Amsterdam, The Netherlands: Elsevier, 2014.

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

[75] H. Martín, A. M. Bernardos, J. Iglesias, and J. R. Casar, "Activity logging using lightweight classification techniques in mobile devices," *Pers. Ubiquitous Comput.*, vol. 17, no. 4, pp. 675–695, Apr. 2013.

[76] 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 *Proc. Int. Conf. Netw. Netw. Appl. (NaNA)*, Dec. 2020, pp. 310–315.

[77] L. Xu, W. Yang, Y. Cao, and Q. Li, "Human activity recognition based on random forests," in *Proc. 13th Int. Conf. Natural Comput., Fuzzy Syst. Knowl. Discovery (ICNC-FSKD)*, Jul. 2017, pp. 548–553.

[78] X. Yang, A. Dinh, and L. Chen, "Implementation of a wearerable real-time system for physical activity recognition based on naive Bayes classifier," in *Proc. Int. Conf. Bioinf. Biomed. Technol.*, Apr. 2010, pp. 101–105.

[79] A. Jiménez and F. Seco, "Multi-event naive Bayes classifier for activity recognition in the UCAmI cup," in *Proc. MDPI*, Oct. 2018, pp. 1–6.

[80] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, vol. 10. Burlington, MA, USA: Morgan Kaufinann, 2006, pp. 559–569.

[81] N. Ravi, N. Dandekar, P. Mysore, and M. L. Littman, "Activity recognition from accelerometer data," in *Proc. AAAI*, vol. 5, Pittsburgh, PA, USA, 2005, pp. 1541–1546.

[82] 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 Trans. Instrum. Meas.*, vol. 61, no. 7, pp. 2059–2072, Jul. 2012.

[83] 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*. Vitoria-Gasteiz, Spain: Springer, 2012, pp. 216–223.

[84] L. Bao and S. S. Intille, "Activity recognition from user-annotated acceleration data," in *Pervasive Computing*. Berlin, Germany: Springer, 2004, pp. 1–17.

[85] P. Siirtola and H. Röning, "Recognizing human activities user-independently on smartphones based on accelerometer data," *Int. J. Interact. Multimedia Artif. Intell.*, vol. 1, no. 5, pp. 38–45, 2012.

[86] 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*. Salamanca, Spain: Springer, 2009, pp. 796–799.

[87] Y. Kim and T. Moon, "Human detection and activity classification based on micro-Doppler signatures using deep convolutional neural networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 1, pp. 8–12, Jan. 2016.

[88] J. Donahue, L. A. Hendricks, S. Guadarrama, M. Rohrbach, S. Venugopalan, T. Darrell, and K. Saenko, "Long-term recurrent convolutional networks for visual recognition and description," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 2625–2634.

[89] S. Ha and S. Choi, "Convolutional neural networks for human activity recognition using multiple accelerometer and gyroscope sensors," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2016, pp. 381–388.

[90] S.-M. Lee, S. Min Yoon, and H. Cho, "Human activity recognition from accelerometer data using convolutional neural network," in *Proc. IEEE Int. Conf. Big Data Smart Comput. (BigComp)*, Feb. 2017, pp. 131–134.

[91] 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," *Math. Biosci. Eng.*, vol. 19, no. 6, pp. 5671–5698, 2022.

[92] 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, Oct. 2019.

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

[94] Y. Guan and T. Plötz, "Ensembles of deep LSTM learners for activity recognition using wearables," *Proc. ACM Interact., Mobile, Wearable Ubiquitous Technol.*, vol. 1, no. 2, pp. 1–28, Jun. 2017.

[95] Y. Zhao, R. Yang, G. Chevalier, X. Xu, and Z. Zhang, "Deep residual bidir-LSTM for human activity recognition using wearable sensors," *Math. Problems Eng.*, vol. 2018, pp. 1–13, Dec. 2018.

[96] 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 *Proc. Int. Conf. Comput., Math. Eng. Technol. (iCoMET)*, Mar. 2018, pp. 1–7.

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

[98] S. Seto, W. Zhang, and Y. Zhou, "Multivariate time series classification using dynamic time warping template selection for human activity recognition," in *Proc. IEEE Symp. Ser. Comput. Intell.*, Dec. 2015, pp. 1399–1406.

[99] C. O'Rourke and M. Madden, "Activity recognition based on accelerometer data using dynamic time warping with ensembles," in *Proc. 22nd Conf. Artif. Intell. Cogn. Sci.*, 2011, pp. 1–10.

[100] 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 *Proc. 8th Int. Conf. Inf. Technol. Electr. Eng. (ICITEE)*, Oct. 2016, pp. 1–5.

[101] 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 Comput.*, vol. 21, no. 1, pp. 681–690, Mar. 2018.

[102] 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 Trans. Antennas Propag.*, vol. 64, no. 11, pp. 4901–4905, Nov. 2016.

[103] M. H. El-Hoseiny and E. Shaban, "Muslim prayer actions recognition," in *Proc. 2nd Int. Conf. Comput. Electr. Eng.*, vol. 1, Dec. 2009, pp. 460–465.

[104] 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 *Proc. 8th Int. Conf. Comput. Commun. Eng. (ICCCE)*, Jun. 2021, pp. 238–243.

[105] 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, and A. B. M. A. Islam, "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.

[106] M. C. Monteiro de Barros, F. C. Leão, H. V. Filho, G. Lucchetti, A. Moreira-Almeida, and M. F. P. Peres, "Prevalence of spiritual and religious experiences in the general population: A Brazilian nationwide study," *Transcultural Psychiatry*, vol. 59, no. 2, Apr. 2022, Art. no. 136346152210887.

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

[108] 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, 2019, Art. no. e0214377.

[109] 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," *Psychol. Methods*, vol. 18, no. 1, pp. 53–70, 2013.

[110] 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, Dec. 2010.

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

[112] L. Mok and A. Anderson, "The complementary nature of perceived and actual time spent online in measuring digital well-being," *Proc. ACM Hum.-Comput. Interact.*, vol. 5, pp. 1–27, Apr. 2021.

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

[114] 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," *Psychol. Assessment*, vol. 30, no. 3, pp. 287–295, Mar. 2018.

[115] C. Wacharamanotham, L. Eisenring, S. Haroz, and F. Echtler, "Transparency of CHI research artifacts: Results of a self-reported survey," in *Proc. CHI Conf. Human Factors Comput. Syst.*, Apr. 2020, pp. 1–14.

[116] 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," *Proc. ACM Interact., Mobile, Wearable Ubiquitous Technol.*, vol. 3, no. 3, pp. 1–45, Sep. 2019.

[117] F. Shahid, W. Rahman, M. S. Rahman, S. A. Purabi, A. Seddiqa, M. Mostakim, F. Feroz, T. R. Soron, F. Hossain, N. Khan, A. B. Islam, N. Paul, E. Hoque, and A. A. Al Islam, "Leveraging free-hand sketches for potential screening of PTSD," *Proc. ACM Interact., Mobile, Wearable Ubiquitous Technol.*, vol. 4, no. 3, pp. 1–22, Sep. 2020.

[118] R. Likert, "The likert-type scale," *Arch. Psychol.*, vol. 140, no. 55, pp. 1–55, 1932.

[119] P. Bhardwaj, "Types of sampling in research," *J. Pract. Cardiovascular Sci.*, vol. 5, no. 3, p. 157, 2019.

[120] M. Hibberts, R. Burke Johnson, and K. Hudson, "Common survey sampling techniques," in *Handbook of Survey Methodology for the Social Sciences*. New York, NY, USA: Springer, 2012, pp. 53–74.

[121] SeekersGuidance. (2011). *When Should Children Start Praying?* accessed: Jun. 23, 2023. [Online]. Available: https://seekersguidance.org/answers/hanafi-fiqh/when-should-children-start-praying/

[122] (2022). *At What Age Does Prayer Become Obligatory?*. Accessed: Jun. 23, 2023. [Online]. Available: https://www.ashtoncentralmosque.com/wp/index.php/jumuah/salah-prayers/at-what-age-does-prayer-become-obligatory/

[123] *When Does it Become Obligatory to Pray?*. Accessed: Jun. 23, 2023. [Online]. Available: https://islamqa.org/shafii/qibla-shafii/33691/when-does-it-become-obligatory-to-pray/

[124] C. H. Goulden, *Methods of Statistical Analysis*, 2nd ed. New York, NY, USA: Wiley, 1952.

[125] P. E. McKnight and J. Najab, "Mann-whitney U test," in *The Corsini Encyclopedia of Psychology*. Wiley, 2010.

[126] W. Kruskal, "Kruskal and Wallis' test," *J. Amer. Statisticians Assoc.*, vol. 47, no. 260, pp. 583–618, Dec. 1952.

[127] K. I. Cheang, "Effect of learner-centered teaching on motivation and learning strategies in a third-year pharmacotherapy course," *Amer. J. Pharmaceutical Educ.*, vol. 73, no. 3, p. 42, Sep. 2009.

[128] L. Wynter, A. Burgess, E. Kalman, J. E. Heron, and J. Bleasel, "Medical students: What educational resources are they using?" *BMC Med. Educ.*, vol. 19, no. 1, pp. 1–8, Dec. 2019.

[129] *Samsung Galaxy Watch Active 2 (44 mm) Black*. Accessed: Sep. 5, 2023. [Online]. Available: https://www.shopz.com.bd/product/samsung-galaxy-watch-active-2-44mm-black/
[130] Samsung Developers, (2023). *Can't Get Magnetic Sensor Data*. Accessed: Jul. 3, 2023. [Online]. Available: https://forum.developer.samsung.com/t/cant-get-magnetic-sensor-data/26113
[131] 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 Trans. Inf. Technol. Biomed.*, vol. 10, no. 1, pp. 156–167, Jan. 2006.
[132] *The Moving Average Filter*, Sharp, Osaka, Japan, 1977.
[133] 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 *Proc. 5th Int. Conf. Future Inf. Technol.*, May 2010, pp. 1–6.
[134] 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 Trans. Inf. Technol. Biomed.*, vol. 14, no. 5, pp. 1166–1172, Sep. 2010.
[135] B. Quigley, M. Donnelly, G. Moore, and L. Galway, "A comparative analysis of windowing approaches in dense sensing environments," in *Proc. MDPI*, Oct. 2018, p. 1245.
[136] 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, Apr. 2014.
[137] 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 Comput. Surv.*, vol. 54, no. 4, pp. 1–40, May 2022.
[138] L. Cao, Y. Wang, B. Zhang, Q. Jin, and A. V. Vasilakos, "GCHAR: An efficient group-based context—Aware human activity recognition on smartphone," *J. Parallel Distrib. Comput.*, vol. 118, pp. 67–80, Aug. 2018.
[139] 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, Apr. 2015, Art. no. e0124414.
[140] J. Wang, Y. Chen, S. Hao, X. Peng, and L. Hu, "Deep learning for sensor-based activity recognition: A survey," *Pattern Recognit. Lett.*, vol. 119, pp. 3–11, Mar. 2019.
[141] A. Ferrari, D. Micucci, M. Mobilio, and P. Napoletano, "Hand-crafted features vs residual networks for human activities recognition using accelerometer," in *Proc. IEEE 23rd Int. Symp. Consum. Technol. (ISCT)*, Jun. 2019, pp. 153–156.
[142] T. Hastie, R. Tibshirani, J. H. Friedman, and J. H. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, vol. 2. Berlin, Germany: Springer, 2009.
[143] H. Qian, S. J. Pan, and C. Miao, "Latent independent excitation for generalizable sensor-based cross-person activity recognition," in *Proc. AAAI Conf. Artif. Intell.*, vol. 35, 2021, pp. 11921–11929.
[144] N. C. Klapoetke, Y. Murata, S. S. Kim, S. R. Pulver, A. Birdsey-Benson, Y. K. Cho, T. K. Morimoto, A. S. Chuong, E. J. Carpenter, and Z. Tian, "Independent optical excitation of distinct neural populations," *Nature Methods*, vol. 11, no. 3, pp. 338–346, Mar. 2014.
[145] R. Chavarriaga, H. Sagha, A. Calatroni, S. T. Digumarti, G. Tröster, J. D. R. Millán, and D. Roggen, "The opportunity challenge: A benchmark database for on-body sensor-based activity recognition," *Pattern Recognit. Lett.*, vol. 34, no. 15, pp. 2033–2042, Nov. 2013.
[146] D. Micucci, M. Mobilio, and P. Napoletano, "UniMiB SHAR: A dataset for human activity recognition using acceleration data from smartphones," *Appl. Sci.*, vol. 7, no. 10, p. 1101, Oct. 2017.
[147] D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. L. Reyes-Ortiz, "A public domain dataset for human activity recognition using smartphones," in *Proc. Esann*, vol. 3, 2013, p. 3.
[148] N. A. Capela, E. D. Lemaire, and N. Baddour, "Improving classification of sit, stand, and lie in a smartphone human activity recognition system," in *Proc. IEEE Int. Symp. Med. Meas. Appl. (MeMeA)*, May 2015, pp. 473–478.
[149] 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," *Pers. Ubiquitous Comput.*, vol. 15, no. 8, pp. 887–898, Dec. 2011.
[150] S. Nazir, M. H. Yousaf, and S. A. Velastin, "Inter and intra class correlation analysis (IIcCA) for human action recognition in realistic scenarios," in *Proc. 8th Int. Conf. Pattern Recognit. Syst. (ICPRS)*, Jul. 2017, pp. 1–6.
[151] M. Ziaeefard and R. Bergevin, "Semantic human activity recognition: A literature review," *Pattern Recognit.*, vol. 48, no. 8, pp. 2329–2345, Aug. 2015.
[152] 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 *Proc. 15th ACM SIGGRAPH Conf. Virtual-Reality Continuum Appl. Ind.*, Dec. 2016, pp. 233–242.
[153] 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 *Proc. 3rd Int. Conf. Pervasive Technol. Rel. to Assistive Environments*, Jun. 2010, pp. 1–5.
[154] V. Ahanathapillai, J. D. Amor, M. Tadeusiak, and C. J. James, "Wrist-worn accelerometer to detect postural transitions and walking patterns," in *Proc. 13th Medit. Conf. Med. Biol. Eng. Comput.*, L. M. Roa Romero, Ed. Cham, Switzerland: Springer, 2014, pp. 1515–1518.
[155] 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 Trans. Biomed. Eng.*, vol. 63, no. 4, pp. 788–796, Apr. 2016.
[156] T. K. Vintsyuk, "Speech discrimination by dynamic programming," *Cybernetics*, vol. 4, no. 1, pp. 52–57, 1972.
[157] M. Straczkiewicz, P. James, and J.-P. Onnela, "A systematic review of smartphone-based human activity recognition for health research," 2019, *arXiv:1910.03970*.
[158] 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.
[159] Y. Permanasari, E. H. Harahap, and E. P. Ali, "Speech recognition using dynamic time warping (DTW)," *J. Physics: Conf. Ser.*, vol. 1366, no. 1, Nov. 2019, Art. no. 012091.
[160] A. V. Haridas, R. Marimuthu, and V. G. Sivakumar, "A critical review and analysis on techniques of speech recognition: The road ahead," *Int. J. Knowl.-based Intell. Eng. Syst.*, vol. 22, no. 1, pp. 39–57, Mar. 2018.
[161] S. K. Saksamudre, P. P. Shrishrimal, and R. R. Deshmukh, "A review on different approaches for speech recognition system," *Int. J. Comput. Appl.*, vol. 115, no. 22, pp. 23–28, Apr. 2015.
[162] A. Mohod and D. S. Datar, "An efficient and elastic approach for partial shape matching using DTW," *Advent Technol.*, vol. 1, no. 4, pp. 1–5, 2013.
[163] D. Riboni and C. Bettini, "COSAR: Hybrid reasoning for context-aware activity recognition," *Pers. Ubiquitous Comput.*, vol. 15, no. 3, pp. 271–289, Mar. 2010.
[164] D. Berrar, "Cross-validation," in *Encyclopedia of Bioinformatics and Computational Biology*. Amsterdam, The Netherlands: Elsevier, 2019, pp. 542–545.
[165] M. Shoaib, S. Bosch, O. Incel, H. Scholten, and P. Havinga, "Fusion of smartphone motion sensors for physical activity recognition," *Sensors*, vol. 14, no. 6, pp. 10146–10176, Jun. 2014.
[166] 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," *J. Population Ageing*, vol. 13, no. 2, pp. 139–165, Jun. 2020.
[167] R. Proffitt, "Acceptability of in-home ambient sensors for activity recognition and assessment post-stroke," *Arch. Phys. Med. Rehabil.*, vol. 103, no. 12, pp. 53–54, Dec. 2022.
[168] 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, no. 6, pp. 1–34, Dec. 2013.
[169] 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, pp. 1–16, Jul. 2018.
[170] 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, Jul. 2021.
[171] 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.
[172] 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, Jan. 2015.
[173] M. N. Mohamed, *Hajj & Umrah From A to Z*. Riyadh, Saudi Arabia: Dar Eshbelia, 1999.

ISHRAT JAHAN received the B.Sc. and M.Sc. degrees from the Department of CSE, Bangladesh University of Engineering and Technology (BUET), where she is currently pursuing the Ph.D. degree. She is a Graduate Research Assistant with the Department of CSE, BUET, under the supervision of Prof. Dr. A. B. M. Alim Al Islam. Her academic pursuits are enriched by her industry background as she has experience of working as a Software Engineer with Veriflow and later with VMware. Her research interests include HCI, machine learning, recommender systems, and ubiquitous computing. She is passionate about designing technologies to support religious communities in their worship and other practices.

NAJLA ABDULRAHMAN AL-NABHAN received the B.S. (Hons.) and master's degrees in computer science from King Saud University, and the Ph.D. degree in computer science from the College of Computer and Information Science (CCIS), King Saud University. She was a visiting Ph.D. student with George Washington University, from 2012 to 2013. She is an Assistant Professor in computer science with CCIS, King Saud University. Her research interests include the Internet of Things, emergency and crowd management systems, wireless sensor networks, mobile computing, distributed systems, cloud and fog computing, smart grids, and network security.

JANNATUN NOOR received the B.Sc., M.Sc., and Ph.D. degrees in computer science and engineering from the Bangladesh University of Engineering and Technology (BUET). She is an Assistant Professor with the Department of CSE, School of Data and Sciences, BRAC University. Besides, she is a Graduate Research Assistant with the Department of CSE, BUET, under the supervision of Prof. Dr. A. B. M. Alim Al Islam. Prior to BRACU, she was the Team Lead of the IPV-Cloud Team, IPvision Canada Inc. Her research interests include cloud computing, HCI, ML, wireless networking, data mining, big data analysis, and the Internet of Things.

MASFIQUR RAHAMAN received the B.Sc. (Engg.) degree from the Department of Computer Science and Engineering, Bangladesh University of Engineering and Technology. He is currently a Graduate Research Assistant with the Department of CSE, Bangladesh University of Engineering and Technology (BUET), under the supervision of Prof. A. B. M. Alim Al Islam. His research interests include machine learning, human–computer interaction, ubiquitous computing, structural health monitoring, and intelligent transportation.

A. B. M. ALIM AL ISLAM received the B.Sc. and M.Sc. degrees in computer science and engineering from the Bangladesh University of Engineering and Technology, Dhaka, Bangladesh, and the Ph.D. degree in computer science and engineering from the School of ECE, Purdue University, West Lafayette, USA, in 2012. He is a Professor with the Bangladesh University of Engineering and Technology. His areas of research interests include wireless networks, embedded systems, modeling and simulation, computer networks security, and reliability analysis.