DESIGN AND DEVELOPMENT OF AN IMPROVED WI-FI BASED HUMAN ACTIVITY RECOGNITION USING LSTM.

JACKY CHAN CHEN HSIEN

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

DESIGN AND DEVELOPMENT OF AN IMPROVED WI-FI BASED HUMAN ACTIVITY RECOGNITION USING LSTM.

JACKY CHAN CHEN HSIEN

This report is submitted in partial fulfilment of the requirements for the degree of Bachelor of Electronic Engineering with Honours

Faculty of Electronic and Computer Engineering
Universiti Teknikal Malaysia Melaka

2021

DECLARATION

I declare that this report entitled "Design and Development of an Improved Wi-Fi based Human Activity Recognition Using LSTM" is the result of my own work except for quotes as cited in the references.

Signature :

Author : Jacky Chan Chen Hsien

Date : 22 June 2021

APPROVAL

I hereby declare that I have read this thesis and in my opinion this thesis is sufficient in terms of scope and quality for the award of Bachelor of Electronic Engineering with Honours.

Signature :

Supervisor Name : PM Dr. Wong Yan Chiew

Date : 24 Jun 2021

DEDICATION

Special dedication to my loving parents, Chan Ming Chin and Tiw Chew Umg, my kind-hearted supervisor, Prof Madya Dr. Wong Yan Chiew and thanks to my dearest friends.

ABSTRACT

Wi-Fi-based human motion sensor technology has been broadly developed from time to time. By comparing with traditional human behaviour recognition system, the benefits of using Wi-Fi-based motion technology are that it had large coverage of signal with unconstraint by any obstacles in terms of dead angle and sensitivity of light. This paper proposes an improved Wi-Fi based human activity recognition by using a deep learning method of Long Short-Term Memory (LSTM). A total of 139 data from first datasets and 140 from second datasets were obtained from internet have been tested by LSTM. Before testing, denoise process and feature extraction by discrete wavelet transform (DWT) were underwent first to extract significant feature of dataset. The performance of the extracted features will be analysed by using Bi-LSTM. The result showed 96.4% and 89.3% of top highest accuracy rate by using proposed signal pre-processing method. For the first and second raw dataset, they achieved 64.3% and 60.7% respectively since the raw datasets that had not going through process of denoise and feature extraction. Besides, the paper also discussed performance of consistency and time processing of LSTM classification between the datasets. Some recommendation and limitation also provided before the end of the paper.

ABSTRAK

Teknologi sensor gerakan manusia berasaskan Wi-Fi telah dikembangkan secara meluas dari semasa ke semasa. Dengan membandingkan dengan sistem pengenalan tingkah laku manusia tradisional, manfaat menggunakan teknologi gerakan berasaskan Wi-Fi adalah ia mempunyai liputan isyarat yang besar dan tidak dapat dibendung oleh halangan dari sudut mati dan kepekaan cahaya. Makalah ini mencadangkan peningkatan pengiktirafan aktiviti manusia berasaskan Wi-Fi dengan menggunakan kaedah pembelajaran mendalam Memori Jangka Pendek Panjang (LSTM). Sebanyak 139 data dari set data pertama dan 140 dari set data kedua yang diperoleh dari internet telah diuji oleh LSTM. Sebelum menguji, proses denoise dan pengekstrakan fitur dengan transformasi wavelet diskrit (DWT) dilakukan terlebih dahulu untuk mengekstrak ciri set data yang ketara. Prestasi ciri yang diekstrak akan dianalisis dengan menggunakan Bi-LSTM. Hasilnya menunjukkan 96.4% dan 89. 3% daripada kadar ketepatan tertinggi dengan menggunakan kaedah pra-pemprosesan isyarat yang dicadangkan. Untuk set data mentah pertama dan kedua, mereka masing-masing mencapai 64.3% dan 60.7% kerana set data mentah yang belum melalui proses pengekstrakan denoise dan fitur. Selain itu, makalah ini juga membincangkan

prestasi konsistensi dan pemprosesan masa klasifikasi LSTM antara set data. Beberapa cadangan dan batasan juga diberikan sebelum akhir makalah.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank all members of my family, especially my parents, who gave me spiritual support and encouragement. So, I can get some achievement and honor to complete these four years of studies at Universiti Teknikal Malaysia Melaka (UTeM). Throughout the final year project, I have received a lot of help from my supervisor, panels, lecturer, and friends. Without their help and support, this research would never be accomplished.

Furthermore, I would like to express my sincere gratitude to my supervisor, Prof Madya Dr. Wong Yan Chiew, who has given me valuable guidance in the writing of this final year project. In addition, my sincere appreciation goes to my panels too, Ts Dr. Muhammad Noorazlan Shah bin Zainudin and En. Muhammad Raihaan bin Kamarudin, who gave me a lot of advice and idea for my final year project. I could not have completed my project without their instruction, impressive kindness, and patience.

Moreover, I would also like to thank all the lecturers of the Faculty of Electronic and Computer Engineering, who helped me to develop the fundamental and essential academic competence within my four years of study.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

HAR	:	Human Activity Recognition
CSI	:	Channel State Information
MIMO	:	Multiple-Input and Multiple-Output
RSSI	:	Received Signal Strength Indicator
LoS	:	Line of Sight
CNN	:	Convolutional Neural Network
MD-DTW	:	Multiple-Dimensional Dynamic Time Warping
EMD	:	Earth Move Distance
HMM	:	Hidden Markov Model
SVM	:	Support Vector Machine
CWT	:	Continuous Wavelet Transform
SAC	:	Soft Actor Critic
PCA	:	Principal Component Analysis
LSTM	:	Long Short-Term Memory
DWT	:	Discrete Wavelet Transform
Bi-LSTM	:	Bi-directional Long Short-Term Memory

CHAPTER 1

INTRODUCTION

1.1 Overview

This thesis proposes the design and development of an improved Wi-Fi human activity recognition system by using LSTM. This chapter will discuss the background of project, problem statement, objectives, research problem, scope of project, project significant and chapter reviews.

1.2 Background of Project

Human activity recognition plays a crucial role in interpersonal relationship and including human-to-human interaction. This is because human activity recognition system not just provides unique features with information of a person but also containing their personality and psychological state of the person [1]. This technology is widely used in various industry that relevant to the field of military security, medical diagnostics, homecare surveillance, public security system, tracking activities for older people and others. Therefore, human activity recognition inspires and offers opportunities for many researchers to invent many applications that associate with this topic.

In general, HAR process involves several steps and procedures which is start from collecting raw data by using sensor or other appliances until the final conclusion about the desired activity. Those processes are pre-processing data, segmentation of data, feature extraction, dimensionally reduction of data and classification by using the core machine learning and reasoning [2]. Human activities actually can be classified into several decomposition which including of gestures, atomic actions, behaviors, events, group actions, and also human-to-human or human-to-object interactions. These decompositions will escalate the complexity level when it undergoes analyzation and optimization of the human activity recognition. Figure 1.0 illustrates the decomposition of human activities based on the complexity [3].

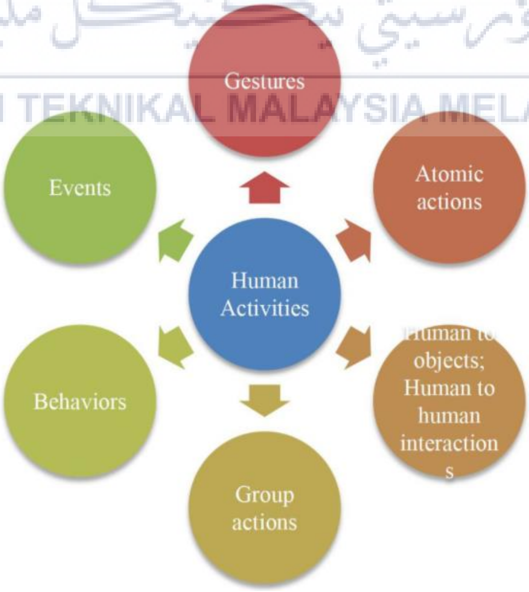


Figure 1.1: Decomposition of human activities [3].

Due to the emerging of Wi-Fi system in this new era, all the HAR applications can be linked with the element and characteristic of Wi-Fi network when the researchers are triggering their inventions. HAR with Wi-Fi network signal is originated from the idea of radio signal propagation in which radio signal can travelling freely with reflect by the obstacles or walls. When the antenna of the receiver is receiving the signal from two or more paths of sources, this will constitute multipath phenomenon. Thus, HAR can be associated with Wi-Fi signal since Wi-Fi signal also applies the multipath phenomenon [4].

Currently, Wi-Fi-based human motion sensor technology has broad developed from time to time. By compared with traditional human behavior recognition system, the benefits of using Wi-Fi-based motion technology are that it has large coverage of signal with unconstraint by any obstacles in term of dead angle and sensitivity of light [5]. Besides, Wi-Fi-based sensing applications are supported by Channel State Information (CSI) measurement since Wi-Fi network is Multiple-Input Multiple-Output (MIMO) system that provided CSI for each transmit and receive antenna [6]. The traditional motion recognition is mainly depending on the received signal strength (RSS), but RSS method is poor at motion recognition and low stability. Therefore, CSI is then replaced RSS method to improve the accuracy by its ability of describing the amplitude attenuation and phase shift of a wireless signal.

1.3 Problem Statement

Nowadays, security of human identification system is difficult to secure due to high advanced of technology such as hacking. Even though there is the existing of the biometric identification system, but some of the biometrics such as fingerprint, foot pressure, face, iris and voice can also be easily obtained by hacker or somebody else.

Besides, recognition system by using camera will also constrain by environment which is light dependency. Therefore, some emerged human activity recognition applications such as accelerometer that required wearable sensor is widely used to perform activity recognition to obtain better performance. However, this technology causes inconvenience to users because it requires them to carry the sensor to undergo recognition. Moreover, the process will be more complex and become much slower to analyze due to a large data signal will contain a lot of unnecessary information such as noise and interference.

Thus, a new type of human activity recognition that utilizing Wi-Fi signal to avoid breaching of user's privacy while still able to reduce the time of processing is highly demanded. This will also allow all the human behavior based on Wi-Fi signal be analyzed easily without the need of additional sensors.

1.4 Objectives and Aims of Project

1. To analyze and visualize the significant features in Wi-Fi based human activity recognition from established dataset using spectrogram.

2. To extract significant features of human activity from data signals by using discrete wavelet transform (DWT).

3. To analyze the performance of extracted features from discrete wavelet transform (DWT) by using Bi-LSTM.

1.5 Research Question

While conducting this project, there are several research questions have been derived. Those questions are mainly related with the efficiency of the coding in order to analyze and classify the desired subjects of the project. Since these questions have not been able to find any research about it, this project is conducted to hope that the coding or algorithms can be improved during implementation of the entire project. The following research questions have been suggested so as to fetch up the research gap:

1. What are the features needed to be considered in order to improve efficiency of the algorithms?

2. How to improve the algorithms to obtain the output?

3. Is that the improved algorithm can be worked for all types of Wi-Fi signal?

4. How the improved algorithm affected the final outcome of this project?

By solving the research questions above, the final outcome of the algorithms can increase the availability especially in the field of analyzation and classification of any signal that relevant to Wi-Fi network. Furthermore, if the algorithms workable for all Wi-Fi signals data, there will be less troublesome in deploying the sensor in certain area. This is because the human activity recognition system nowadays that related with biometrics such as facial and iris recognition are required users to close contact with sensor and they have less guarantee in securing users' privacy. Therefore, the improvement of the algorithms of this project must be considered so that this system can be broadly use in this new era of generation.

1.6 Scope of Project

This project will focus on the performance of an improved Wi-Fi human activity recognition system by using LSTM algorithm. No data collection and prototyping in this work. The performance will be testified by analyzed the different established dataset from the internet. The activities of first dataset [30] are walking, lying on bed, falling, picking up, running, sitting down and standing up. These datasets will undergo various signal processing and also deep learning to obtain recognition rate. Later on, the designed algorithms will be tested by using another established dataset from internet. The second dataset [32] is similar in term of activity samples but different in user or subject. The dataset will be classified using deep learning of LSTM model that developed in MATLAB.

1.7 Project Significant

By implementation of this project, the outcome will be the improvement of the accuracy of the recognition percentage and also reduce time processing of classification of LSTM. This mean that user can obtain higher accuracy of recognition level with shorter time taken when performing human activity recognition by using Wi-Fi signal data.

1.8 Thesis Outline

Chapter 2 will discuss the literature review on the general knowledge for analyzing and classification of the signal data. Besides, it also covers some related work and their differences. Chapter 3 is methodology of this project. It discussed about the overall flow of project, software design and development and some project management. For chapter 4, the results and discussion will be discussed in proper manner after implementation of simulation of project. All the results are explained and analyzed based on the results that displayed in term of graph and confusion matrix. Lastly, Chapter 5 is conclusion and recommendation. This chapter is summarized all the results and discussion regarding the objectives of the project. Besides, some recommendation also provided for further enhancement of the project.

CHAPTER 2

BACKGROUND STUDY

2.1 Introduction

In this chapter, it discusses about the general framework of human activity recognition (HAR). Later, some research about application of human activity recognition based on biometric will be discussed. It then followed by the review of Wi-Fi properties and related work that have been proposed before conducting this project. In order to distinguish between the previous works, the research gaps also been provided to clarify their methods used and accuracy rate.

2.2 General Framework of Human Activity Recognition (HAR)

Human activity recognition has been an active field of research of computer vision as well as human-to-human and human-to-computer interactions. Since human activity recognition has the ability to determine human activity and gait action by interpret the human motion and gesture, this technology is then widely used in various application. Therefore, its framework become the highest priority to focus in order to optimize the human activity recognition system. Figure 2.1 visualizes the general structure of human activity recognition system [7]. Generally, the information of human gesture or motion will be captured by a sensor. Then the captured information data will be interpreted, analyzed and classified by using algorithms or machine learning tools so that it can determine the type of activity has been performed.

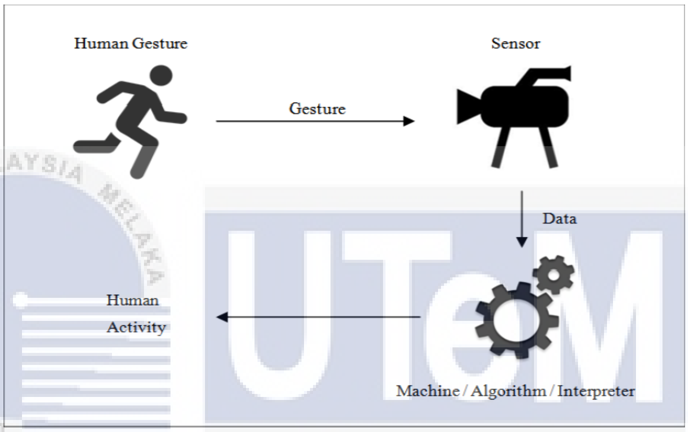


Figure 2.1: General structure of HAR system [7].

Due to the advancement of technology nowadays, human activity recognition system has been widening its field into different approaches. These approaches can be categorized into vision-based and sensor-based as shown in Figure 2.2 [8].

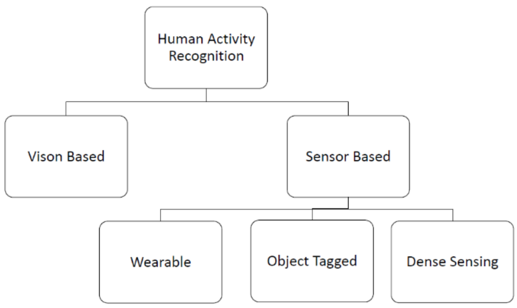


Figure 2.2: Classification of HAR approaches [8].

2.2.1 Vision-based Human Recognition System

Vision-based human recognition system actually means that it used camera as the sensor to capture the human motion or gesture. Even though this approach can obtain good result and easy to use, but this approach still has weakness in privacy issues. Besides, this method is also a light dependency which mean the camera cannot capture the clearer image when there is absence of light as its sources [9]. For further understanding about vision based HAR, Figure 2.3 below will be briefly discussed.

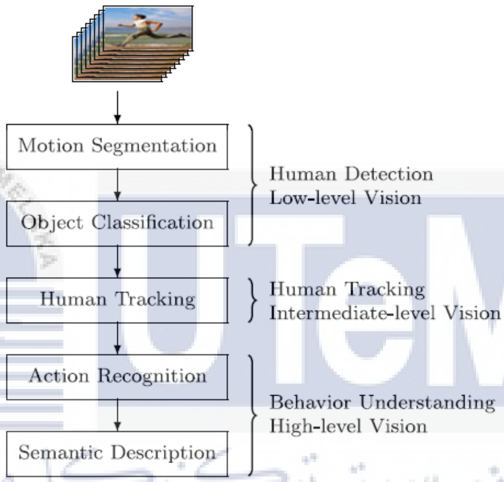


Figure 2.3: Level of HAR of vision-based [9].

Since every HAR systems start with human detection, task for human detection such as motion segmentation and object classification play a crucial role in vision based HAR systems [10]. Motion segmentation is the task of identifying the moving objects in the video and separating them from the background motion. Motion segmentation actually divided into few parts in which called, background subtraction [11], statistical methods, temporal differencing [12] and also optical flow. Meanwhile, object classification refers to the task of distinguished moving target from others moving object in an image sequence. Object classification will also be categorized by shape-based classification [13] and motion-based classification [14].

The aim of human tracking system is to using computer vision system to locate and follow people in the video imagery. This system contains strong parallaxes in which it is rich in information about the 3D nature of the scene as well as moving ground objects [15]. In human tracking system, it also distributed into different approaches such as region-based tracking, contour-based tracking, feature-based tracking, model-based tracking, hybrid tracking, and optical flow-based tracking.

The higher-level vision hierarchy of HAR system will the behavior understanding. Behavior understanding is involving the recognition and analysis of motion pattern in order to produce high-level of description action among the interaction between human and object [16]. The abnormal behavior can be detected in video based on certain classes and action that have been saved in the system. However, abnormal behavior is difficult to analyze the activity clearly.

2.2.2 Sensor-based Human Activity Recognition (HAR)

Recent years ago, most of the research in HAR system are diverted into sensor-based approach. This is due to the factors of low cost and advancement in sensor technology. Sensor-based HAR are categorized into three major aspects of deployment which are wearable, object-tagged, and dense sensing as shown in Figure 2.2 [8] above. A sensor-based HAR also similar to vision-based HAR system in term of general structure which it is just using a sensor to measure or capture the data from the users instead of camera that are light dependency. The general structure for a sensor-based human activity recognition system is visualized in Figure 2.4.

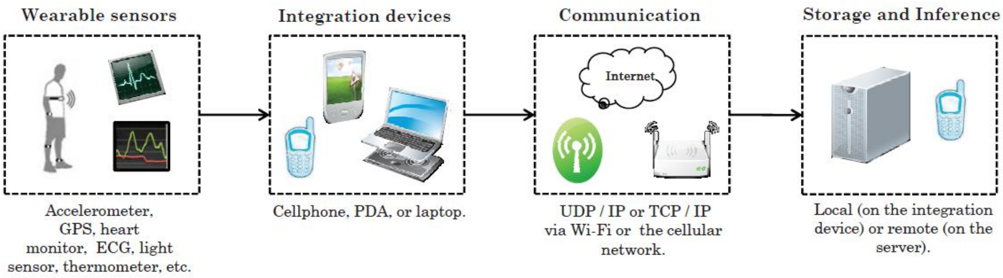


Figure 2.4: General structure of sensor based HAR system [17].

In wearable method, it is required a user to carry the sensor during the entire process of the recognition. Even though this method can surpass the camera methods in term of cost and environment issue, but this wearable sensor will still have its own weakness which is convenience when using this sensor. Some users will forget and refuse to wear the sensor when underwent any recognition system. In order to overcome this weakness, object-tagged approach can be corrected method of wearable sensor HAR.

For an object-tagged approach, the sensor will attach with the object that the users will be often use in daily such as watch, smartphone, earphones and other relevant items. Object-tagged method can be very useful to discriminate between activities such as cooking, chopping, pouring and others. For more advanced method of object-tagged approach, there is a radio frequency identification (RFID) tag has been used to optimize this technique [17]. However, this method is still bounded the users to use the tagged object which is still not feasible for all the users to implement the process of recognition.

A dense sensing approach also called as device-free approach can resolve all the problem that stated above because it does not require any tagged-object or to carry any device for them [18]. Therefore, the device-free method will become more practical due to unnecessary of users to wear and carry any devices or sensor when undergo certain activity of recognition. Since nothing is perfect, this approach also has its own challenge which is the interference from the surrounding environment. The data that collected from user will be vary depend on the surrounding condition and environment.

2.3 Human Activity Recognition using Wi-Fi signal.

Human gait is one of the unique features for human identification of a person. It can determine by a person's physical behavioral and character that similar to aspect of weight, height, and their walking habits. Since wireless network is emerging and widely deployed with high demand of wireless data, the availability of wireless signal is bringing new opportunities to human gait recognition. Wireless sensing can reuse the wireless communication infrastructure in which it is easier to use and only required low cost when compared to sensor-based and vision-based human activity recognition system. Besides, this technology also eliminated the constraint of Line of Sight (LoS) and light dependency of vision based HAR system. In term of privacy issues, wireless technology will be more secured than other method of human activity sensing [19]. One of the wireless technologies that contributed the growth of human gait recognition will be the Wi-Fi network. Figure 2.5 illustrated how the Wi-Fi network work in human gait recognition system.

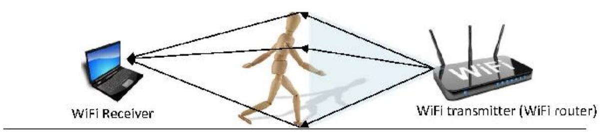


Figure 2.5: How Wi-Fi work in human gait system [19].

2.3.1 Received Signal Strength Indicator (RSSI) in Human Activity Recognition System

The Received Signal Strength Indicator (RSSI) main characterized the attenuation of signal during transmission of data and it is widely used in indoor navigation system. In typical indoor environment, the transmitted signal will be propagated to the receiver of the system via multiple paths. Each path will have its own attenuation and delayed of the signal in certain value. Therefore, a formula of signal voltage is manipulated to measure the signal of the receiver at specific time [20]:

$$V = \sum_{i=1}^{N} |V_i| e^{-j\theta_i} \tag{2.1}$$

where Vi is the amplitude of the ith multipath component and θi is the phase of the ith multipath component. In order to obtain RSSI in term of received power form (dB), simply apply the formula below:

$$RSSI = 10 \log_2 |V|^2 \tag{2.2}$$

In term of power, RSSI is mapped into the distance from the transmitter by the Log-normal Distance Path Loss (LDPL) model. This value can be calculated by using [21]:

$$PL(d)[dB] = \overline{PL(d_0)} + 10n \log(\frac{d}{d_0}) + X_\sigma \tag{2.3}$$

where PL(d) denotes the measured path loss at distance d. PL(d0) is the average path loss at reference point d0 and n is the path loss exponent. Xσ is a zero-mean normal random variable reflecting the attenuation in decibel caused by shadowing.

LDPL model visualize the variation of received signal power over distance due to path loss and shadowing. Path loss stems from the attenuation of transmission power in the propagation channel, while shadowing caused from the obstacles that attenuate signal power through absorption, reflection, scattering, and diffraction as shown in Figure 2.6.

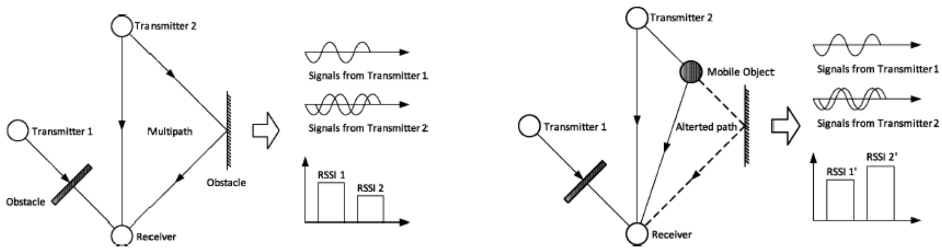


Figure 2.6: LDPL model in term of path loss [21].

Due to large path loss caused by the RSSI in LDPL model, this method has limited and constraint its accuracy [22]. Even though the possibility of RSSI to improve its accuracy to 72%, this accuracy level still much lower when compared to CARM due to the lack of frequency diverse from CSI.

2.4 Related Work

In this session, several previous works that related with human activity recognition in Wi-Fi network will be discussed. Later on, those related works are compared with each other to determine their pros and cons in order to optimize this project in the future. The topics of the previous works will be WiWho and WiFiU.

2.4.1 WiWho

WiWho has been proposed by Zeng et al [23] with a framework that can identified a small group of people with a device-free environment. It showed by the using the Channel State Information (CSI) in Wi-Fi to identify a person's walking gait. The overall structure of WiWho can be referred to Figure 2.7. There are two stationary endpoints in the room in where the WiWho will be deployed at. These endpoints can perform transmission and receiving process between each other when it is switched on. One of the endpoints will be Wi-Fi AP and the other will be equipped device such as laptop.

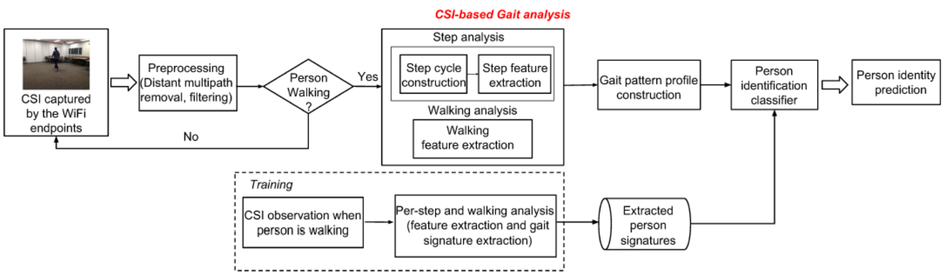


Figure 2.7: General architecture of WiWho [23].

When the WiWho detected there is someone walking, the CSI samples will the input to the gait analysis method. This gait analysis consists of two parts: step analysis and walk analysis. In the step analysis, the step cycle is constructed from the CSI data and for each of the detected step, and various features of its shape are derived. For walk analysis, it analyzes the overall walking behavior of the person for the entire walk segment (multiple steps). This provides information on various body movements that can be different from person to person.

WiWho using experiments at multiple locations with a total of 20 volunteers and show that it can identify a person with average accuracy of 92% to 80% from a group of 2 to 6 people respectively. It also showed that in most cases walking as few as 2-3 meters is sufficient to recognize a person's gait and identify the person. However, WiWho still has its own limitation such as feasibility, number of people, diverse set of people and also environment.

2.4.2 WiFiU

WifiU is proposed by Wang et al. [24] which uses commercial WiFi devices to capture fine-grained gait patterns to recognize humans. The objective is that due to the differences in gaits of different people, the WiFi signal reflected by a walking human generates unique variations in the Channel State Information (CSI) on the WiFi receiver. The Figures 2.8 and 2.9 showed the application scenario and also data collection environment of WiFiU.

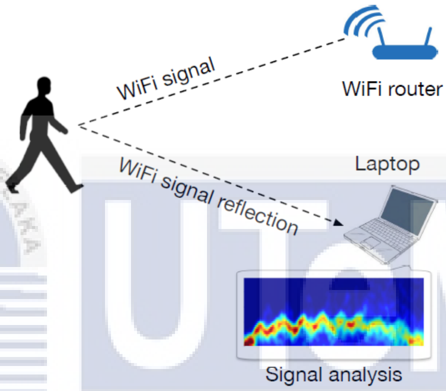


Figure 2.8: Application scenario [24].

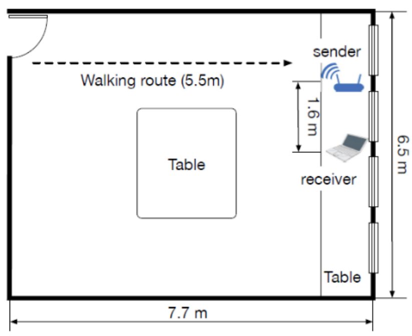


Figure 2.9: Data collection environment [24].

Later on, verification of human movement using CSI will require some signal processing techniques which is generated spectrograms from CSI measurements so that the resulting spectrograms are similar to those generated by specifically designed Doppler radars. To extract features from spectrograms that best characterize the walking pattern, the autocorrelation is performed on the torso reflection to remove imperfection in spectrograms. WifiU is evaluated on a dataset with 2,800 gait instances collected from 50 human subjects walking in a room with an area of 50 square meters. Experimental results show that WifiU achieves top-1, top-2, and top-3 recognition accuracies of 79.28%, 89.52%, and 93.05%, respectively.

Despite of that, WifiU also has two limitation and constraints during current implementation. Firstly, the users must walk on a predefined path in a predefined walking direction. The classification models trained for a given walking path and direction cannot be used for testing samples obtained on different walking paths and directions. This is because Doppler spectrograms are sensitive to the walking directions. Therefore, the current implementation of WifiU is only suitable for confined spaces, such as a corridor or a narrow entrance. Secondly, WifiU cannot allow multiple users to walk at the same time. Although static users have no impact to the performance of WifiU, but it is difficult to separate the gait signals from multiple moving users.

2.4.3 WiFi-ID

Zhang et al. [25] had proposed WiFi-ID which is a potential device-free WiFi sensing for human recognition system. A transmitter node continuously sent packets to a receiver node which passively records CSI data from the received packets. The CSI data captured the aggregate impact of multi-path, shadowing and interference on the WiFi signals in a given environment. Considering Figure 2.10 as the general operation of scenario for WiFi-ID. When a person passed through it, their gait impacts the environment in a unique manner, which changes the effect of these phenomena on the WiFi signal. It is expected that these are in turn manifested as unique perturbations in the CSI data.

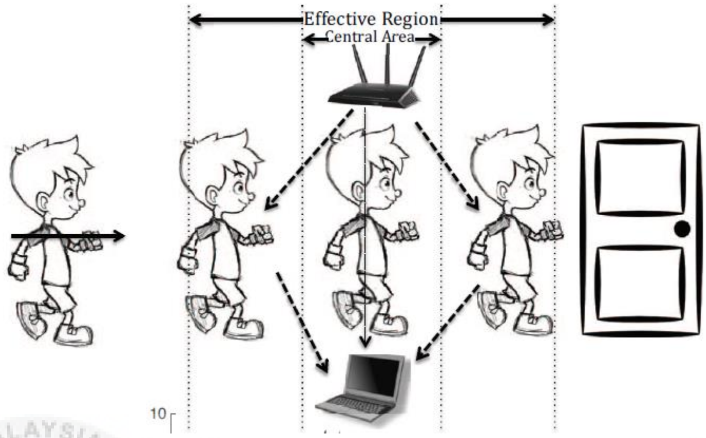


Figure 2.10: General operation for WiFi-ID [25].

Since the CSI data is rich enough to capture unique signatures of a person's gait, the perturbations caused by the motion of human limbs while walking is typically concentrated in certain specific RF bands. Using empirical measurement, it demonstrated that the effects of human gait are most pronounced in the 20-80 Hz frequency bands for 5GHz WiFi. In order to find the unique patterns of each person, WiFi-ID computed multiple features from the CSI data. These features maintained both time and frequency domain information of the CSI time series in order to capture the effects of the gait on an individual. However, given the dimension of the CSI data, the resulting feature set is very large. Therefore, ReliefF will be applied to efficient feature selection algorithm to rank the various features. A silence removal method also designed which analyzed the short time energy of the CSI time series data to determine the start and end points of the effective region. Figure 2.11 illustrates the overview of

WiFi-ID system which consists of its steps required to complete the recognition system.

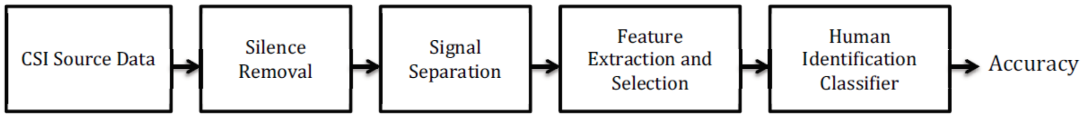


Figure 2.11: Overview of the WiFi-ID system [25].

WiFi-ID analyzed the perturbations observed in the CSI data to identify unique features that allow identification of individuals. WiFi-ID achieved 93% to 77% human identification accuracy for 2 to 6 individuals in a group, respectively. For further information, several limitations are listed out for future challenge in WiFi-ID. Firstly, the scenario must be considered where in the participants directly cut across the Line of Sight (LoS) path between the transmitter and receiver. Secondly, WiFi-ID is just considered a simple setting of identifying a person from a maximum group size of 6 people. Therefore, it will lead to false identification if a large size of group is used on WiFi-ID.

2.4.4 EI

A deep learning-based device free activity recognition system named EI is proposed by Jiang et al [26]. Its framework can remove the environment and subject specific information contained in the activity data. It also extracted environment or subject-independent features that shared by the data collected on different subjects under different environments. The core of EI is an adversarial network, which consists of three main components: feature extractor, activity recognizer, and domain discriminator. The feature extractor, which is a Convolutional Neural Network (CNN), cooperates with the activity recognizer to carry out the major task of recognizing human activities. Figure 2.12 illustrated the framework of EI.

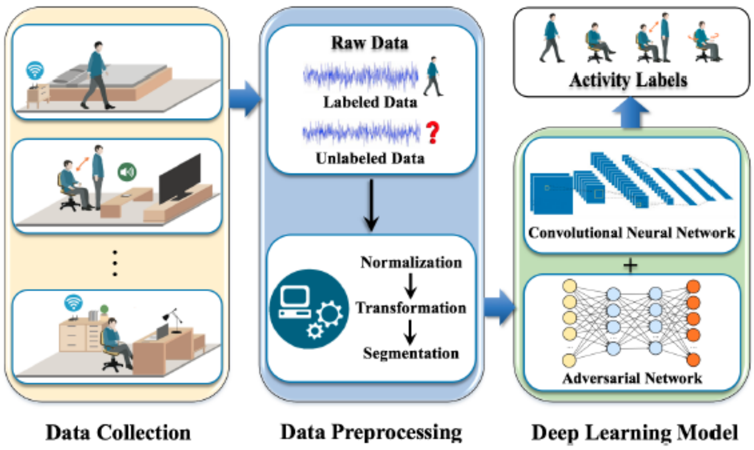


Figure 2.12: Framework of EI [26].

EI employ CNNs to extract activity features, which are widely used in the human activity recognition task. In the proposed approach, a three-layer stacked CNNs are used to extract features. In each layer of CNNs, 2D kernels are used as the filters, followed by a batch norm layer to normalize the mean and variance of the data at each layer. At last, a rectified linear unit (ReLU) is added to introduce nonlinearity and a max-pooling layer to reduce the size of representation. Next, a deep learning model, which incorporates an adversarial network, to predict the label of unlabeled activities. The proposed deep learning model can not only make use of labeled data, but also take advantage of the information contained in the unlabeled data that can help improve the predictive performance. Lastly, an accuracy of 75% had been achieved in this system in recognizing activities such as wiping the whiteboard, walking, moving a suitcase, rotating the chair, sitting, standing up and sitting down.

2.4.5 E-eyes

Wang et al. [27] proposed E-eyes which is a device-free location-oriented activity identification at home using existing Wi-Fi access points and Wi-Fi devices. The low-cost system takes advantage of the ever more complex web of Wi-Fi links between such devices and the increasingly fine-grained channel state information that can be extracted from such links. It examined channel features and can uniquely identify both in-place activities and walking movements across a home by comparing them against signal profiles. Signal profiles construction can be semi-supervised and the profiles can be adaptively updated to accommodate the movement of the mobile devices and day-to-day signal calibration.

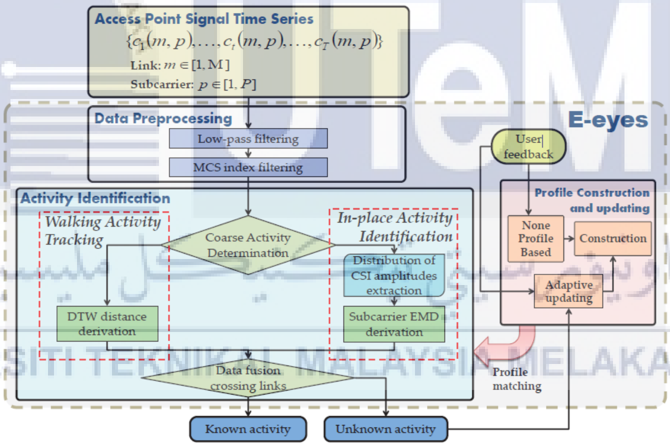


Figure 2.13: System flow of E-eyes [27].

The core of E-eyes are the activity identification, the profile construction and its updating feature. Activity identification encompasses two different activity matching approaches to address the generality challenge. The system distinguishes between walking activities and in-place activities. In general, a walking activity causes significant pattern changes of the CSI amplitude over time. Next, the system identified activities by calculating the similarity between such a CSI segment and the preconstructed activity profiles. Based on the characteristics of walking and in-place activities, two separate similarity metrics and classifiers are designed. For walking activities, Multiple-Dimensional Dynamic Time Warping (MD-DTW) technique is used, which can align a trace with larger CSI changes to the profile while correcting for differences in speed. While for in-placed activity, the Earth Move Distance (EMD) is used to qualify the similarity of two CSI histogram distributions. The experimental evaluation in two apartments of different size demonstrates that the approach can achieve over 96% average true positive rate and less than 1% average false positive rate to distinguish a set of in-place and walking activities with only a single Wi-Fi access point.

2.4.6 CARM

Wei et al. [28] proposed CARM, a CSI based human activity recognition and monitoring system. CARM has two theoretical underpinnings, a CSI-speed model and CSI-activity model. CSI-speed model quantified the correlation between CSI value dynamics and human movement speeds while CSI-activity model is used which quantified the correlation between the movement speeds of different human body parts and a specific human activity. By these two models, a system is quantitatively built the correlation between CSI value dynamics and a specific human activity. CARM used this correlation as the profiling mechanism and recognizes a given activity by matching it to the best-fit profile. The CARM basically using commercial Wi-Fi devices and evaluated it in several different environments as shown in Figure 2.14.

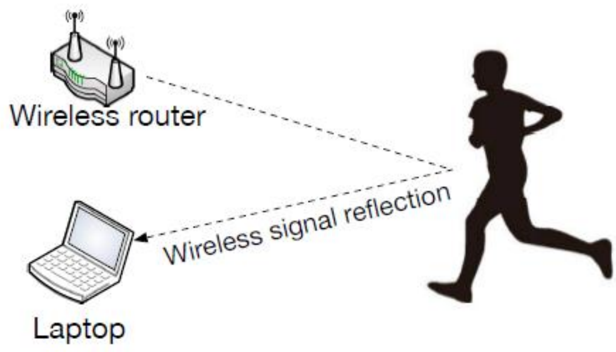


Figure 2.14: CARM system [28].

This system used Hidden Markov Model (HMM) to build in CSI activity models that consist of multiple movement states. The HMM captured information from all training samples and thus worked very well even when there is high within-class variance. Provided that enough representative training samples of an activity are available, an HMM can be constructed that implicitly models all of the many sources of variability inherent in the activity. The results showed that CARM achieves an average accuracy of greater than 96%.

2.4.7 Survey on Behavior Recognition Using Wi-Fi Channel State Information.

A survey of recent advances in passive human behavior recognition in indoor areas using the channel state information (CSI) of commercial Wi-Fi systems is performed by Siamak et al. [29]. The human behavior can be recognized successfully by analyzed the data streams of CSIs for different activities. The data streams are then used machine learning techniques to build models and classifiers. The techniques from the literature that are presented have great performance Besides, this paper also proposed to use deep learning techniques such as long-short term memory (LSTM) recurrent neural networking (RNN) to improve performance when compared to other machine learning technique. Figure 2.15 illustrated the general structure of common activity recognition technique.

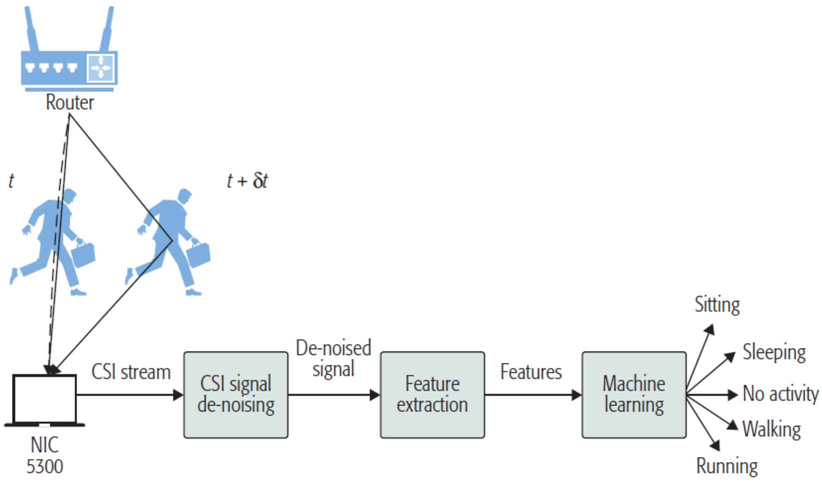


Figure 2.15: General structure of common activity recognition technique [29].

After that, the performance of LSTM using Tensor flow in Python is evaluated by author. The input feature vector is the raw CSI amplitude data, which is a 90-dimensional data. The LSTM approach is different from conventional approaches because it did not used PCA and STFT during classification. Nevertheless, LSTM can extract features from CSI directly. The number of hidden units is chosen to be 200 where it considered only one hidden layer. For numerical minimization of cross entropy, the stochastic gradient descent (SGD) with batch size of 200 and learning rate of 10–4 is used for configuration of LSTM. The result is showed after classification where the accuracy is 90.5%.

2.5 Research Gaps

In this session, the related topics above are being compared in term of pros and cons. Table 2.1 illustrates their limitation and constraints so that it is easier to be compared. Even though all the related works have achieved high accuracy in recognition training of deep learning, but there still have disadvantages and limitation based on their works. Therefore, this project will focus not just the accuracy of the recognition rate but also including the efficiency of the algorithms. In other words, it determined whether the improved algorithms can be used for all Wi-Fi signal from established dataset while maintaining high accuracy of recognition rate.

Table 2.1: Comparison between related work.

Topics	Method	Accuracy
WiWho [23]	Decision Tree	80% – 92%
WifiU [24]	SVM	79% – 93%
WiFi- ID [25]	CWT + SAC	77% – 93%
EI [26]	CNN	75%
E-eyes [27]	EMD + MD-DTW	96%
CARM [28]	HMM	96%
Survey [29]	LSTM	90.5%

2.6 Chapter Summary

This chapter is started with the general framework of human activity recognition that discussed the different of vision based and sensor based HAR based on some articles and journals. Next, it then discussed the human gait recognition by using Wi-Fi signal. Certain ways and formulae are given in order to calculate the RSSI and also path loss during propagation of Wi-Fi signals. Lastly, the chapter ended with some related works that have been established so that can easily analyze and compare between proposed method and previous works.

CHAPTER 3

METHODOLOGY

3.1 Introduction

In this chapter, it will discuss about the method to implement this project with title of improved Wi-Fi based human activity recognition using LSTM. This session will be included of flowchart, overall flow of project, software design and development, project management, environmental and sustainability.

3.2 Overall Flow of Project

Based on the Figure 3.1, the project will be started with the importation of Wi-Fi signal dataset [30] into MATLAB. The sources must be complied with the format as the requirement of the MATLAB. If the sources are found out incompatible in format, another data source need to determine from the internet again. After the importation succeed, the data will undergo pre-procession. This step will reduce some unnecessary information of the raw data which will be very helpful in processing a data.

Later, the that will continued proceed to the next stage which is feature extraction. In this session, the significant part of the signal data will be filtered out in order to improve the accuracy and also time taken to analyze it. Next, the extracted features will be tested and trained by using LSTM algorithms in MATLAB. The real data will be compared with the predicted data in term of confusion matrix form so that to clarify the accuracy between them. If the result is not satisfied, the coding or algorithms will be modified to achieve desired results.

Last but not least, the whole algorithms will be modified in order to enhance and simplified the data signal processing. Besides, the algorithms will have the possibility to change so that the whole process of Wi-Fi signal processing will become more advanced.

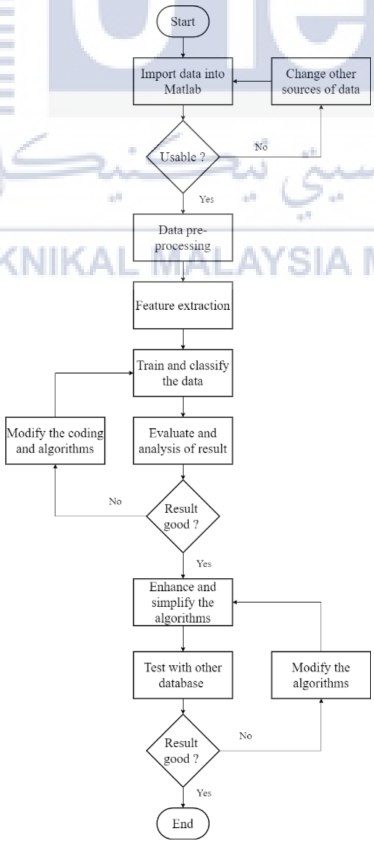


Figure 3.1: Flowchart of the project.

3.3 System Design and Development

In this proposed project, MATLAB software is using for this entire project of title of improved Wi-Fi based human activity recognition using LSTM. MATLAB software is a powerful tool and high-level programming language that allow us to do things that relevant data analysis. It also provided useful tips and training for the beginner user such as deep learning in MATLAB. This software not just good in data analysis but also explore its possibility in wireless communication, computer vision, signal processing, robotics, control system, quantitative finance and also risk management. Therefore, MATLAB is appropriate for this proposed system. Figure 3.2 shown the icon of MATLAB software used in this project.

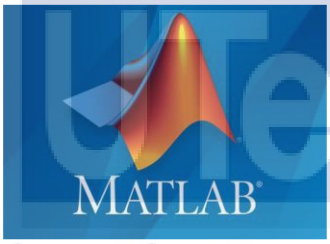
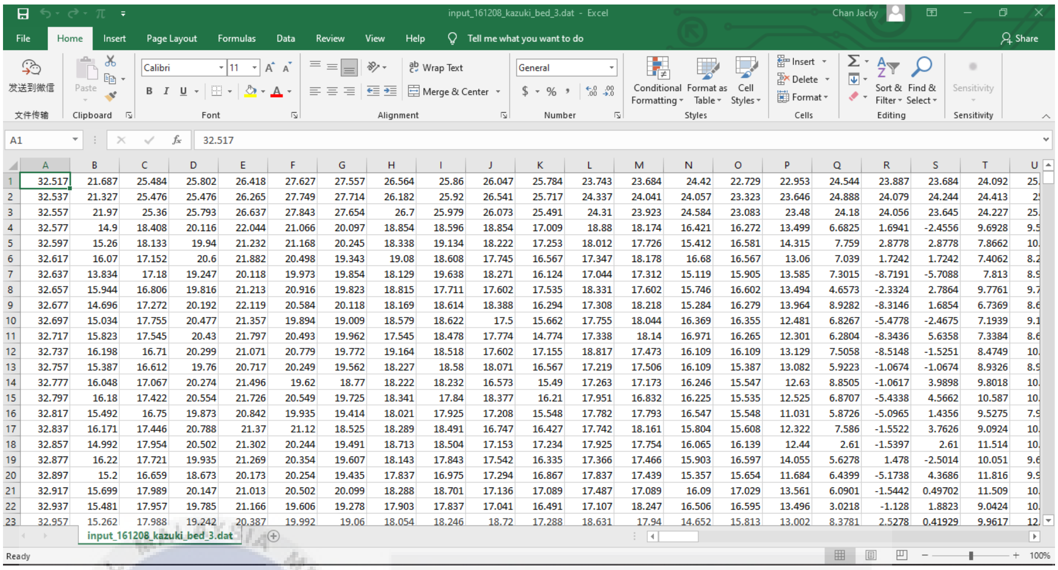


Figure 3.2: MATLAB icon [31].

3.3.1 Data Preparation

Since there is no prototype or hardware in this project, the data will be obtained from internet so that the project can analyze the signal data efficiently. There are several Wi-Fi data on the internet, the first dataset that will be analyze is named as “LSTM_wifi_activity_recognition” by Hirokazu-Narui et al. [30]. This author has provided plenty of data and categorized every action accordingly in his/she files. The second dataset that obtained is from author Siamak Yousefi et al. [32] with title of “Wifi_Activity_Recognition using LSTM”. Besides, the files are in the format of

".csv" which is a Microsoft Excel file and it is suitable for the MATLAB to open it. Figure 3.3 shown the data of Wi-Fi signal used in this project.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	32.517	21.687	25.484	25.802	26.418	27.627	27.557	26.564	25.86	26.047	25.784	23.743	23.684	24.42	22.729	22.953	24.544	23.887	23.684	24.092	25.
2	32.537	21.327	25.476	25.476	26.265	27.749	27.714	26.182	25.92	26.541	25.717	24.337	24.041	24.057	23.323	23.646	24.888	24.079	24.244	24.413	2
3	32.557	21.97	25.36	25.793	26.637	27.843	27.654	26.7	25.979	26.073	25.491	24.31	23.923	24.584	23.083	23.48	24.18	24.056	23.645	24.227	25.
4	32.577	14.9	18.408	20.116	22.044	21.066	20.097	18.854	18.596	18.854	17.009	18.88	18.174	16.421	16.272	13.499	6.6825	1.6941	-2.4556	9.6928	9.5
5	32.597	15.26	18.133	19.94	21.232	21.168	20.245	18.338	19.134	18.222	17.253	18.012	17.726	15.412	16.581	14.315	7.759	2.8778	2.8778	7.8662	10.
6	32.617	16.07	17.152	20.6	21.882	20.498	19.343	19.08	18.608	17.745	16.567	17.347	18.178	16.68	16.567	13.06	7.039	1.7242	1.7242	7.4062	8.2
7	32.637	13.834	17.18	19.247	20.118	19.973	19.854	18.129	19.638	18.271	16.124	17.044	17.312	15.119	15.905	13.585	7.3015	-8.7191	-5.7088	7.813	8.9
8	32.657	15.944	16.806	19.816	21.213	20.916	19.823	18.815	17.711	17.602	17.535	18.331	17.602	15.746	16.602	13.494	4.6573	-2.3324	2.7864	9.7761	9.7
9	32.677	14.696	17.272	20.192	22.119	20.584	20.118	18.169	18.614	18.388	16.294	17.308	18.218	15.284	16.279	13.964	8.9282	-8.3146	1.6854	6.7369	8.6
10	32.697	15.034	17.755	20.477	21.357	19.894	19.009	18.579	18.622	17.5	15.662	17.755	18.044	16.369	16.355	12.481	6.8267	-5.4778	-2.4675	7.1939	9.1
11	32.717	15.823	17.545	20.43	21.797	20.493	19.962	17.545	18.478	17.774	14.774	17.338	18.14	16.971	16.265	12.301	6.2804	-8.3436	5.6358	7.3384	8.6
12	32.737	16.198	16.71	20.299	21.071	20.779	19.772	19.164	18.518	17.602	17.155	18.817	17.473	16.109	16.109	13.129	7.5058	-8.5148	-1.5251	8.4749	10.
13	32.757	15.387	16.612	19.76	20.717	20.249	19.562	18.227	18.58	18.071	16.567	17.219	17.506	16.109	15.387	13.082	5.9223	-1.0674	-1.0674	8.9326	8.9
14	32.777	16.048	17.067	20.274	21.496	19.62	18.77	18.222	18.232	16.573	15.49	17.263	17.173	16.246	15.547	12.63	8.8505	-1.0617	3.9898	9.8018	10.
15	32.797	16.18	17.422	20.554	21.726	20.549	19.725	18.341	17.84	18.377	16.21	17.951	16.832	16.225	15.535	12.525	6.8707	-5.4338	4.5662	10.587	10.
16	32.817	15.492	16.75	19.873	20.842	19.935	19.414	18.021	17.925	17.208	15.548	17.782	17.793	16.547	15.548	11.031	5.8726	-5.0965	1.4356	9.5275	7.9
17	32.837	16.171	17.446	20.788	21.37	21.12	18.525	18.289	18.491	16.747	16.427	17.742	18.161	15.804	15.608	12.322	7.586	-1.5522	3.7626	9.0924	10.
18	32.857	14.992	17.954	20.502	21.302	20.244	19.491	18.713	18.504	17.153	17.234	17.925	17.754	16.065	16.139	12.44	2.61	-1.5397	2.61	11.514	10.
19	32.877	16.22	17.721	19.935	21.269	20.354	19.607	18.143	17.843	17.542	16.335	17.366	17.466	15.903	16.597	14.055	5.6278	1.478	-2.5014	10.051	9.6
20	32.897	15.2	16.659	18.673	20.173	20.254	19.435	17.837	16.975	17.294	16.867	17.837	17.439	15.357	15.654	11.684	6.4399	-5.1738	4.3686	11.816	9.9
21	32.917	15.699	17.989	20.147	21.013	20.502	20.099	18.288	18.701	17.136	17.089	17.487	17.089	16.09	17.029	13.561	6.0901	-1.5442	0.49702	11.509	10.
22	32.937	15.481	17.957	19.785	21.166	19.606	19.278	17.903	17.837	17.041	16.491	17.107	18.247	16.506	16.595	13.496	3.0218	-1.128	1.8823	9.0424	10.
23	32.957	15.262	17.988	19.242	20.387	19.992	19.06	18.054	18.246	18.72	17.288	18.631	17.94	14.652	15.813	13.002	8.3781	2.5278	0.41929	9.9617	12.

Figure 3.3: Example of Wi-Fi signal collected data format.

3.3.2 Data Pre-Processing

Based on the data in Figure 3.3, the dataset contained of 181 column of data and 1000 row for the time stamp. The first column will be the time stamp of the dataset whereas the rest will be divided into CSI amplitude value and CSI phase value. CSI amplitude value will be started from column 2 to 91 while the CSI phase magnitude is started from column 92 to 181. Inside the CSI amplitude and phase magnitude, it only provided 30 subcarriers CSI value. Therefore, this data actually collected total up to 3 antennas data in which contain 30 subcarriers each.

Signal pre-processing that used in this project will be the denoise function of raw datasets. Denoising function allow user to filter the background noise that caused by the environment or any unwanted noise caused by other factors. After denoising the signals, the complexity of the signals will be reduced and easier to analyze the significant parts of the signal based on certain human activity.

3.3.3 Spectrogram Visualization

Spectrogram is a very useful tool in visualization of signal data. A spectrogram is a visual way of representing the signal strength of a signal over time at various frequencies present in a particular waveform. Spectrograms are basically two-dimensional graphs, with a third dimension represented by colors. Time will run from left to right along the horizontal axis whereas vertical axis represents frequency, which can also be thought of as a pitch or tone, with the lowest frequencies at the bottom and the highest frequencies at the top. In MATLAB software, spectrogram can be plotted by using command of "s = spectrogram(x)" in order to analyze desired signal waveform. Therefore, the significant parts of signal that after pre-processing can be easily detected by using spectrogram since the color can indicated based on different level of signal strength. Figure 3.4 illustrated the example of a spectrogram in MATLAB.

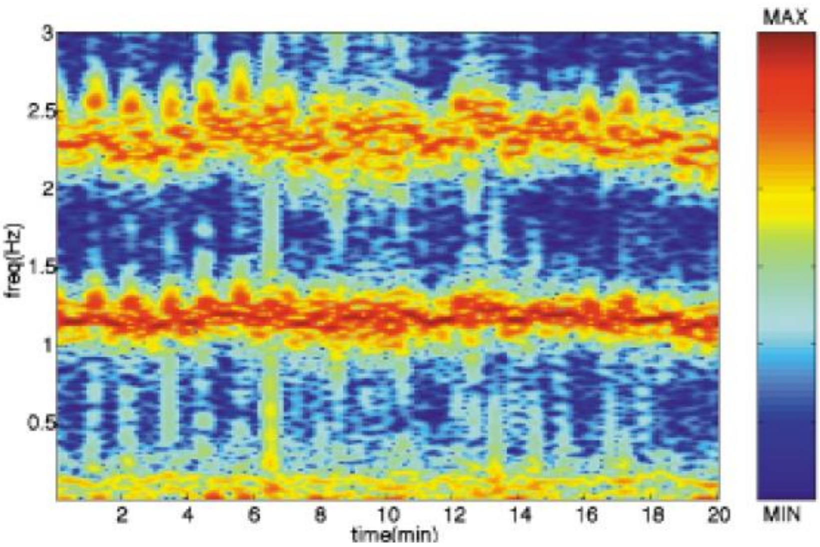


Figure 3.4: Example of spectrogram in MATLAB [32].

3.3.4 Feature Extraction

Feature extraction is a technique that transforming raw data into numerical feature in which MATLAB can be processed while still preserving the important feature of original dataset. It also can be explained as dimensionality reduction process of raw data is then divided and reduced to more manageable groups that represent the key feature of the original data. There are many applications of feature extraction such as Principal Component Analysis (PCA), Independent Component Analysis (ICA), Linear Discriminant Analysis (LDA) and others. The feature extraction method proposed for this project is Discrete Wavelet Transform (DWT).

Discrete wavelet transform (DWT) is a technique that decomposed a given signal into a number of sets where the set is a time series of coefficients describing signal corresponding to the frequency band. DWT produced two coefficient sets which are approximation coefficient (cA) and detailed coefficient (cD). These coefficients convolving the signal that is underwent low pass filter for approximation while high pass filter for detailed part. After extracted the features, the data can be used for training and testing by using deep leaning algorithms. Figure 3.5 illustrated the decomposition of DWT.

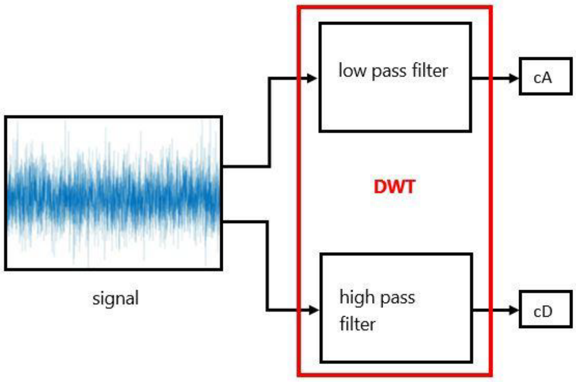


Figure 3.5: Decomposition of DWT.

3.3.5 Deep Learning of Data

The deep learning method that used for this project is Long Short-Term Memory (LTSM). LSTM is a recurrent neural network (RNN) architecture based that used in the field of deep learning. The core element for LSTM network is the sequence input layer which will input sequence or time series data into LSTM network. Figure 3.6 below demonstrated the architecture of the LSTM network. In order to predict the class labels, the network normally end with a fully connected layer, a softmax layer and also a classification output layer.

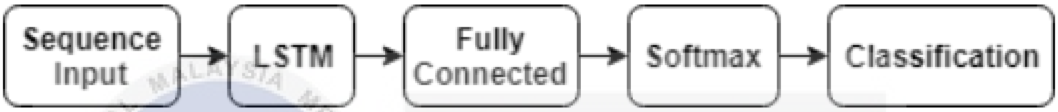


Figure 3.6: Architecture of LSTM network.

3.3.5.1 Training and Testing Dataset

In order to classify human activity by using LSTM, the raw dataset needed to distribute into two categories which is training and testing subsets. The training subset is used for LSTM to learn the pattern or key feature of the data input whereas the testing subset is used to evaluate the performance of LSTM. Table 3.1 below shown the distribution of datasets into ratio of 8:2 as training and testing datasets.

Table 3.1: Distribution of datasets for training and testing subsets.

Dataset	Training Subset	Testing Subset	Total
First dataset [30]	112	27	139
Second dataset [32]	112	28	140

3.3.5.2 Proposed LSTM

LSTM also categorized into various types such as uni-directional LSTM, Bi-LSTM and Cascaded LSTM. In this project, bi-directional Long Short-Term Memory (Bi-LSTM) is proposed. Figure 3.7 below illustrated the structure of a Bi-LSTM network. Bidirectional LSTM structure allow networks to have both backward and forward information about the sequence at every time step. When compared with uni-LSTM, bi-LSTM will perform well due to its mechanism that allow the data to read backward and forward which will increase the predicted score during classification.

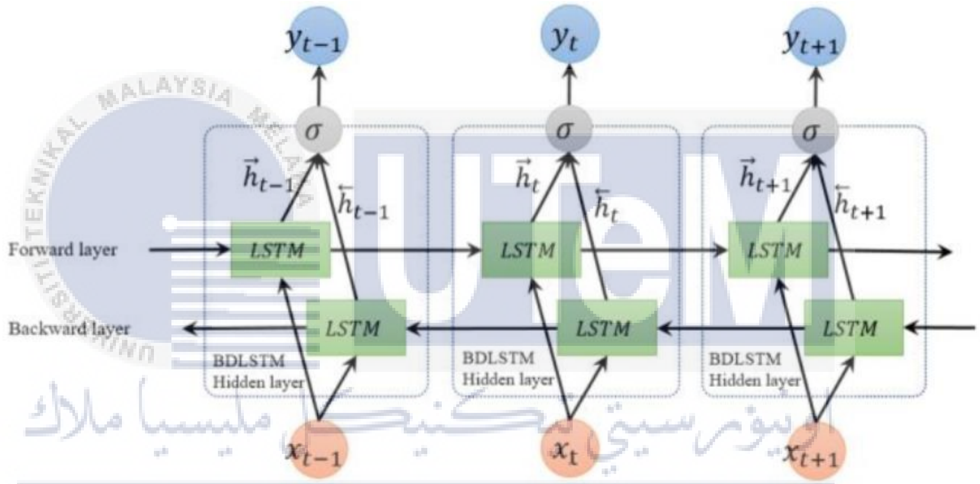


Figure 3.7: Structure of Bi-LSTM network.

3.3.5.3 Implementation of LSTM

First of all, the sequence input data needed to determine in order to procced the classification of LSTM. Figure 3.8 below demonstrated the input data of LSTM for this project. “Signals.mat” file represented the Wi-Fi signal that been underwent signal pre-processing and feature extraction whereas “Labels.mat” file is represented as the correspond ground truth label of each signal. Be aware that the format for signals data and ground truth label must be the same data type so that the LSTM can proceed with the classification.

```
load('Labels.mat')
load('Signals.mat')
```

Figure 3.8: Sequence input data for LSTM.

Next, the dataset will be distributed randomly in term of desired ratio. Figure 3,9 below shown the code for dataset distribution into training and testing subset in LSTM algorithm. The ratio is set as 8:2 due to the performance of the accuracy can be improved by the increment of the training subsets. The number of dataset for each human activity can be displayed in command window by using "summary(Labels)" as shown in Figure 3.10.

```
[trainBed,~,testBed] = dividerand(20,0.8,0.0,0.2);
[trainFall,~,testFall] = dividerand(20,0.8,0.0,0.2);
[trainPickup,~,testPickup] = dividerand(20,0.8,0.0,0.2);
[trainRun,~,testRun] = dividerand(20,0.8,0.0,0.2);
[trainSitDown,~,testSitDown] = dividerand(20,0.8,0.0,0.2);
[trainStandUp,~,testStandUp] = dividerand(20,0.8,0.0,0.2);
[trainWalk,~,testWalk] = dividerand(19,0.8,0.0,0.2);
```

Figure 3.9: Coding for dataset distribution into training and testing subset in LSTM.

```
>> summary(Labels)
     Bed            20
     Fall           20
     Pick Up        20
     Run            20
     Sit Down       20
     Stand Up       20
     Walk           19
```

Figure 3.10: Total number of datasets for each human activity.

Besides, the configuration of LSTM needed to be set up based on different type of LSTM. Figure 3.11 illustrated the layer configuration of LSTM. The sequence input layer is depended on the signal dataset. The input is set as 90 due to the dataset had 90x1000 double type data for each human activity. Therefore, this setting will depend on the input data type and number. Moreover, Bi-LSTM is used in this project and the fully connected layer represented the output layer in which total number of 7 was set because there had 7 different types of human gait action in the datasets. In term of training parameter, Figure 3.12 shown the setting for training parameter in LSTM. The training parameter is used “adam” as main core to train the datasets. In additional, others configuration such as “MaxEpoch”, “MiniBatchSize” and “InitialLearnRate” are set as shown in Figure below before running the classification of LSTM.

```
layers = [ ...
    sequenceInputLayer(90)
    bilstmLayer(100,'OutputMode','last')
    fullyConnectedLayer(7)
    softmaxLayer
    classificationLayer
    ];
```


Figure 3.11: Layer configuration of LSTM.

```
options = trainingOptions('adam', ...
    'MaxEpochs',70, ...
    'MiniBatchSize', 32, ...
    'InitialLearnRate', 0.001, ...
    'SequenceLength', 'longest', ...
    'GradientThreshold', 1, ...
    'ExecutionEnvironment','auto',...
    'plots','training-progress', ...
    'Verbose',true);
```

Figure 3.12: Training parameter of LSTM.

Lastly, the accuracy of the testing part will implement once the training process is done for LSTM. The Figure 3.13 demonstrated the coding for calculating the accuracy of the LSTM for this project. After that, the result is displayed in the form of matric which called as confusion matrix. Confusion matrix is used for describing the performance of a classification model on certain set of data. The result of accuracy based on the confusion matrix will be discussed and displayed in Chapter 4.

```
testPred = classify(net,XTest,'SequenceLength','longest');

testLSTMAccuracy = sum(testPred == YTest)/numel(YTest)*100
```

Figure 3.13: Coding for accuracy calculation in LSTM.

3.4 Project Management

This project will focus on analyzing, verifying and enhancing on MATLAB software only. Hence, there will be zero cost for this project.

3.5 Environment and Sustainability

In this era of the emerging of Wi-Fi, this project can sustain a very long period due to the high demand of Wi-Fi system. Besides, new era of 5G network is also promote the growth of Wi-Fi system to this world. Therefore, this project will be functioned once the presence of Wi-Fi system is guarantee. In term of environmentally, this project will not cause any disturbance to surrounding environment. Thus, this project is environmentally friendly to the surrounding.

3.6 Chapter Summary

In this chapter, the overall project workflow is discussed and displayed by using a flowchart. Next, the method of implementation of this project also decided and explained clearly in this chapter. The system software used for this project is

MATLAB software only which required no cost for this project. The dataset that used is downloaded on internet. These raw data is then undergoing pre-processing of denoise, feature extraction of DWT and classified by using a Bi-LSTM algorithm to obtain the accuracy which is displayed in term of confusion matrix. All the results of each process will be shown in Chapter 4. Lastly, the environment and the sustainability of this project is also discussed before ended the chapter.

CHAPTER 4

RESULTS AND DISCUSSION

4.1 Introduction

This chapter will discuss all the result related to the process had been implemented by using MATLAB. Firstly, the visualization of CSI signal data based on different types of human gait action are presenting by using spectrogram. Next, subsection 4.3 discussed the effect by using pre-processing of denoise function. Subsection 4.4 is focused on the feature extraction part of the data before underwent classification and subsection 4.5 is discussed all the LSTM results based on accuracy, consistency and also time processing. Lastly, the comparison between previous work will be discussed in subsection 4.6.

4.2 Visualization of CSI data signals

The original dataset of Wi-Fi signal of gait action is visualize using plotting method. Figure 4.1 showed the original dataset is plotted in term of CSI amplitude versus packet index. The data is large because it contained 90 channels that having 1000 CSI amplitude value each.

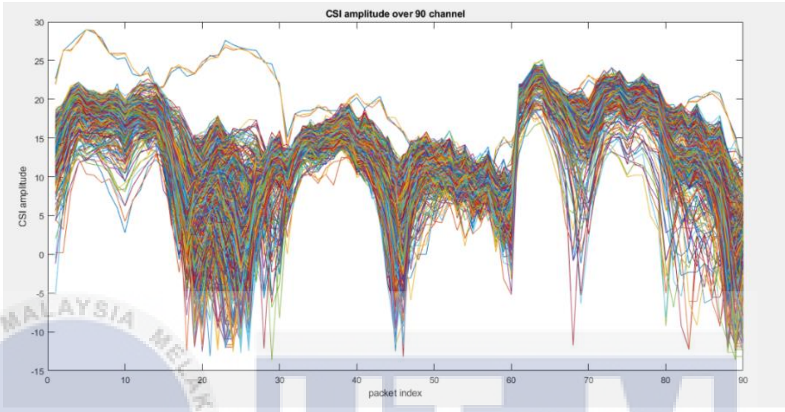


Figure 4.1: CSI amplitude over 90 channels.

Since the dataset is combining 3 antennas as receiver, the data was further distributed into 3 different antennas and different subcarrier so that it can be easily analyzed. Figure 4.2 showed the CSI amplitude of first subcarrier for 3 antennas whereas Figure 4.3 illustrates CSI phase of first subcarrier for 3 antennas.

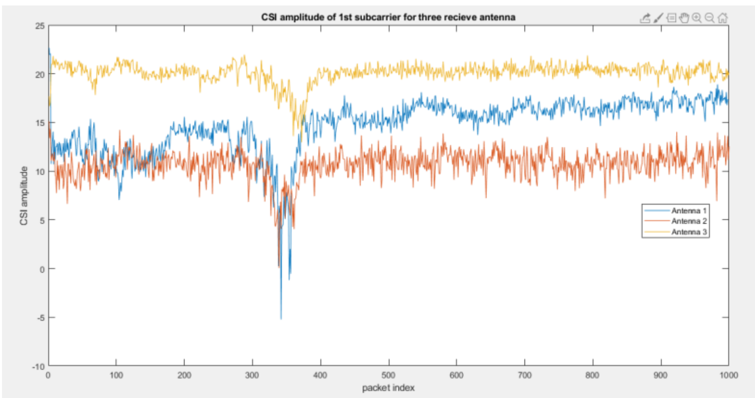


Figure 4.2: CSI amplitude of 1st subcarrier for 3 antennas.

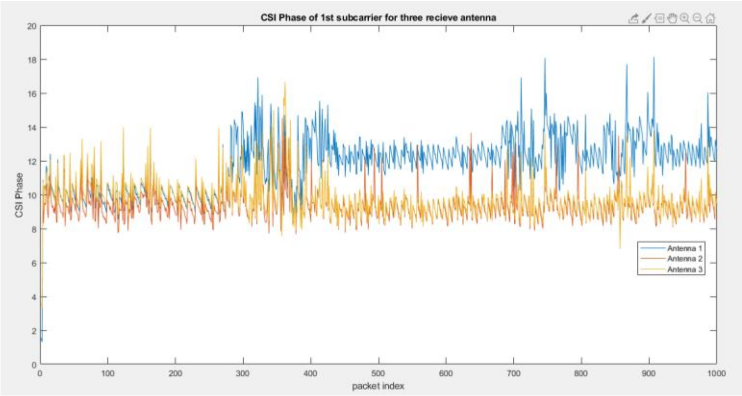
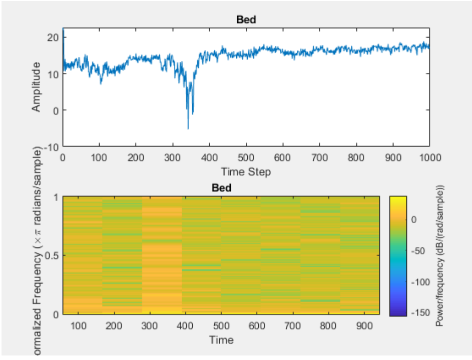
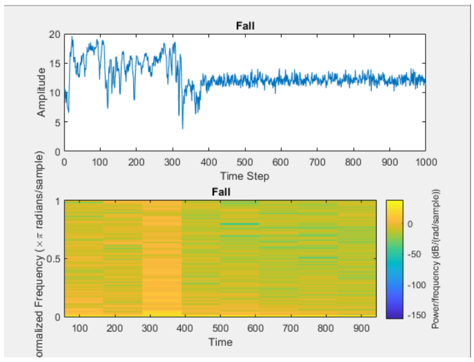


Figure 4.3: CSI phase of 1st subcarrier for 3 antennas.

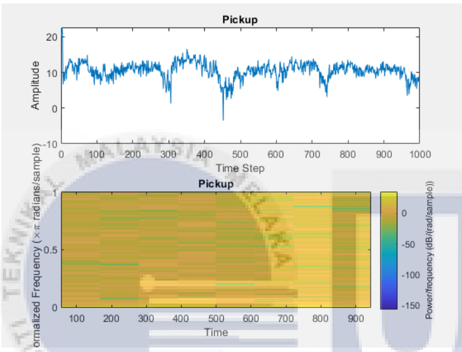
For further analysis, spectrogram function in MATLAB is used in order to determine the main feature for every human avtivity. Figure 4.4 illustrates the spectrogram results of each human activity. Based on the result, each human gait action had their own specified region that contain the significant feature among the time steps. Spectrogram allowed the significant features displayed in term of signal strength which indicated by yellow spot region. For example, action of lying on bed had its significant part around time between 300 and 400 seconds whereas the falling action located around 100 to 400 seconds. At this stage, the significant parts for all human activity are not clearly displayed yet due to the raw data signal itself contain a lot of noises and some interferences that caused by surrounding, Therefore, signal pre-processing of denoise function is needed to reduce the background noise so that the significant features can be shown out clearly by using spectrogram.


Bed
Amplitude
Time Step
Bed
ormalized Frequency (×π radians/sample)
Time
Power/frequency (dB/(rad/sample))

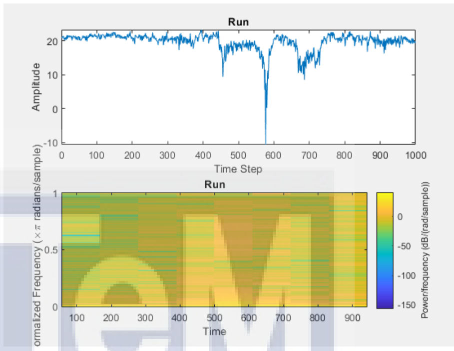
(a)


Fall
Amplitude
Time Step
Fall
ormalized Frequency (×π radians/sample)
Time
Power/frequency (dB/(rad/sample))

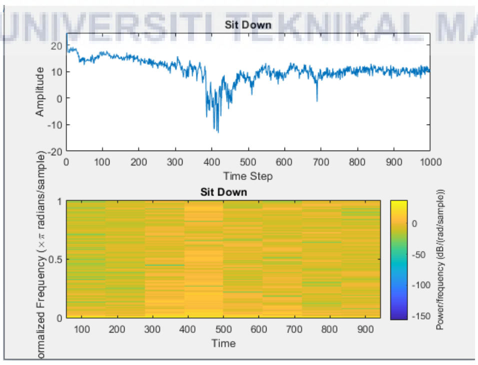
(b)


Pickup
Amplitude
Time Step
Pickup
ormalized Frequency (×π radians/sample)
Time
Power/frequency (dB/(rad/sample))

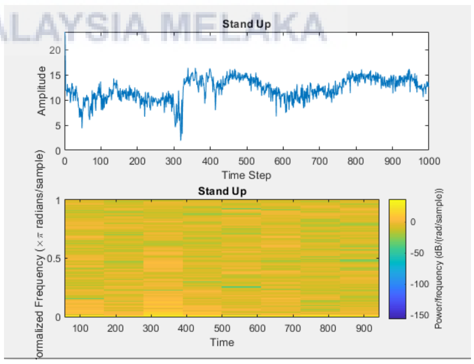
(c)


Run
Amplitude
Time Step
Run
ormalized Frequency (×π radians/sample)
Time
Power/frequency (dB/(rad/sample))

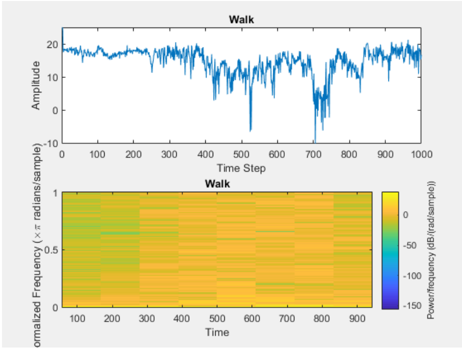
(d)


Sit Down
Amplitude
Time Step
Sit Down
ormalized Frequency (×π radians/sample)
Time
Power/frequency (dB/(rad/sample))

(e)


Stand Up
Amplitude
Time Step
Stand Up
ormalized Frequency (×π radians/sample)
Time
Power/frequency (dB/(rad/sample))

(f)



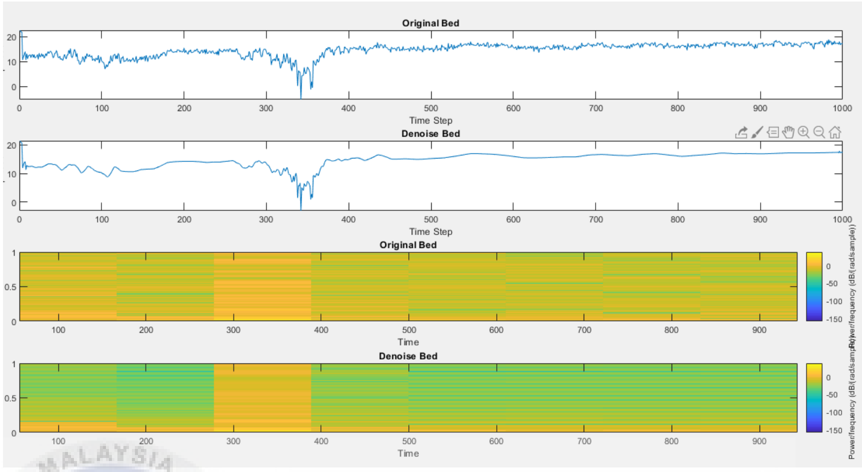
(g)

Figure 4.4: Spectrogram results for action (a) lying on bed, (b) falling, (c) pickup, (d) run, (e) sit down, (f) stand up and (g) walk.

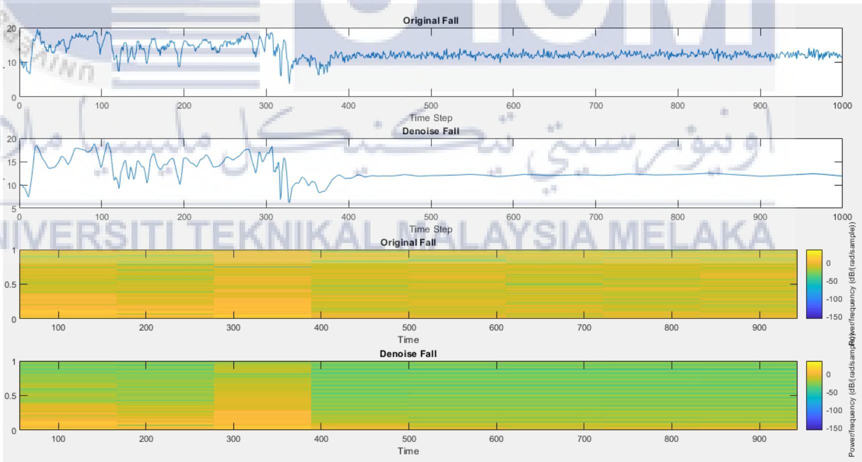
4.3 Denoise Pre-processing's Result

Denoise function is necessary for this project due to the raw datasets contain noises and other factor that caused interference by the surrounding environment. The denoise function in MATLAB is simply using command "wdenoise" with all default value such as 2dB for this project. This is because high dB value will reduce more detail information from original raw data. Since the significant features is still needed for further classification, default setting for command "wdenoise" is used in order to obtain higher accuracy during classification by LSTM.

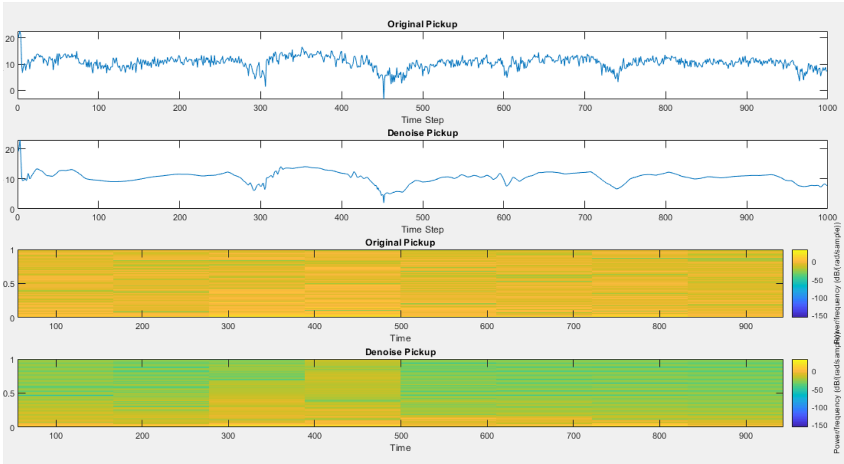
Figure 4.5 below shows the result of denoise signals with respect of their spectrogram results for each human activity. Based on the figure, it showed that the unnecessary noises and data has been removed by using denoise function. The leftover parts still containing significant features in which the indicated by yellow region of the spectrogram. Moreover, after denoising, the significant parts for each activity are clearly displayed out since the noise had been removed. For instance, the main feature part of lying on bed action can be spotted around 280 to 390 seconds whereas the action of falling allocated around 0 until 390 seconds. Besides, the signal strength of falling action is more compact during 270 until 390 seconds when refer to figure 4.5 (b).



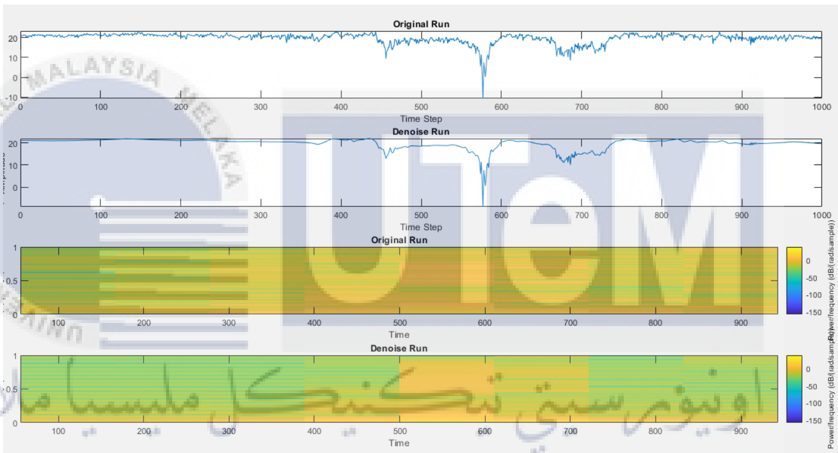
(a)



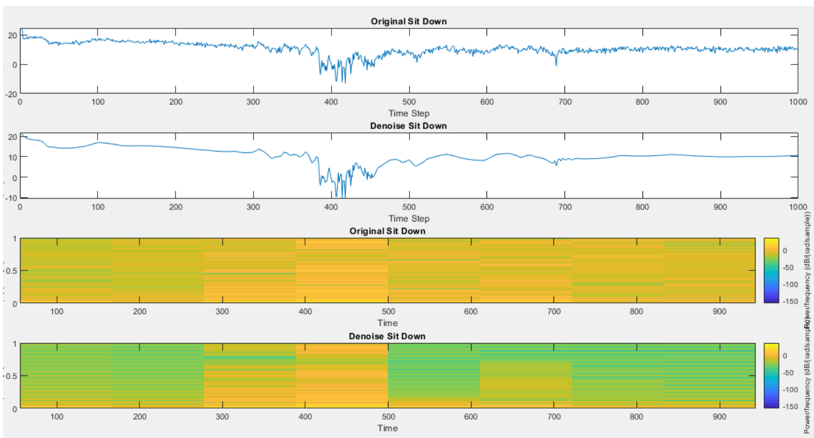
(b)


Original Pickup
Time Step
Denoise Pickup
Time Step
Original Pickup
Time
Denoise Pickup
Time
Power/frequency (dB/(rad/sample))

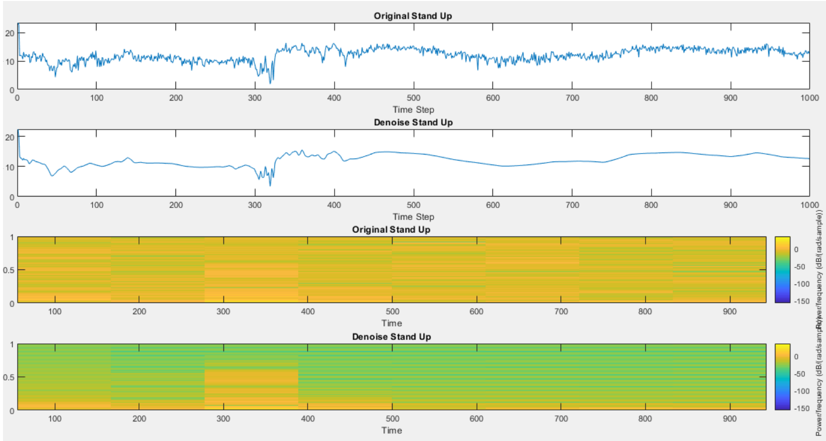
(c)


Original Run
Time Step
Denoise Run
Time Step
Original Run
Time
Denoise Run
Time
Power/frequency (dB/(rad/sample))

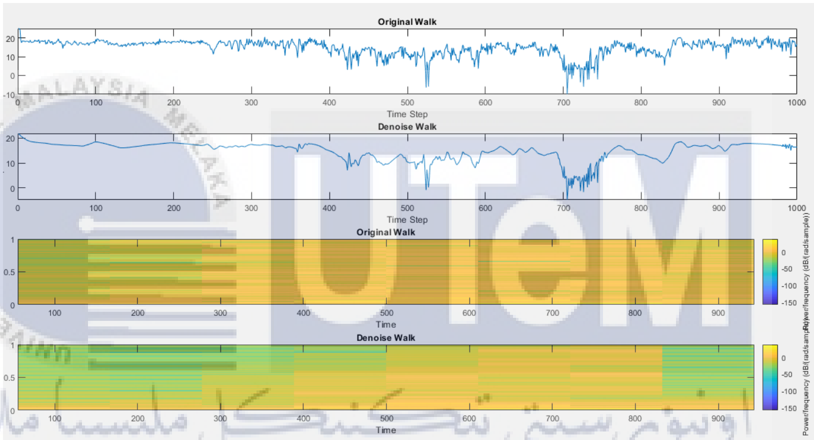
(d)


Original Sit Down
Time Step
Denoise Sit Down
Time Step
Original Sit Down
Time
Denoise Sit Down
Time
Power/frequency (dB/(rad/sample))

(e)



(f)



(g)

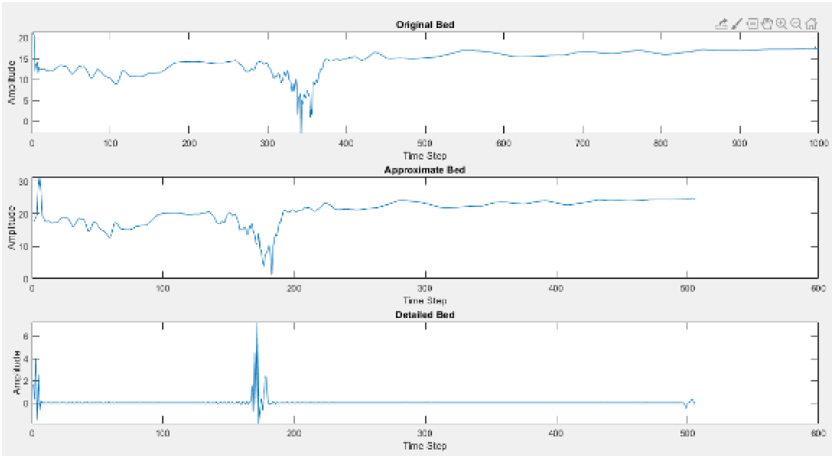
Figure 4.5: Result after denoise function of spectrogram for activities of (a) lying on bed, (b) falling, (c) pickup, (d) run, (e) sit down, (f) stand up and (g) walk.

All the significant feature is located in different time with different amplitude and various in term of signal strength. Hence, different human gait action can be classified easily since all activities will have different effect on CSI amplitudes.

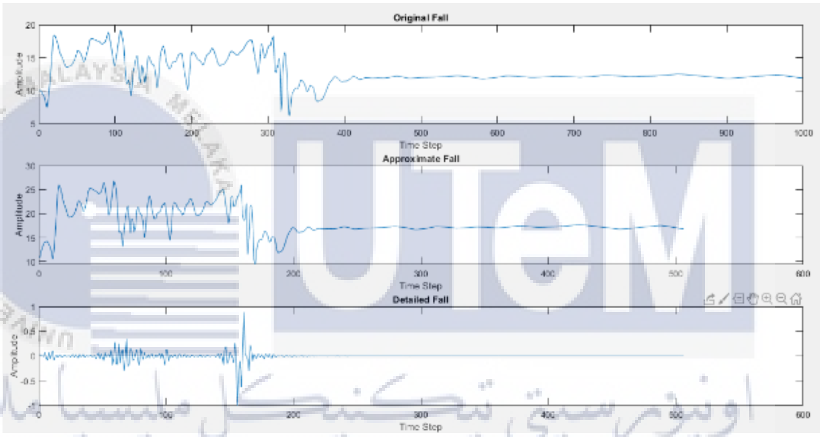
4.4 Performance of Discrete Wavelet Transform (DWT)

After denoised process of signal data, it shows some improvement in term of spectrogram since all the features can be displayed clearly. These data are then proceeded with feature extraction. As discussed in Chapter 3, feature extraction that had been used in this project is discrete wavelet transform (DWT). DWT will decompose two components named as approximated and detailed signals. A single level of discrete wavelet transform is used for this project by using command of "dwt". For other cases which required more level and detail feature can be used command of "wavedec" with specified certain decomposition level of DWT function.

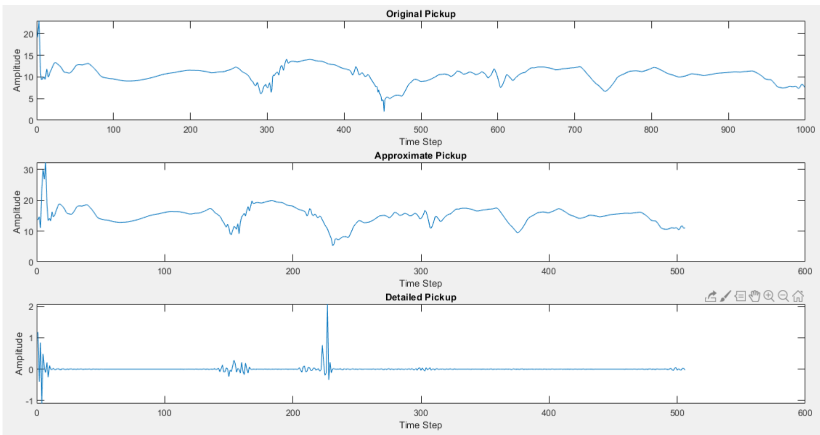
DWT is a method to reduce the dimension of the data while maintaining the significant feature of the data. A single level DWT is used due to not over reduced the dimension of original data. Figure 4.6 shown the result of DWT after denoise function in MATLAB. As observed, the dataset dimension had been reduced to half of its original dataset which is from 1000 to 501 times steps. Even though the dimension had been halved, but the significant parts are remained the same. Besides, the DWT provided decomposition of approximated and detailed signal data. The approximated signals data are used for further classification instead of detailed parts because detailed parts contained zero value which affected LSTM had low performance in term of accuracy and time processing. Moreover, detailed parts may have problem in over reduced the element of the dataset. Therefore, approximated signals are then used for classification by using LSTM algorithm.


Original Bed
Amplitude
Time Step
Approximate Bed
Amplitude
Time Step
Detailed Bed
Amplitude
Time Step

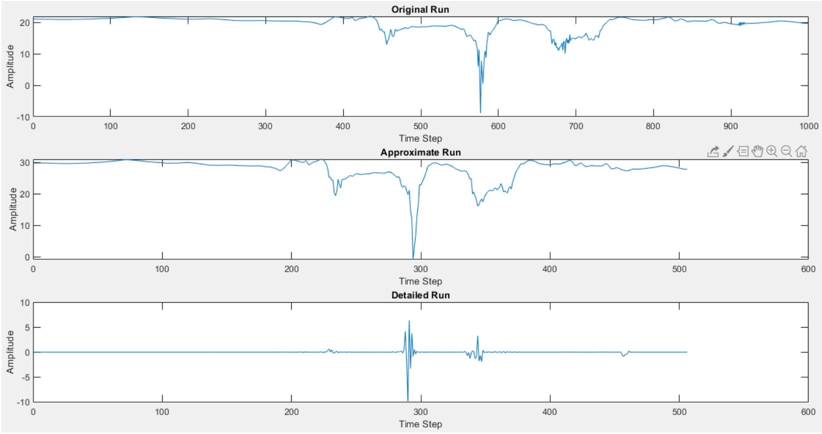
(a)


Original Fall
Amplitude
Time Step
Approximate Fall
Amplitude
Time Step
Detailed Fall
Amplitude
Time Step

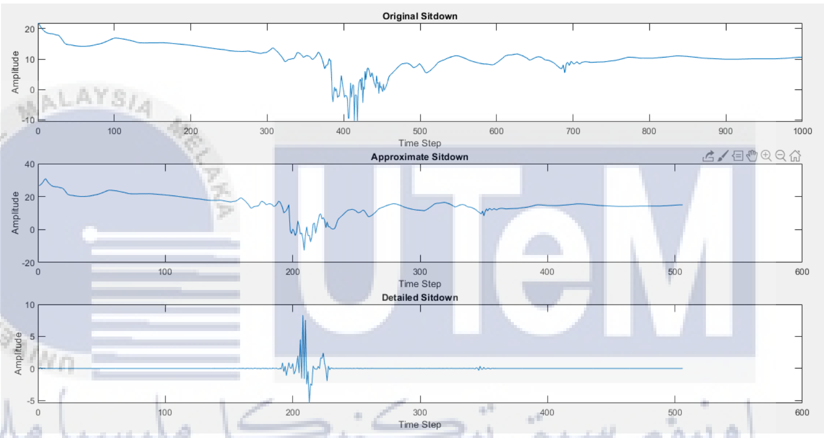
(b)


Original Pickup
Amplitude
Time Step
Approximate Pickup
Amplitude
Time Step
Detailed Pickup
Amplitude
Time Step

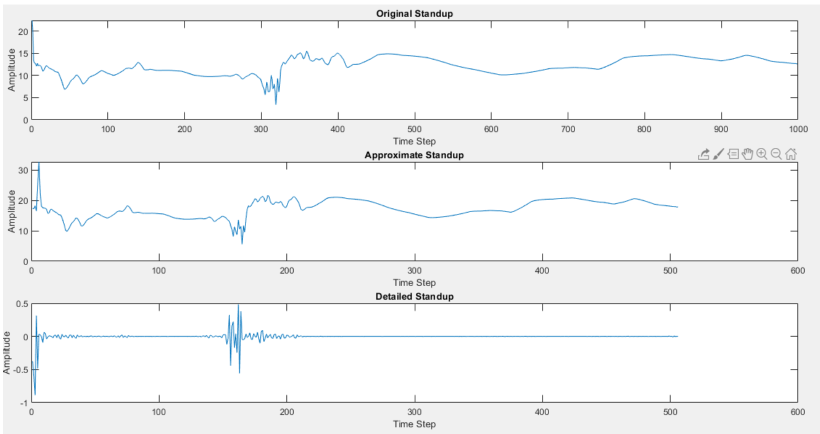
(c)


Original Run
Amplitude
Time Step
Approximate Run
Amplitude
Time Step
Detailed Run
Amplitude
Time Step

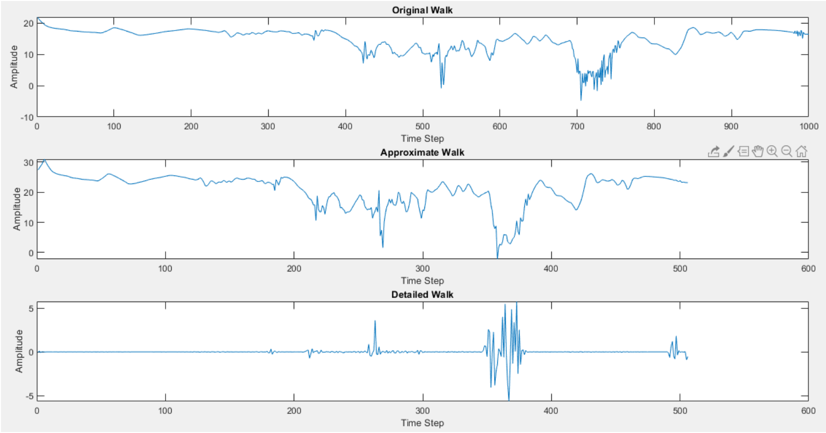
(d)


Original Sitdown
Amplitude
Time Step
Approximate Sitdown
Amplitude
Time Step
Detailed Sitdown
Amplitude
Time Step

(e)


Original Standup
Amplitude
Time Step
Approximate Standup
Amplitude
Time Step
Detailed Standup
Amplitude
Time Step

(f)



(g)

Figure 4.6: Result of one level discrete wavelet transform for activity (a) lying on bed, (b) falling, (c) pickup, (d) run, (e) sit down, (f) stand up and (g) walk.

4.5 LSTM Analysis

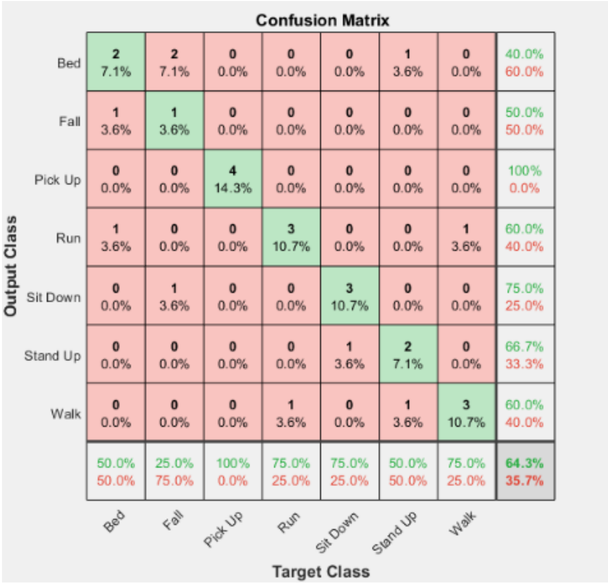
In this session, several parameters such as accuracy, consistency and time processing of LSTM are discussed. These parameters results can be obtained by using MATLAB in term of confusion matrix and training tables. As mentioned in Chapter 3, the proposed LSTM is Bi-LSTM which known as Bi-directional Long Short-Term Memory algorithm. Since Bi-LSTM can read the information from backward and forward direction, the accuracy of the classification should be higher in the end of the results.

4.5.1 LSTM's Accuracy

When all of the pre-processing and feature extraction are done, the modified CSI data signals are then underwent classification by using deep learning of LSTM. Figure 4.7 below demonstrated the result of LSTM's accuracy in term of confusion matrix. Raw datasets and modified datasets for has been undergo same classification of LSTM in order to compare the differences between their results. Based on the observation from the Figure, it showed that the first raw dataset and second raw dataset that without any feature extraction only achieved highest accuracy rate of 64.3% and 60.7% respectively. In other hand, the first and second dataset can achieve highest accuracy of 96.4% and 89.3% respectively.

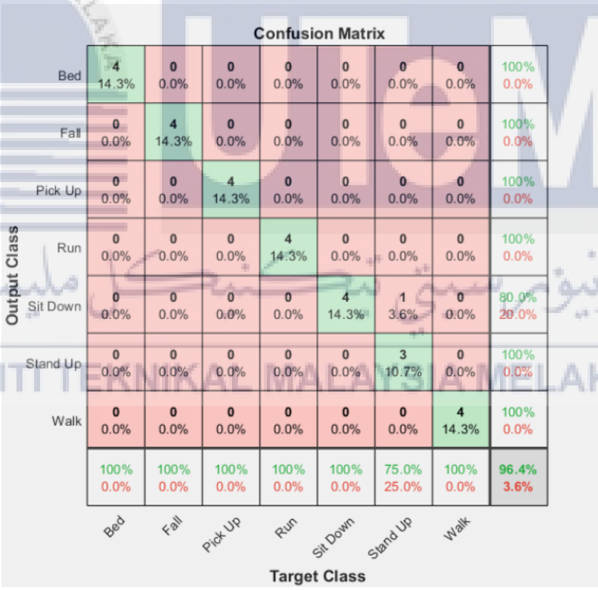
The raw datasets had lower accuracy rate because the raw datasets contain a lot of unnecessary information. That information caused the training progress cannot determined the human activity properly. Hence, the accuracy rate is decreased if just using raw data as input dataset. Meanwhile, the first modified dataset based on proposed method can perform well in LSTM algorithm due to the significant feature that represent the human activity has been extracted. Besides, the background noise of the raw datasets also eliminated during the process of denoising. Therefore, the first modified dataset can achieve greater accuracy rate.

However, second modified dataset can only achieve highest accuracy rate of 89.3%. This is because second dataset had larger dimension of dataset when compared to first dataset. The first dataset only consists of 90 x 1000 double data type whereas second dataset consist of 90 x 20k double data type. By comparing these two data sizes, second datasets had more complexity when compared to first datasets. Thus, the accuracy will not as high as first dataset when undergoing classification of LSTM.



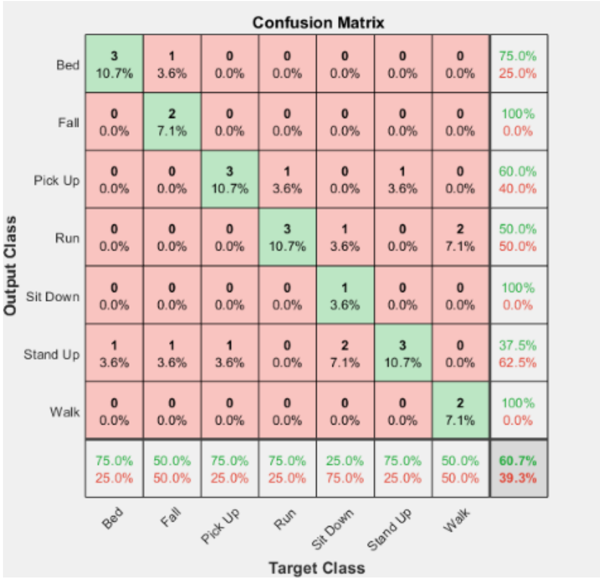
Confusion Matrix
Output Class
Bed
Fall
Pick Up
Run
Sit Down
Stand Up
Walk
Target Class

(a)

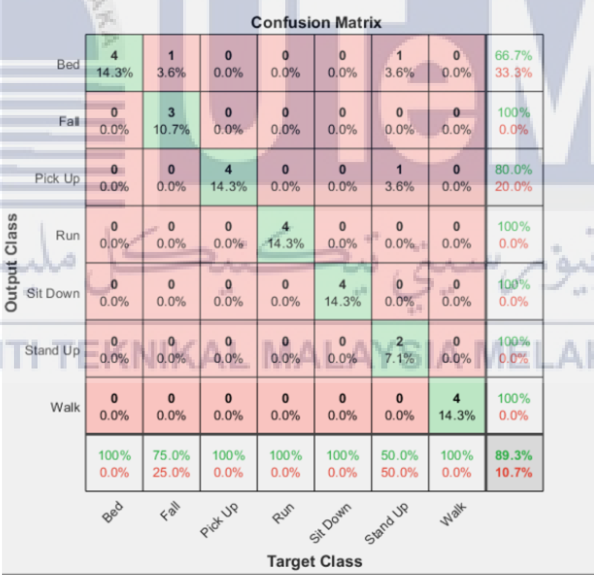


Confusion Matrix
Output Class
Bed
Fall
Pick Up
Run
Sit Down
Stand Up
Walk
Target Class

(b)



(c)



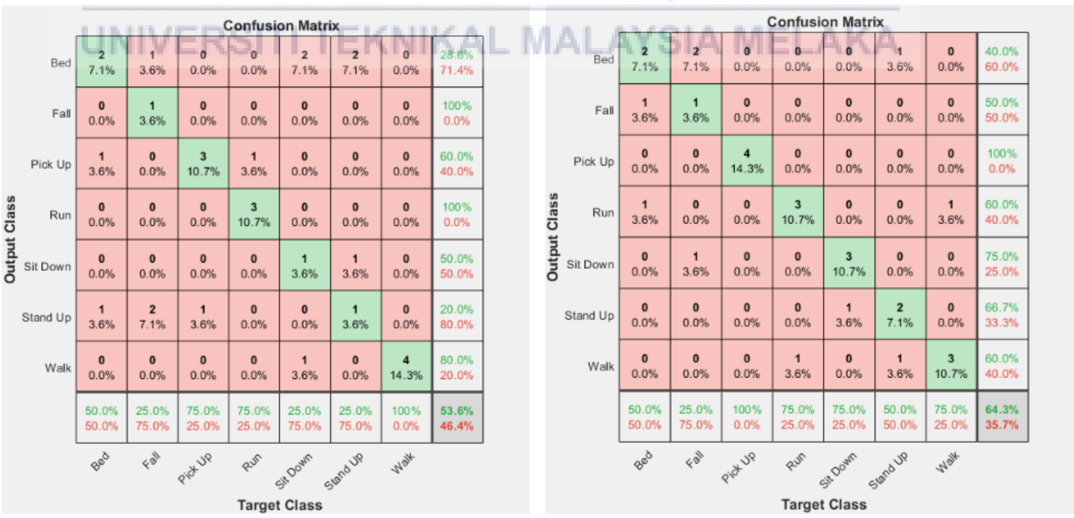
(d)

Figure 4.7: Confusion matrix result of LSTM for (a) first raw dataset, (b) first modified dataset, (c) second raw dataset, (d) second modified dataset.

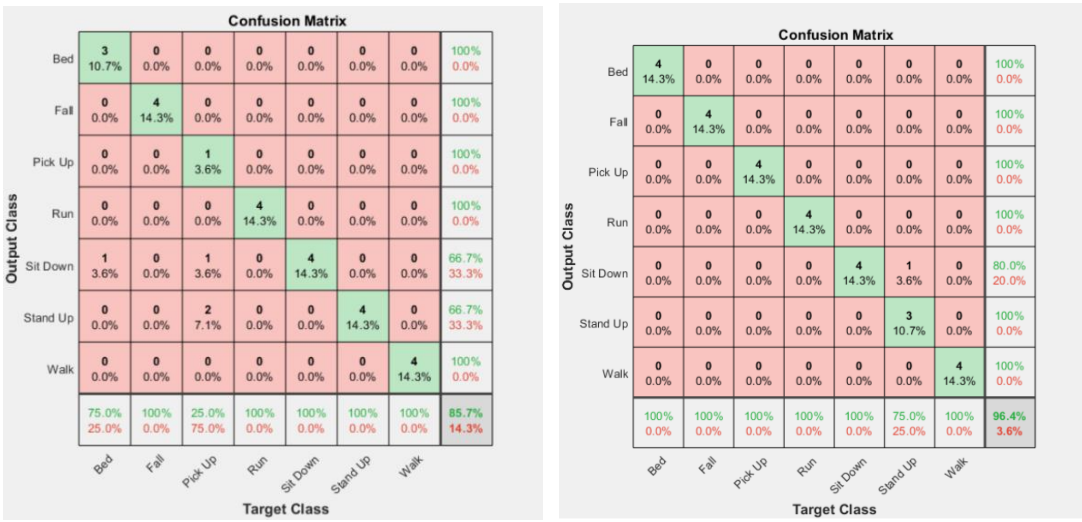
4.5.2 LSTM's Consistency

Due to the characteristic of LSTM algorithm that processed classification randomly, the accuracy rate will be different during each time of training and testing processes. Therefore, this session discussed the consistency of the LSTM's accuracy after running several times of classification. In the end of this session, the consistency range of each dataset that been achieved will be listed out.

After running multiple times of classification for raw data and modified datasets, Figure 4.8 shows the results of confusion matrix for each dataset in term of certain range. Based on the results, the consistency for first and second raw datasets is within the range of 53.6% to 64.3% and 42.9% to 60.7% respectively. For first modified dataset, the consistency is between 85.7% and 96.4% whereas second modified datasets achieved consistency of 75% to 89.3% accuracy rate. As observed, LSTM will produce different accuracy rate since the LSTM is distributed training and testing subset randomly during each time of classification process.



(a)


Confusion Matrix
Output Class
Bed
Fall
Pick Up
Run
Sit Down
Stand Up
Walk
Target Class
Confusion Matrix
Output Class
Target Class

(b)


Confusion Matrix
Output Class
Bed
Fall
Pick Up
Run
Sit Down
Stand Up
Walk
Target Class
Confusion Matrix
Output Class
Target Class

(c)

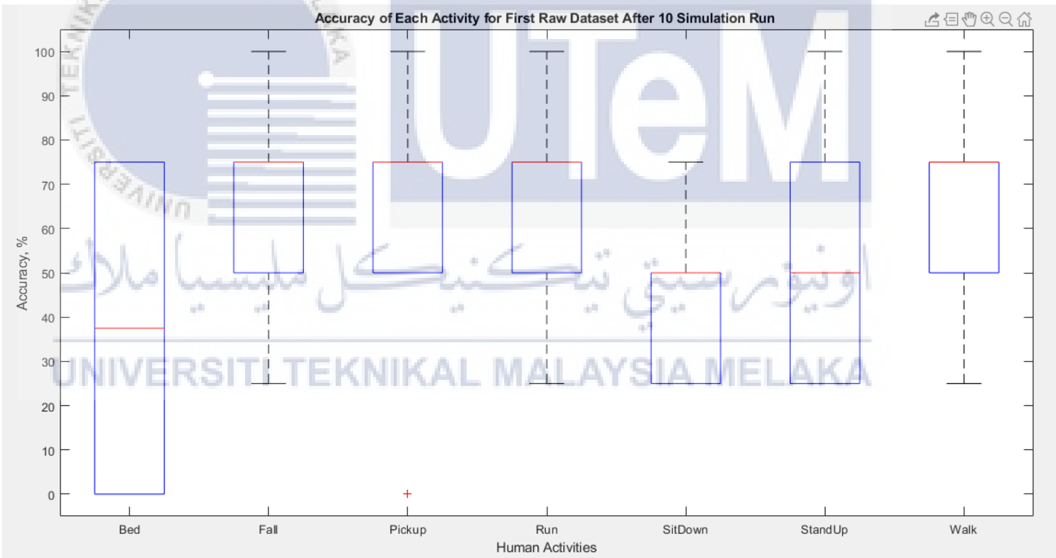
(d)

Figure 4.8: Consistency of accuracy rate for (a) raw data, (b) first modified dataset, (c) second raw dataset, (d) second modified dataset.

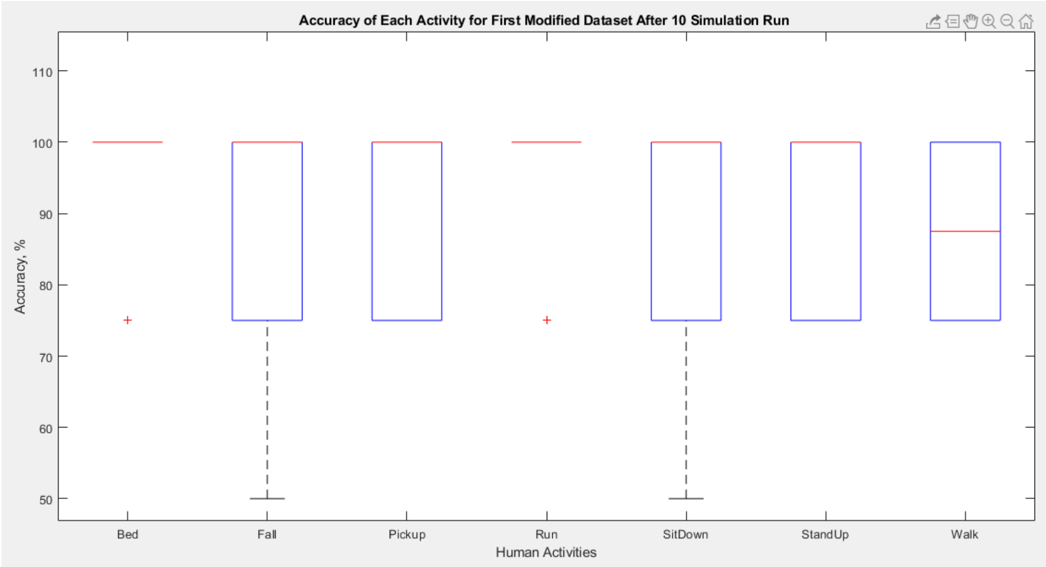
For further analyzing, the boxplot graph is used to identify the accuracy of each activity after 10 simulations ran in MATLAB as shows in Figure 4.9. Boxplot graph is a method to visualize the summary statistics for certain dataset. On each box, the central line represented the median of data. The top and bottom edge of the box is indicated as the 75th and 25th percentiles of data, respectively. Then, the whiskers line extended is denoted as the maximum data point whereas the outlier is expressed as symbol of "+".

Based on the Figure 4.9 (a), the activity of bed achieved the lowest median accuracy among all other activities even though the maximum accuracy it can achieve is 75%. Therefore, the overall accuracy is dropped due to action of bed has lower accuracy after 10 simulations ran. Meanwhile, the second raw dataset has lowest median for activity of sitting down at 25% which affected overall accuracy has been decrease during each classification of LSTM.

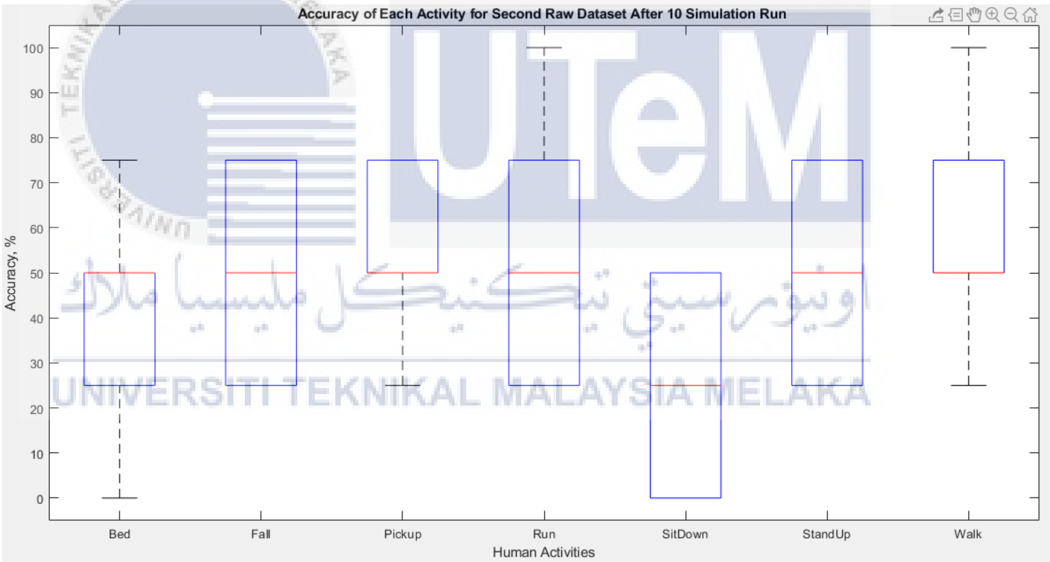
For first modified dataset, the median for most of the actions have achieved 100% accuracy except walking action. The activity of running and lying on bed have few outliers in the boxplot graph. This is because all the results are mostly 100% accuracy which caused the graphs have no boxes but a median line only. In term of consistency, the first modified dataset is mostly maintained at range of 75% to 100%. Hence, the first modified dataset can attain higher accuracy when compared to other datasets. As observed, the second modified dataset has retained consistency between 50% and 100% after 10 simulations ran. The accuracy of second modified dataset lower because of standing up action has lower median value which is 60% only. Therefore, it can only achieve 89.2% which considered lower than the first modified dataset.



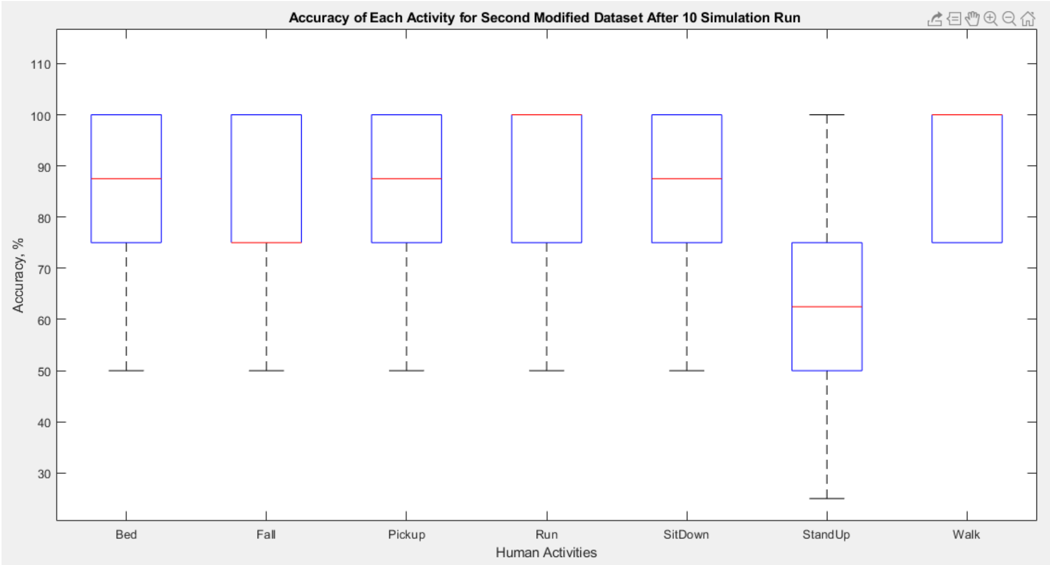
(a)


Accuracy of Each Activity for First Modified Dataset After 10 Simulation Run
Accuracy, %
110
100
90
80
70
60
50
Bed
Fall
Pickup
Run
SitDown
StandUp
Walk
Human Activities

(b)


Accuracy of Each Activity for Second Raw Dataset After 10 Simulation Run
Accuracy, %
100
90
80
70
60
50
40
30
20
10
0
Bed
Fall
Pickup
Run
SitDown
StandUp
Walk
Human Activities

(c)



(d)

Figure 4.9: Boxplot graph of accuracy of each activity after 10 simulations run for (a) raw data, (b) first modified dataset, (c) second raw dataset, (d) second modified dataset.

Even though Bi-LSTM can read data from backward and forward direction of dataset, but consistency will be dropped due to the way of datasets collection. First and second datasets are collected data from three different antennas. Each antenna had its own characteristics in term of different range as shown in Figure 4.2. Moreover, both datasets contained few different persons walking gait action profiles which will reduce the accuracy and consistency of the classification results. This is because different person will have different characteristics walking gait action profiles. At the same time, the distribution of training and testing subset is randomly. Hence, the LSTM will have some loss in term of accuracy and consistency when undergoing classification every time.

4.5.3 LSTM's Time Processing Analysis

In term of time processing of LSTM, time taken to process is very crucial since this aspect is highly demanded in any data processing. Therefore, this session discussed the time processing of LSTM for each dataset. Figure 4.9 below demonstrated the time taken for each dataset when undergoing same training parameter. It showed that the time elapsed of first raw dataset is 13 minutes 23 seconds while the time elapsed for second raw dataset is 12 minutes 55 seconds. However, the first and second modified dataset achieved time elapsed of 6 minutes 25 seconds and 7 minutes 31 seconds respectively.

Different training parameters will generate different output of classification in term of time elapsed and accurate rate of the data. Therefore, the training parameter is standardized into same parameters as shown in figure so that it can be compared easily among the datasets. The results showed that the time elapsed for modified datasets had reduced successfully due to the fact of dimensionality reduction by the DWT in feature extraction process. When the data size is small, the processing of the classification will be reduced at the same time.

Among those training parameters, the important parameters such as “MaxEpoch”, “MiniBatchSize” and “InitialLearnRate” will affect the output results. An epoch is the full pass of the training algorithms over entire training set. Based on the result in Figure, as the epoch increased, the training performance also increased. But, higher in epoch value is not necessarily generated better performance due to it depend on the characteristic of the dataset. Mini batch is subset of training set that used to evaluate gradient loss of function. Better performance can be achieved if a smaller batch size is used. Learning rate refer to the amount of weight that are updated during training process. If the learning rate is too low, it will take longer time to process. In other hands, the training might reach suboptimal result if it is too high.

Epoch	Iteration	Time Elapsed (hh:mm:ss)	Mini-batch Accuracy	Mini-batch Loss	Base Learning Rate
1	1	00:00:05	25.00%	2.1826	0.0010
17	50	00:03:15	93.75%	0.8369	0.0010
34	100	00:06:26	100.00%	0.1633	0.0010
50	150	00:09:37	100.00%	0.1079	0.0010
67	200	00:12:45	100.00%	0.0517	0.0010
70	210	00:13:23	100.00%	0.0303	0.0010

(a)

Epoch	Iteration	Time Elapsed (hh:mm:ss)	Mini-batch Accuracy	Mini-batch Loss	Base Learning Rate
1	1	00:00:03	0.00%	2.0110	0.0010
17	50	00:01:38	96.88%	0.5612	0.0010
34	100	00:03:07	81.25%	0.7219	0.0010
50	150	00:04:35	100.00%	0.3155	0.0010
67	200	00:06:07	100.00%	0.0940	0.0010
70	210	00:06:25	100.00%	0.1519	0.0010

(b)

Epoch	Iteration	Time Elapsed (hh:mm:ss)	Mini-batch Accuracy	Mini-batch Loss	Base Learning Rate
1	1	00:00:05	13.33%	2.1693	0.0010
17	50	00:03:06	86.67%	0.8043	0.0010
34	100	00:06:13	100.00%	0.1759	0.0010
50	150	00:09:16	100.00%	0.1058	0.0010
67	200	00:12:19	100.00%	0.0705	0.0010
70	210	00:12:55	100.00%	0.0484	0.0010

(c)

Epoch	Iteration	Time Elapsed (hh:mm:ss)	Mini-batch Accuracy	Mini-batch Loss	Base Learning Rate
1	1	00:00:05	12.50%	2.1377	0.0010
17	50	00:01:50	90.63%	0.6337	0.0010
34	100	00:03:36	96.88%	0.3936	0.0010
50	150	00:05:21	100.00%	0.1367	0.0010
67	200	00:07:09	100.00%	0.0642	0.0010
70	210	00:07:31	100.00%	0.0529	0.0010

(d)

Figure 4.10: Time elapsed for training parameters for (a) raw data, (b) first modified dataset, (c) second raw dataset, (d) second modified dataset.

4.6 Comparison with Previous Work

Table 4.1 below shown the accuracy comparison between proposed method and previous works. By using proposed method, signal pre-processing and feature extraction are implemented right before classification of LSTM. Those method can reduce dimensionality of the original dataset which will achieve better performance in accuracy and time elapsed. Moreover, the proposed LSTM method is Bi-LSTM which the input data can run in two ways, forward and backward while preserved the information for both past and future. For second dataset, the dimensionality of data is much higher when compared to first dataset. This caused the data had higher complexity which leaded to poor performance in LSTM classification. In fact, the LSTM is a deep learning-based classification method that do not required any pre-processing and feature extraction method. This is due to the reason of LSTM can extract and analyze the feature automatically. However, the performance is not that well as shown in subsection 4.5.1. Therefore, proposed method is suggested so that the entire system can manage to achieve higher accuracy and time processing when underwent classification of LSTM.

Table 4.1: Accuracy comparison table between previous works.

Topics	Method	Accuracy	Time Elapsed
WiWho [23]	Decision Tree	80% – 92%	-
WifiU [24]	SVM	79% – 93%	-
WiFi- ID [25]	CWT + SAC	77% – 93%	-
EI [26]	CNN	75%	-
E-eyes [27]	EMD + MD-DTW	96%	-
CARM [28]	HMM	96%	-
Survey [29]	LSTM	90.5%	-
First raw dataset [30]	DWT + Bi-LSTM	53.6% – 64.3%	13m23s
First modified dataset [30]	DWT + Bi-LSTM	85.7% – 96.4%	06m25s
Second raw dataset [32]	DWT + Bi-LSTM	42.9% – 60.7%	12m55s
Second modified dataset [32]	DWT + Bi-LSTM	75% – 89.3%	07m31s

4.7 Chapter Summary

In the first part of this chapter, the CSI data signal is visualized and analyzed in detailed manner in term of graph and spectrogram. Next, the denoised signal data also displayed out clearly and the significant feature can be analyzed easily based on spectrogram. The third part of chapter discussed about performance of DWT that can reduce dimensionality of dataset while maintaining feature part of data. Then, the analysis of LSTM in term of accuracy, consistency and also time processing had been analyzed and explained in detail manner. Lastly, the results are compared comprehensively with previous works in form of table.

CHAPTER 5

CONCLUSION AND FUTURE WORKS

5.1 Introduction

In this chapter, the conclusion of results and discussion of this project will be summarized. The recommendation of this project will also be provided in order to improve the quality of the project.

5.2 Conclusion

In conclusion, a deep learning-based recognition system for improved Wi-Fi based human activity recognition system is successfully developed by using MATLAB. All the signal processing is conducted by using datasets that obtained from the internet. Next, the raw data signal contained a lot of unnecessary data in which the data is needed pre-processing such as denoise. Even though the denoise function helped a lot in term of filtered noises, but it still not enough when come to detect the significant feature of the datasets. Therefore, discrete wavelet transform (DWT) is used as feature extraction method. In order to determine the feature parts of signals, spectrogram also used to display the main feature for each activity. The significant is indicated in term of signal strength color bar as shown in Chapter 4. DWT helped in dimensionality reduction while still maintaining the significant parts of the signals in datasets. Bi-LSTM is proposed since this method can run input data in two ways, from forward to backward and from backward to forward. This method can enhance the performance of classification because it will still be preserved the data for past and future even though it ran the input data in two ways. Accuracy rate for first modified dataset is achieved around 85.7% to 96.4% whereas accuracy between 75% and 89.3% is achieved by using second modified dataset. Since second modified dataset had larger dimensionality in data size, the performance will not as good as the first modified dataset. Moreover, in term of time processing, both first and second modified datasets had attained better time elapsed during classification of LSTM. This is because the raw data do not undergo any pre-processing and feature extraction in which it leaded to increase the time elapsed for the LSTM classification.

5.3 Recommendation

In this project, there is no data collection and hardware design. Therefore, the dataset can only obtain from the internet. The dataset from the internet is nonlinear among the three different antennas which will increase the complexity of the datasets. Those nonlinearly datasets not just affected the accuracy rate of classification but also increased the process to fix the linearity problem that existed inside the dataset. Hence, it is suggested that maintaining the linearity during collection of data or reduced the antennas of Wi-Fi receivers so that the linearity of datasets can be maintained.

Besides, the datasets also contained the phase of Wi-Fi signals. For further enhancement in accuracy rate, the phase of Wi-Fi signal can represent the main feature for each walking gait action. Several approach that relevant to phase different and phase change can be implemented together with the amplitude of the Wi-Fi signal might also helped in generate better performance during classification of LSTM.

As a conclusion, the deep learning-based human gait recognition system in this project holds a great potential in current time. With the rise of 5G technology, the usage of Wi-Fi system can be amplified. More work can be done in improving the system so that it is possible to be implemented on all vehicles. Especially during this pandemic of COVID-19 because this system does not require close contact with other people or object.

REFERENCES

[1] M. Vrigkas, C. Nikou, and I. A. Kakadiaris, “A review of human activity recognition methods,” *Front. Robot. AI*, vol. 2, no. NOV, pp. 1–28, 2015, doi: 10.3389/frobt.2015.00028.

[2] S. Ranasinghe, F. Al Machot, and H. C. Mayr, “A review on applications of activity recognition systems with regard to performance and evaluation,” *Int. J. Distrib. Sens. Networks*, vol. 12, no. 8, p. 155014771666552, Aug. 2016, doi: 10.1177/1550147716665520.

[3] W. Zhang, C. Su, and C. He, “Rehabilitation exercise recognition and evaluation based on smart sensors with deep learning framework,” *IEEE Access*, vol. 8, pp. 77561–77571, 2020, doi: 10.1109/ACCESS.2020.2989128.

[4] H. Li, X. He, X. Chen, Y. Fang, and Q. Fang, “Wi-Motion: A Robust Human Activity Recognition Using WiFi Signals,” *IEEE Access*, vol. 7, pp. 153287–153299, 2019, doi: 10.1109/ACCESS.2019.2948102.

[5] Z. Hao, Y. Duan, X. Dang, and T. Zhang, “CSI-HC: A WiFi-Based Indoor Complex Human Motion Recognition Method,” *Mob. Inf. Syst.*, vol. 2020, 2020, doi: 10.1155/2020/3185416.

[6] Y. Ma, G. Zhou, and S. Wang, “WiFi sensing with channel state information:

A survey," *ACM Comput. Surv.*, vol. 52, no. 3, 2019, doi: 10.1145/3310194.

[7] O. C. Ann and L. B. Theng, "Human activity recognition: A review," *Proc. - 4th IEEE Int. Conf. Control Syst. Comput. Eng. ICCSCE 2014*, no. November, pp. 389–393, 2014, doi: 10.1109/ICCSCE.2014.7072750.

[8] Z. Hussain, M. Sheng, and W. E. Zhang, "Different approaches for human activity recognition- A survey," *arXiv*, pp. 1–28, 2019.

[9] R. Poppe, "A survey on vision-based human action recognition," *Image Vis. Comput.*, vol. 28, no. 6, pp. 976–990, 2010, doi: 10.1016/j.imavis.2009.11.014.

[10] G. R. Bradski and J. Davis, "Motion segmentation and pose recognition with motion history gradients," *Proc. IEEE Work. Appl. Comput. Vis.*, vol. 2000-Janua, no. 2002, pp. 238–244, 2000, doi: 10.1109/WACV.2000.895428.

[11] M. Piccardi, "Background subtraction techniques: A review," *Conf. Proc. - IEEE Int. Conf. Syst. Man Cybern.*, vol. 4, pp. 3099–3104, 2004, doi: 10.1109/ICSMC.2004.1400815.

[12] G. Tesauro, "Practical issues in temporal difference learning," *Mach. Learn.*, vol. 8, no. 3–4, pp. 257–277, 1992, doi: 10.1007/bf00992697.

[13] R. T. Collins and T. Kanade, *Introduction to the Special Section on Video Surveillance*, vol. 22, no. 8. 2000.

[14] R. Cutler and L. Davis, "Real-time periodic motion detection, analysis, and applications," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2, no. 8, pp. 326–332, 1999, doi: 10.1109/cvpr.1999.784652.

[15] J. Fan, W. Xu, Y. Wu, and Y. Gong, "Human tracking using convolutional neural networks," *IEEE Trans. Neural Networks*, vol. 21, no. 10, pp. 1610–1623, 2010, doi: 10.1109/TNN.2010.2066286.

[16] P. V. K. Borges, N. Conci, and A. Cavallaro, "Video-based human behavior understanding: A survey," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 23, no. 11, pp. 1993–2008, 2013, doi: 10.1109/TCSVT.2013.2270402.

[17] J. Wu, A. Osuntogun, T. Choudhury, M. Philipose, and J. M. Rehg, "A scalable approach to activity recognition based on object use," *Proc. IEEE Int. Conf. Comput. Vis.*, 2007, doi: 10.1109/ICCV.2007.4408865.

[18] W. Wang, A. X. Liu, M. Shahzad, K. Ling, and S. Lu, "Device-Free Human Activity Recognition Using Commercial WiFi Devices," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 5, pp. 1118–1131, 2017, doi: 10.1109/JSAC.2017.2679658.

[19] M. Mohamed and M. Cheffena, "Received signal strength based gait authentication," *IEEE Sens. J.*, vol. 18, no. 16, pp. 6727–6734, 2018, doi: 10.1109/JSEN.2018.2850908.

[20] Z. Yang, Z. Zhou, and Y. Liu, "From RSSI to CSI," *ACM Comput. Surv.*, vol. 46, no. 2, pp. 1–32, 2013, doi: 10.1145/2543581.2543592.

[21] K. Kasantikul, C. Xiu, D. Yang, and M. Yang, "An enhanced technique for indoor navigation system based on WIFI-RSSI," *Int. Conf. Ubiquitous Futur. Networks, ICUFN*, vol. 2015-Augus, no. I, pp. 513–518, 2015, doi: 10.1109/ICUFN.2015.7182597.

[22] W. Shuaieb *et al.*, "RFID RSS fingerprinting system for wearable human activity recognition," *Futur. Internet*, vol. 12, no. 2, pp. 1–12, 2020, doi: 10.3390/fi12020033.

[23] Y. Zeng, P. H. Pathak, and P. Mohapatra, "WiWho: WiFi-Based Person Identification in Smart Spaces," *2016 15th ACM/IEEE Int. Conf. Inf. Process. Sens. Networks, IPSN 2016 - Proc.*, 2016, doi: 10.1109/IPSN.2016.7460727.

[24] W. Wang, A. X. Liu, and M. Shahzad, "Gait recognition using WiFi signals," *UbiComp 2016 - Proc. 2016 ACM Int. Jt. Conf. Pervasive Ubiquitous Comput.*, pp. 363–373, 2016, doi: 10.1145/2971648.2971670.

[25] J. Zhang, B. Wei, W. Hu, and S. S. Kanhere, "WiFi-ID: Human identification using WiFi signal," *Proc. - 12th Annu. Int. Conf. Distrib. Comput. Sens. Syst. DCOSS 2016*, pp. 75–82, 2016, doi: 10.1109/DCOSS.2016.30.

[26] W. Jiang *et al.*, "Towards environment independent device free human activity recognition," *Proc. Annu. Int. Conf. Mob. Comput. Networking, MOBICOM*, pp. 289–304, 2018, doi: 10.1145/3241539.3241548.

[27] Y. Wang, J. Liu, Y. Chen, M. Gruteser, J. Yang, and H. Liu, "E-eyes: Device-free location-oriented activity identification using fine-grained WiFi signatures," *Proc. Annu. Int. Conf. Mob. Comput. Networking, MOBICOM*, pp. 617–628, 2014, doi: 10.1145/2639108.2639143.

[28] W. Wang, A. X. Liu, M. Shahzad, K. Ling, and S. Lu, "Understanding and modeling of WiFi signal based human activity recognition," *Proc. Annu. Int. Conf. Mob. Comput. Networking, MOBICOM*, vol. 2015-Septe, pp. 65–76,

2015, doi: 10.1145/2789168.2790093.

[29] S. Yousefi, H. Narui, S. Dayal, S. Ermon, and S. Valaee, "A Survey on Behavior Recognition Using WiFi Channel State Information," *IEEE Commun. Mag.*, vol. 55, no. 10, pp. 98–104, 2017, doi: 10.1109/MCOM.2017.1700082.

[30] "LSTM_wifi_activity_recognition/161208_activity_data at master · Hirokazu-Narui/LSTM_wifi_activity_recognition · GitHub." https://github.com/Hirokazu-Narui/LSTM_wifi_activity_recognition/tree/master/161208_activity_data (accessed May 28, 2021).

[31] "matlab-icon -3." http://pafkiet.edu.pk/pdc/matlab-icon-3/ (accessed Jan. 12, 2021).

[32] "GitHub - ermongroup/Wifi_Activity_Recognition: Code for IEEE Communication Magazine (A Survey on Behaviour Recognition Using WiFi Channle State Information)." https://github.com/ermongroup/Wifi_Activity_Recognition (accessed Jun. 10, 2021).

[33] A. K. Barros and N. Ohnishi, "Heart instantaneous frequency (HIF): An alternative approach to extract heart rate variability," *IEEE Trans. Biomed. Eng.*, vol. 48, no. 8, pp. 850–855, 2001, doi: 10.1109/10.936361.

[34] F. A. Furfari(tony), "The Transformer," *IEEE Ind. Appl. Mag.*, vol. 8, no. 1, pp. 8–15, 2002, doi: 10.1109/2943.974352.