**CHAPTER - 1**

**INTRODUCTION**

**1.1 Motivation and Goal:**

With the advancement of technology, machine learning and deep learning methods are being applied in numerous fields. Among these Human activity recognition (HAR) has been the one of the most significant topics in the active research field. Human activity recognition intends to monitor, recognize the activity of a person based on series of observations and surrounding environment. In easy words, HAR technology is used in recognizing human daily activities, from simple activities like standing, sitting, walking upstairs, running to complex activities like cooking while standing, watching TV while sitting or lying [1].

Human activity recognition, due to availability of devices and wearable sensors in low cost has become an integral part of people’s daily lives and is being applied broadly in common domains like health management - elderly monitoring, disease prevention, rehabilitation, in the idea of smart cities – domestic activity monitoring. Furthermore, HAR is applied in security concerns like through individual activity monitoring solutions, crowd anomaly detection. Besides, wearable and inertial sensors combined with embedded systems are being used in sports activities [2,3].

In this project our goal is to develop a model to recognize daily and simple human activities real time. For simplicity our goal is to recognize 6 activities and we used 2 types of sensor data for this (3-axis accelerometer and 3-axis gyroscope).

**1.2 Approaches to human activity recognition**

In order to recognize human activities, different methods maybe applied. Methods and way of approaching may vary from, type of collecting data and data processing to detection approaches. Based on collecting activity data, HAR can be divided into two types:

1. Computer Vision based
2. Sensor based

**1.2.1 Computer Vision based approach**

In computer vision based approach, images and videos are collected by means of optical sensors like cameras, CCTVs and then captured images or videos are analyzed. Vision based data being affordable and collectible hardly with any trouble, vast majority of researches have been conducted on vision based HAR [4]. Besides, contact based like wearable sensor based approach sometimes requires sophisticated equipment and also it has to be affordable in cost, correct size and also the user’s acceptability. Moreover, vision based system will not require the user to wear devices uncomfortable to them in different parts of the body. In this type of approach mainly three stages are followed which includes – 1. detection (first stage), which determines the part of the body to recognize or follow (methods like using skin color, shape, pixel values etc. are used), 2. Tracking (second stage), where link between successive images are provided (methods like feature tracking, contour tracking, optimal estimation etc. are used) and 3. Classification, which is the final stage, and where different machine learning and deep learning algorithms are used to finally recognize the activity [5].

**1.2.2 Sensor based approach**­

As mentioned, human activity recognition can be carried out in a sensor based approach. Sensors of different kinds are used in data acquisition for the recognition to be performed. Sensors can be integrated in a device or can be used separately. Based on the platforms used, sensors can be:

1. Wearable Sensors
2. Smartphone Sensors

**(a) Wearable Sensors:**

This type of sensors is used for only one particular purpose/function. These can’t take any other measurements other than the ones these are made for. Usually these sensors are integrated into a device used for the given task only. Wearable sensors can further be classified into more types – *Inertial sensors* include accelerometer, gyroscope, magnetometer, *Physical health sensors* includes heart rate sensors(HR), skin temperature, oxygen saturation (SPO2) etc., *Environmental sensors* include, temperature, barometer, humidity, light sensors etc. Sometimes these wearable sensors are not used in applications, for example in human activity recognition due to size, price and acceptability to carry by the user [6].

**(b) Smartphone Sensors:**

Since smartphones, as well as smartwatches, are easy to carry and use, they are used widely to collect data. Today smartphones and smartwatches are embedded with sensors like accelerometer, gyroscope, barometer, GPS, temperature etc. Smartphones and watches are now also used in other fields alongside human activity recognition, like, health monitoring, monitoring sports activities etc. [3,6].

**1.3 Challenges in sensor based human activity recognition:**

In conducting human activity recognition, over the years, variety of challenges have been faced. Challenges varies from type of sensors to use datasets to be used to methods or models and many more.

First of all, the problem arises and much challenge is faced in choosing a suitable sensor or set or sensors as well as the type of sensors that will be used to collect raw data.

Sensor based human activity recognition eliminates the drawback of vision based system, where data has to be collected by means of external sensing method like camera. When collecting data by means of camera or CCTV the user or the subject has to stay within a particular range, which limits the reliability. Furthermore, data taken by camera are influenced by the background, daylight, weather etc. And thus, sensor based systems are preferable which eliminates all of these drawbacks [14,15].

Now the challenge arises in choosing the suitable type of sensors. Sensors can be a separate wearable sensor or sensors used in smartphones. Each have their own trade off. A wearable sensor is used for only a particular purpose whereas smartphones have multiple sensors embedded in it. The first challenge in wearable sensors is to select which sensor to use that is whether to use only one sensor or multiple sensors at a time. If only one sensor is used the data processing part becomes trouble free. But the recognition accuracy and thus the performance might not be satisfactory. On the other hand, with the use of multiple sensors, the accuracy and performance of the system will be more acceptable and better. Moreover, single sensor based system might be good for recognizing simple activities like walking, sitting, stand to sit etc., though not ensured to good accuracy, but will face much challenge in detecting complex activities and will not give high accuracy. Hence, multi-sensor based system is preferable over single sensor. Multiple sensors like accelerometer, gyroscope, magnetometer etc. are used to collect data and recognize complex activities [13,16].

Though multi-sensor based wearable sensors have shown high accuracy and good performance, challenges are also faced in this method. First if to be mentioned, users may feel uncomfortable in wearing or carrying these sensors while data is being collected. And it is problematic to carry multiple sensors and data becomes difficult to process. Besides, in unfavorable environment hindrances are faced while calibrating and collecting data [3,17]. To overcome these challenges smartphones and smartwatches have been preferred in recent years. Specially, with the advancement of technology, powerful smartphones and smartwatches have been introduced, which have high computation capacities along with powerful sensors. Smartphones are used in our daily lives and it is not only a communicating devices but also a good sensor platform and sensors are calibrated easily as well as data can be transferred easily by means of Bluetooth, Wi-Fi etc. [12,18].

However, challenges are faced in using smartphones, more precisely in collecting sensor data. Moreover, one another problem with smartphone is it rapid loss of battery energy. The main challenge that is assumed is the variation of location of smartphones. In quite a number of papers like [8,18-21] this problem has been discussed and as well as solutions have been proposed. Unlike wearable sensors, smartphones aren’t kept in a fixed position and it is not reasonable to do so either. Smartphones are often kept in shirt or pant pockets, bags and in different pockets in bag etc. places and also these are kept in either right hand or left hand along with various orientation. Signals read will vary in these conditions which makes it arduous to recognize the particular activity. If activity recognition is done using smartphone data while keeping it in only a particular position, orientation or place, low accuracy and wrong recognition will result. And, it is the greatest challenge to make the recognition system with smartphone a position/location independent system. And also data collected from smartphone may contain much noise. Hence, researches have been conducted and is still being conducted focusing on such problems with location of smartphones [8,12,18].

Apart from the above mentioned challenges, one of the biggest challenges is the method to choose for recognition i.e. whether to use a hand-crafted featured model or a deep learning model. Whether hand-crafted featured model or deep learning model is used, selecting and finding appropriate features for the recognition system is the most significant task in human activity recognition. Feature extraction is more challenging as different activities can have similar characteristics and hence it becomes strenuous to obtain unique features for each activities [7,11]. Hand-crafted featured models since requires features to be selected manually, knowledge equivalent to an expert is required. Moreover, hand-crafted features might not work for other similar applications and also will face much difficulties while recognizing complex activities [22,23]. Furthermore, hand-crafted featured human activity recognition system large and most importantly labelled data from sensors are vital for good performance. Otherwise, expected accuracy and performance will not be obtained. Additionally, in order to collect such data and hence prepare proper datasets sophisticated infrastructure might be required leading to more cost and time consumption. Also, many datasets have labelled data indeed but those labels might be assigned to data having information from secondary activities. And, training errors might occur in machine learning algorithms due to such data and degrade the performance, specially while a class or activity is being assigned to a data segment [9,24].

In order to tackle the challenges and drawbacks of hand-crafted features, deep learning models are used in recent years. In deep learning models, necessary features are automatically extracted from raw sensor data. Unlike popular machine learning algorithms such as, KNN (K nearest neighbor), SVM (Support Vector Machine) etc. deep learning models don’t need carefully engineered hand crafted feature as it can develop most efficient features from raw data [10]. Deep learning models contain, multiple layers and the path deepens and hence the name deep learning [2]. Deep learning methods are being used in human activity recognition(HAR) to automatically extract useful features from raw sensor data using multiple layers of abstraction. And it is applied in HAR concerning the movements of human being hierarchical [24]. Numerous deep learning models using deep learning networks like CNN (Convolutional Neural Network), LSTM (Long Short Term Memory), Bi-LSTM (Bidirectional Long Short Term Memory), RNN (Recurrent Neural Network) etc. have been introduced in large number of papers. For example, in CHIH-TA YEN *et al.* a deep learning model using CNN has been proposed [2], in Xiaokang Zhou *et al.* LSTM based model [13] and in [24] RNN based model has been proposed to overcome previous challenges and obtain better performances.

However, implementation of deep learning models doesn’t mean that the system will not have lack in performances and efficiency. Challenges are faced even when using deep learning models. As we know deep learning models have numerous layers, these require large number of initializations and parameter tuning. All of these escalates computational costs, time and requires powerful processing units for fast computations. Moreover, due to such large computations, low processing devices and low energy mobile devices aren’t suitable [10]. Hence, pre-processing and dimensionality reduction are significant in the recognition system. Dimensionality reduction, can lessen the computational complexity mentioned above. Now, one of the greatest challenges is the pre-processing of the data to achieve optimum performance. For this, various processing techniques such as normalization, standardization etc. have been used and need further experimentation to find out suitable accuracies and performances which is the challenge. Moreover, challenges remain in issues with hyper-parameters like learning rate optimization, kernel filter size, reduction of data size etc. [26]. In order to improve the performance data augmentation method is used to generate more training examples from existing small dataset and in order to reduce and sometimes prevent overfitting. For data augmentation techniques like permutation of location with sensor events, arbitrary rotation etc. are used. This method is still a challenge for implementing motion sensors’ (accelerometer, gyroscope) data for improving performance [27].

Apart from all the challenges that have been mentioned above, challenge, that cannot be denied, remains in using proper datasets. Custom datasets can be prepared collecting data using wearable sensors, smartwatches or smartphones. But preparing custom datasets is challenging because, the raw data need to be pre-processed properly for the deep learning layers to extract useful features and perform recognition. Since sometimes much challenges are faced in preparing custom datasets, publicly available datasets can be found and used for human activity recognition. In KAIXUAN CHEN *et al.* information and also challenges about publicly available datasets have been given. For example, the most common three datasets used are UCI HAR, WISDM Activity Prediction, OPPORTUNITY and their challenges that for the first and third one have been mentioned to multimodality of the data whereas for the second dataset the challenge is said to class imbalance [7].