Human Activity Recognition And Classification Using Sensory Data With MLP

Archana Sarathe[1], Prof. Anshul Khurana[2]
[1] Dept of CTA
[2]Professor, Dept of CSE
[1, 2] Shri Ram Institute of Technology Jabalpur, Madhya Pradesh, India.

Abstract- *HAR detects and identifies human activity in the real environment by learning useful information from original sensor data or videos containing human activities. Sensor based HAR uses sensors to obtain activity information and classify activities without acquiring video images, which not only protects the privacy of users, but also reduces the computational complexity and expands the use scenarios. With the development of pervasive computing and the improvement of sensor technology, the characteristics of small size, low cost, high capacity and portability have attracted people's attention. More and more mobile terminals such as mobile phone and sports bracelets have embedded such as triaxial accelerometer, gyroscope and magnetometer sensor, which greatly promoted the development of HAR based on sensor.*

The purpose of this research was to conduct an investigation of deep learning and dimensionality reduction techniques in human activity recognition and behavioural prediction using MHEALTH data. These techniques combined with multiple sensor data aim to classify daily activities. Previous work in HAR has focused on using multiple accelerometers placed on different parts of the body, with more recent work focused on sensors embedded in smartphones to classify activities. This research classifies activities from utilising the data from the following sensors – accelerometer, gyroscope and magnetometer.

Keywords- Human Activity, Sensor data, Machine Learning, Deep Learning, Multilayer Perceptron, Accelerometer and Gyroscope.

I. INTRODUCTION

Continuous growth in the health sector has led to astronomical advancements in the field of medicine. Due to this continuous growth, the quality of life has greatly increased when compared to one hundred years ago. Everything from life expectancy, physical health, education, safety and freedom has vastly improved. This rise in health sector growth has led to an increase in healthcare costs. Increased healthcare costs in monitoring patients with chronic illness, monitoring the elderly along with many more instances has led to drastic cost-cutting measures employed by healthcare providers worldwide.

Technological advancements in healthcare can contribute unquestionably in reducing healthcare costs by ensuring clinicians, doctors and other medical staffs operates and conduct their daily activities more efficiently in the hospital vicinity. Since the turn of the 21st century, Human Activity Recognition (HAR) [1] has undergone significant research in the healthcare domain. Human activity recognition utilised with powerful technologies can potentially benefit remote patient monitoring, the elderly, patients suffering from chronic illness and ambient assisted living.

Simple activities such as cycling, running and jogging have been successfully recognized and classified to date. Complex activities are proving increasingly difficult to monitor, with continuous active research conducted in this area of HAR. The main goal of HAR is to predict common activities in real-life surroundings. Researchers are exploring pattern recognition and human-computer relationships [2] due to its applicability in the real world, such as a Human Activity Recognition Healthcare Framework. Successfully classifying human activities through wearable sensors generates endless individual information, which provides insight about the individuals' functional ability, lifestyle and health [3].

1.1 Basic Building Blocks HAR systems: The figure below details the basic building blocks of any human activity recognition (HAR) system.

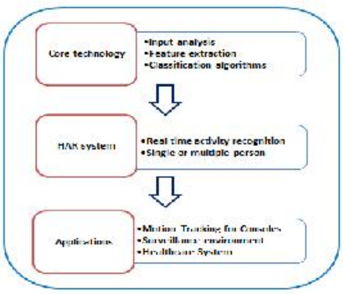

Figure 1.1: Basic building blocks of almost every system [4].

Human Activity Recognition (HAR) becomes a very popular and active research area for researchers from the last two decades. However, it still remains a complex task due to some irresolvable issues such as sensor movement, sensor placement, background clustering, and the inherent variability of how different people perform activities. Determining detailed activities is beneficial in many areas of human-centric applications, such as home care support, postoperative trauma rehabilitation, abnormal activities, gesture detection, exercise, and fitness. Most of the person's daily tasks can be simplified or automated if recognized by the HAR system. Usually, HAR systems are based on either unsupervised or supervised learning. A supervised system requires pre-training using special datasets, but unsupervised systems have a set of rules during development [5].

HAR system can be divided into several modules, sensing, segmentation, feature extraction, classification and post preprocessing. Generally, according to the sensing method, we can classify the HAR framework into two sorts: vision-based and speeding up based strategies. Vision-based technique, as a rule, utilizes at least one cameras to gather information, this goes under outside detecting which perceives complex exercises like shower, drinking mug of espresso, washing dishes and so forth however acknowledgment of human exercises through video or picture is testing task because of which it is less well known when contrasted with different sensors and if video is viewed as it is generally for security reason. Then again increasing the speed-based strategy approach clients to wear a few accelerometers for information gathering.

The advantage of the vision-based system is that it works without placing any sensors with users, but its recognition performance highly depends on the light condition, visual angle and other outer factors. On the contrary, the acceleration-based system requires users to wear a device, but almost eliminates all those outer interferences. With this type of system, we use accelerometer which is the most common device for activity tracking, counting the number of steps and also assess their quality [6, 7]. HAR is reflected as a significant component in several fields like in Surveillance System, Human-computer interaction, anti-terrorists, anti-crime securities, Healthcare as well as life logging and assistance etc.

1.3 Various Approaches to HAR

To achieve the goal of recognizing human activity, a HAR system is required. The two most commonly used techniques for this purpose are sensor-based and vision-based activity recognition. We can classify them, as shown in Fig. 1.2.

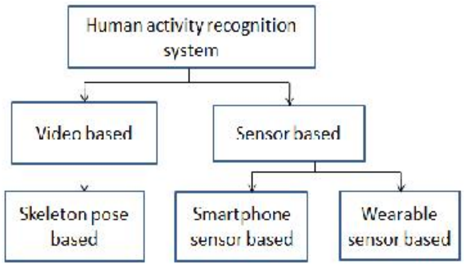

Figure 1.2: Classification of HAR system based on their approaches [4].

A. Pose Based Approach

Poses are important for analyzing videos, which include humans, and there is strong evidence that body posture concepts are very effective for various tasks such as activity recognition, content extraction, etc. This approach classifies human actions based on the coordinate information of the body parts. Basically, HPE refers to the process of assessing the composition of a part of the human body (3D poses) or the projection onto an image plane (2D HPE). It covers all issues related to the human body, from understanding the entire human body to the detailed localization of body parts [8]. It is formulated as a regression problem that can be modelled with a simple CNN. It takes the entire image as input and shows the\ pixel coordinates of the body's key points. There are 15 body joints: Neck, LKnee, LAnkle, RShoulder, RWrist, Relbow, LShoulder, RHip, LElbow, LWrist, Chest, and 14 joint connections. The classification problem can be formulated as a multi-class classification problem that can also be modeled by using neural networks. CNN accepts the body joints location as input and generates a number vector representing the probability of each activity labels accordingly. Some popular dataset is also available such as MPII, which contain more than 20,000 labelled images of 410 specific subcategories activities under 20 activity categories [9].

B. Smartphone Sensor- Based

Smartphones are the most useful tool in our everyday lives, and advanced technology is enabling us to meet the needs and expectations of customers every day. To make these devices more functional and powerful, designers are adding new modules and devices to the hardware. Sensors enhance the capabilities of smartphones and play a major role in understanding the environment. As a result, most smartphones have a variety of built-in sensors that can collect a wealth of useful data about the human’s daily life. Sensors retrieve information from body gestures and then recognized the activities. The most commonly used sensors are accelerometer and gyroscope etc. Accelerometer sensor is used for measures the change in speed, and gyroscope is used to measures the orientation of the body. Some techniques of HAR through smartphones used in recent studies are SVM, k-NN, Bagging, Ada Boost [10].

C. Wearable Sensor Based

The wearable technique uses sensing devices to be mounted on the subject to collect data from the sensors. As human activity contains actions of different bodily positions, the research of human activity needs to capture information from more than one sensor installed on the different parts of the body of the person. Wearable devices must be designed with user accessibility in mind. Lightweight, modern, and comfortable wearing devices with embedded sensors are used for activity monitoring. Activity monitoring sensors are used in multiple datasets. The most commonly used sensors are an accelerometer, gyroscope, magnetometer, and RFID tag [11]. After feature extraction and modelling, human activities can be recognized through statistics, and machine learning algorithm is applied. How to map low-level sensor data to higher-level abstractions is the key to activity recognition.

II. RELATED WORK

Modern environments can benefit greatly from the continuous monitoring of workers current activities. The data generated from sensors in these activity recognition applications allow information to be analysed and pro-actively used. Modern applications, such as in the aircraft industry, yield these benefits. Aircraft maintenance procedures with the aid of activity recognition applications using wearable sensors continue to develop. Nicolai, Sindt, Witt, Reimerdes, and Kenn [12] employ a wearable computing system with the aim of reducing aircraft maintenance. The system incorporates management techniques such as visual assisted maintenance techniques, deployment of electronic logbook as well as aircraft industry manuals.

Lampe, Strassner and Fleisch [13] present a ubiquitous computing environment for Aircraft Maintenance to ensure maintenance time is minimised and resources such as staff, tools and techniques are efficiently used.

Maurtua, Kirisci, Stiefmeier, Sbodio, Witt [14] present an activity recognition prototype using wearable electronics that aids training activities in automotive production. Results showed that the prototype was highly effective in allowing automotive production to be flexible. Workers guided by the prototype, ensure automotive production is maximised compared to operating alone. Monitoring the workers procedures and actions led the prototype guided the worker in performing appropriate tasks for error handling.

Aleksy and Rissanen [15] reports on how wearable electronics can yield benefits in modern working environments. A detailed account of how efficiency improves in various sectors with constant interaction between the user and the device is given. The study presents a case study where processes in an industrial plant aided by wearable technology to analyse if productivity increases.

Cheok et al. [16] presents a wearable electronic entertainment system, which is a simulator evolving around a human Pacman entertainment system. The system is a mobile, wide-area entertainment system based on physical, social, and ubiquitous computing. Cheok et al. [16] builds an architecture that is capable of monitoring human motion with the data generated from body worn inertial sensors. Building an entertainment system that clearly mimics the real and virtual world is difficult, [45] successfully accomplishes this.

Tobita and Kuzi [46] present an unusual entertainment application. The wig-based wearable computing device enhances communication and provides entertainment to users. This obscure, natural looking device employs wearable sensors to allow to parties to communicate with each other. SmartWig is one of the first wearable electronic applications for effective communication in the entertainment sector.

2.9 Activities

When presented with a human activity recognition problem, choosing the correct sensor to complement the working environment or application is important. The amount of different types of sensors to choose from is endless, with advanced sensors developed each year. Continuous research in the activity recognition field has led more companies trying to reap benefits of predicting activities to improve communication and improve productivity. This section discusses the different types of activities performed for activity recognition.

Zhu, Xu, Guo, Liu, and Wu [48] classify physical exercise activities such as running, jogging, standing still and powerwalking using body-worn inertial sensors in related work. These activities, when performed, generate a specific type of range in body motion with the acceleration measured being relatively similar when individuals of different characteristics perform the activity.

III. PROPOSED WORK

For the Figure 4.1 below shows the data flow and approach implemented to classify the activities detailed in chapter 4. All three sensors data (gyroscope, accelerometer, magnetometer and electrocardiogram) are classified together while predictions are generated from each individual classification model. A classification model including all six classification predictions is generated, followed by an in-depth evaluation procedure. The best overall prediction model is then identified.

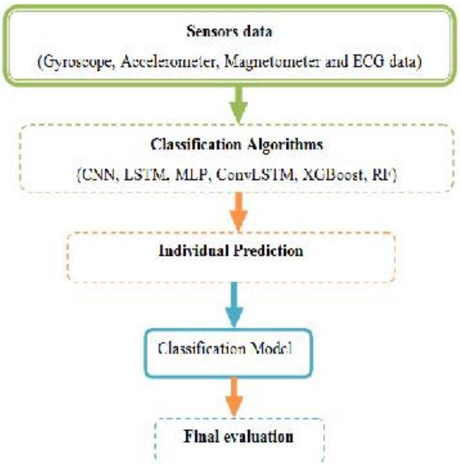

Figure 4.1: Proposed Model Architecture.

4.1 Proposed Model steps:

The steps of the proposed model are described below:

Stage 1: Business Understanding

The first goal of this research is to build six different types of prediction models and conduct a comparative study on the performance of each model to identify which model yields the greatest results in terms of classifying human activities. The six models are CNN, LSTM, ConvLSTM, MLP, XGBoost and AE with RF.

Stage 2: Data Understanding

For this research, The MHEALTH dataset is analysed. The dataset includes body motion and vital signs recordings for ten differing characteristics while performing several physical activities. Twelve different activities with three sensors attached to their body (gyroscope, accelerometer and magnetometer) are recorded. Each activity leads the three sensors to record twenty-three signals. There are no missing values in this classification task. The goal is to predict each activity given the twenty-three signals recorded. The data collected for each volunteer is stored in individual log files: 'mHealth_subject.log'. There are ten log files altogether. Each file contains the data samples recorded for all three sensors.

Stage 3: Data Preparation:

Data preparation involves feature extraction, encoding labels to one-hot form, converting the raw data into the right shape for input into the model, normalising the data and finally splitting the data into training and testing. Each volunteers' data was analysed separately for various reasons.

Stage 4: Modelling

Building each model involved extracting the features and labels, hyperparameters setting, compiling the model and model evaluation. Following models are build.

- CNN: The process for building the CNN model.
- LSTM: The process for building the LSTM model.
- ConvLSTM: The process for building the ConvLSTM model.
- MLP: The process for building the MLP model.
- XGBoost: The process for building the XGBoost model.
- RF: The process for building the RF model.

Stage 5: Evaluation

This stage involves evaluating each model. Each classification model is evaluated using accuracy, precision, recall, F1-Score, feature importance and confusion matrix. All six-classification models are evaluated based on their categorical cross entropy values. The model, which best suits the data is chosen based on the evaluation procedures.

DBSCAN and t-SNE are evaluated based on how well each one visualises clusters in this high-dimensional dataset.

Stage 6: Deployment

The final stage involves the deployment of each model into Python to analyse the MHEALTH dataset. Each models code is not only applicable to the MHEALTH dataset, but can be applied to any health prediction dataset, as each model is built on multivariate analysis. Finally each models experimental result is analysed.

4.2 Feature Extraction

The next step is feature extraction. The MHEALTH dataset consist of 10 log files, with each log file corresponding to each of the ten subjects. In order to extract the features (signal attributes) and labels (activities) of each log file, the following feature extraction method is used.

1. Each file is opened by reading it into the Python Editor.
2. Each line in the file is then processed and read in until the final line in the file is reached.
3. Each line is split to extract the labels for each line.
4. Once the label is found, it is split and removes the word. If there is a space or comma, it detects it.
6. A sub list is created store all the values of each line.
7. Each line is split into values.
8. An array is then created from the sublist.

This loop successfully extracts all the features and labels of each subjects log file.

IV. RESULTS WORK

Hyperparameters setting

This section gives an overview of training the model and hyperparameters setting.

- In the training process, the 'fit ()' function is used to train the proposed model.
- 'X_train' represents the training data.
- 'y_train' refers to the target data.
- 'X_test,y_test' represent the validation data.
- The model is trained on 706317 parameters.

While training the model:

- The Learning rate is set as 0.0005.
- Batch size is set as 32.
- Training process is run for 20 epochs.

The generalisation between each subjects activity performed was measured in performance. Performance is measured through precision, recall, f1 score and accuracy. The following table depicts a detailed evaluation summarisation of each class, the figures are normalised to the percentage of data. Table below shows the comparison of the proposed and existing models.

Table 5.1: Comparison Table.

Algorithms	**Evaluation parameters**			
	Accuracy	**Precision**	**Recall**	**F-Score**
CNN	84.16%	83.91%	84.16%	83.28%
XGBoost	90.024%	90.155%	90.024%	89.85%
CNN-LSTM	84.61%	84.32%	84.61%	84.14%
Random Forest	83.3%	83.5%	83.9%	81.5%
Proposed Model	90.53%	91.71%	90.53%	90.76%

REFERENCES

[1] Valliani, A., & Ranti, D., & Oermann, E. (2019) Deep learning and neurology: A systematic review, Neurology and Therapy, Springer, vol. 8, issue. 2, pp. 351–365.

[2] A. Charidimou, A. Krishnan, D. J. Werring, and H. R. Jager, "Cerebral microbleeds: a guide to detection and clinical relevance in different disease settings," Neuroradiology, vol. 55, no. 6, pp. 655–674, 2013.

[3] A. Charidimou and D. J. Werring, "Cerebral microbleeds: detection, mechanisms and clinical challenges," Future Neurology, vol. 6, no. 5, pp. 587–611, 2011.

[4] M. Vernooij, A. van der Lugt, M. A. Ikram, P. Wielopolski, W. Niessen, A. Hofman, G. Krestin, and M. Breteler, "Prevalence and risk factors of cerebral microbleeds the rotterdam scan study," Neurology, vol. 70, no. 14, pp. 1208–1214, 2008.

[5] C. Cordonnier, R. A.-S. Salman, and J. Wardlaw, "Spontaneous brain microbleeds: systematic review, subgroup analyses and standards for study design and reporting," Brain, vol. 130, no. 8, pp. 1988–2003, 2007.

[6] A. Charidimou and D. J. Werring, "Cerebral microbleeds and cognition in cerebrovascular disease: an update," Journal of the neurological sciences, vol. 322, no. 1, pp. 50–55, 2012.

[7] Z. Wang, Y. O. Soo, and V. C. Mok, "Cerebral microbleeds are antithrombotic therapy safe to administer? Stroke, vol. 45, no. 9, pp. 2811–2817, 2014.

[8] M. Akter, T. Hirai, Y. Hiai, M. Kitajima, M. Komi, R. Murakami, H. Fukuoka, A. Sasao, R. Toya, E. M. Haacke et al., "Detection of hemorrhagic hypointense foci in the brain on susceptibility-weighted imaging: clinical and phantom studies," Academic radiology, vol. 14, no. 9, pp. 1011–1019, 2007.

[9] J. D. Goos, W. M. van der Flier, D. L. Knol, P. J. Pouwels, P. Scheltens, F. Barkhof, and M. P. Wattjes, "Clinical relevance of improved microbleed detection by susceptibility-weighted magnetic resonance imaging," Stroke, vol. 42, no. 7, pp. 1894–1900, 2011.

[10] S. M. Greenberg, M. W. Vernooij, C. Cordonnier, A. Vishwanathan, R. Al-Shahi Salman, S. Warach, L. J. Launer, M. A. Van Buchem, and M. Breteler, "Cerebral

microbleeds: a guide to detection and interpretation," The Lancet Neurology, vol. 8, no. 2, pp. 165–174, 2009.

[11] S. Gregoire, U. Chaudhary, M. Brown, T. Yousry, C. Kallis, H. Jager, and D. Werring, "The microbleed anatomical rating scale (MARS) reliability of a tool to map brain microbleeds," Neurology, vol. 73, no. 21, pp. 1759–1766, 2009.

[12] A. Gebrehiwot, L. Hashemi-Beni, G. Thompson, P. Kordjamshidi, and T. E. Langan, "Deep Convolutional Neural Network for Flood Extent Mapping Using Unmanned Aerial Vehicles Data," Sensors (Basel)., vol. 19, no. 7, p. 1486, 2019.

[13] K. O'Shea and R. Nash, "An Introduction to Convolutional Neural Networks," no. November, 2015.

[14] N. Passalis and A. Tefas, "Learning Bag-of-Features Pooling for Deep Convolutional Neural Networks," Proc. IEEE Int. Conf. Comput. Vis., vol. 2017-Octob, no. September, pp.5766–5774, 2017.

[15] Hyunkwang Lee. "Practical Window Setting Optimization for Medical Image Deep Learning." URL: https://arxiv.org/abs/1812.00572.

[16] N. V. Chawla et al. "SMOTE: Synthetic Minority Over-sampling Technique". En. In: Journal of Artificial Intelligence Research 16 (June 2002), pp. 321– 357. ISSN: 1076-9757. Doi: 10.1613/jair.953. URL:https://www.jair.org/index.php/jair/article/view/10302 (visited on 12/16/2019).