Intelligent Behavior Recognition based on Gyroscope and Accelerometer

YUNG-HAN LU, CHUN-TING LIN, PIN-YU HUANG,

BO-XIAN HE and Tun-Wen Pai *

Department of Computer Science and Information Engineering,
National Taipei University of Technology, Taipei, Taiwan

*twp@ntut.edu.tw

Abstract.
In such an aging society, the rate of dementia has dramatically increased. It is urgent to provide intelligent facilitative devices and appropriate medical care for those who with dementia symptoms, especially for the elders. Accordingly, the primary research target of this study is to automatically detect and monitor behavior of the elderly suffered from dementia. Although previous research has been conducted to identify various human movement patterns, the experimental data all came from cameras. Furthermore, these research applications are not specifically designed to identify the behavior of the elderly with dementia. In order to improve challenges of previous research methods, our study adopted wearable multi-axis sensors to collect corresponding motion data with protected privacy. In addition to collecting raw data containing time series signals, frequency and spatial domain features were applied to detect human behaviors, and Random Forest algorithms were used to construct a motion prediction model. The classification system for real-time motion recognition contained two main actions of users including holding/putting hand gestures and walking model. Finally, an additional camera on the chest was designed to be triggered and take a shot according to previous settings. It could enhance practical applicability of the proposed intelligent system to assist the elderly with dementia to find her/his lost objects. In order to extend the subsequent applications, this study developed an APP to record individual's behavior and to integrate with a chat robot to answer the location and time of the lost objects, or combine it with a medical advice of the individual's rehabilitation training to record all the activities during the rehabilitation processes, which was used as a reference basis for the patient's returning to hospital visits and providing medical staffs with appropriate recommendations for supporting precise medical care and treatments.

Keywords: Internet of behavior, motion detection, gyroscope, accelerometer, dementia

1 Introduction

According to studies by Stopschinski et al [1] and Fu et al [2], the prevalence of dementia in society increases rapidly with the rapid aging of the population. According to a study by Chen and others [3], the prevalence of dementia doubled for each 5-year increase in the average age of the Taiwanese population over the age of 65.

In response to the rapid increase in the elderly and dementia population, the Ministry of Health and Welfare plans to integrate relevant resources to develop a diversified network of community-based care services for the elderly with dementia. To develop a long-term care service network plan for them. Moreover, the training of quality professional caregivers for dementia is ongoing. Both of the knowledge and skills related to dementia prevention and treatment are included in the training courses for professional caregivers in long-term care. Unfortunately, the current dementia medication cannot stop or restore the damaged brain cells. Therefore, technological assistance has become the focus development in long-term care.

With the rapid development of technology, the use of emerging smart technologies such as Internet of Things (IoT), Virtual Reality (VR), Bluetooth positioning and wearable devices. These have made smart technology-assisted medical care feasible, and these technological aids are optimized with the patient in mind. Such as companion robots which are more intelligent, touching the hearts of the elderly during the chatting process, realizing the wisdom of nostalgic treatment, and achieving the possibility of delaying dementia. If people with dementia are willing to let go of their feelings for technology, they will be able to enjoy the experience. If people with dementia are willing to let go of their wariness of technology and engage in sports and social life, their bodies and minds will begin to revitalize and achieve the therapeutic purpose of sports and cognitive training. Technology assistive devices not only encourage cognitive training, but also care for mental health.

However, an important phenomenon of elderly people with dementia is short-term memory loss, often forgetting objects or their own location, causing great disturbance to the patients and even their caregivers in their daily lives. This study will focus on how to automatically analyze and identify the behavioral actions of elderly with dementia and effectively record important information about their daily life processes, with the hope that the function of big data analysis can effectively improve the quality of daily life of elderly with dementia and their caregivers.

The traditional data collection methods mostly use video cameras to record the whole process of a specific person or all the surrounding people at the same time. This data collection method often involves privacy and portrait rights issues, the device is not easy to carry, the types of activities recognized are limited, and the data files and computing resources are too large [4-6].

This study focus on analyzing behavioral actions, analyzing the wearable sensor data in stead of photographic data, which can effectively avoid privacy and portrait rights violations and solve the problem of unportable devices. However, the first important issue in using a wearable sensor for motion recognition research is where the sensor should be placed on the human body. The second challenge is to identify the best combination of features needed to distinguish the motion. The second challenge is to identify the best combination of features needed to distinguish movements. Generally, developers have to continuously analyze and try to find the most important features to distinguish different behaviors and movements. For example, Yu-Jin Yang and others [7] proposed a multiple simple sensor system for real-time action recognition, where the sensors are placed on the right wrist, right arm, chest, left waist, right thigh, and right ankle. Then this study collects nine different actions such as standing, sitting, lying, walking, running, going up the stairs, going down the stairs, lifting a dumbbell, and drinking water. In this paper, when the number of sensors was reduced to two (right wrist and right ankle). The results showed that even if there were only two sensors, the test results were only slightly affected. After comparing the accuracy rate, reducing the number of sensing devices to two can still maintain good recognition results. Therefore, this study extends this result to the activity recognition of elderly people with dementia. In addition, Shu-Yu Chang et al [8] used accelerometer and gyroscopes for motion interpretation of walking on stairs, slopes, and flat surfaces. The study learned that the human walking cycle by the foot received the most obvious message, so this study will be based on the cycle judgment on the foot sensor data, so as to determine the personalized window size (Window Size) selection basis.

The current study by Jukka-Pekka Onnela et al [9] used gyroscope and accelerometer data from smartphones for activity classification, and they introduced a modified version of the so-called movelet method to classify activity types and quantify the uncertainty in the classification, distinguishing between walking, going up stairs, going down stairs, standing, and sitting, but the predicted results were highly correlated with the placement of the phone. We will design a system to improve this problem and increase the accuracy of the prediction.

To overcome the challenges of previous studies, this study will use a wearable inertial measurement unit (IMU) to collect privacy-protected data for data analysis and to determine various behavioral patterns of users. The effect will be easy to carry. In addition to the recognition of daily living behaviors, this study also takes into consideration and enhances the recognition of the movements of the elderly with dementia in picking up and putting down objects, which can enhance the applicability and practicality of the intelligent system to help the elderly with dementia improve their quality of life.

2 Research Methodology

2.1 Subjects and source of dataset

Data collection will be conducted from 2021 to 2022. This dataset is conducted by this study. Eligibility requirements include being 18 years of age or older and being able to walk and pick up/place objects without assistance or other equipment. Four healthy subjects, including one female and three males, were enrolled in the study. The data sets were divided into hands and feet, and each collection was made in 59-second increments, with a total of 274 sets of data, each with about 800 to 1000 data.

2.2 Equipment

Fig. 1. (a), the wearing position of Hand

Fig. 1. (b), foot sensor wearing position

All participants were equipped with two sets of IMU (Inertial Measurement Unit) containing a variety of complex information, including three-axis gyroscope, three-directional accelerometer (Accelerometer) and three-directional magnetic field meter (Magnetic), the above three combined into ESP32, installed in the subject's right wrist and left foot position as shown in Figure 1 ((a)(b) to measure the data. Measurements were performed at a sampling rate of 20 Hz, and the received data were transmitted by Bluetooth HC-05.

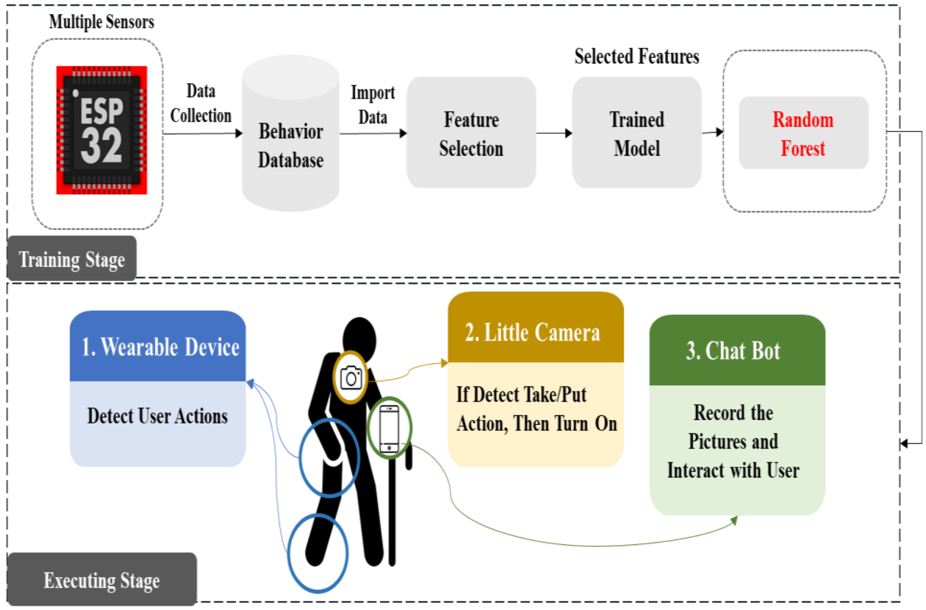

Fig. 2. System Architecture

2.3 Deep Learning Regression Model

The goal of this study is to complete the analysis of key behaviors of people with Alzheimer's disease and to develop a system for searching for forgotten items. Figure 2 shows the system architecture of this study. The subjects of this system will wear two multi-axis sensors, and each sensor will collect 14 returns. (Three gyroscope values, three magnetic field values, three accelerometer values, four rotation offset values, and one time value.) Data will be added to Fourier transform, first order differential, second order differential, and arc-tangent analysis. In addition, the data will be analyzed and compared with previously published literature, and all data will be trained by Random Forest to classify user's individual actions. After the training and verification of the prediction model, the user's daily actions, the location of the objects and the images of the objects will be recorded through the smart phone app and the chat robot.

In the implementation phase, the user wears the sensor device on the wrist and ankle, and uses the ESP32 return value to automatically classify different behavioral actions through the model, and then activates different functions after recognizing specific actions. When users are looking for forgotten items, they can view the photos through the cell phone app and learn the actual location of the items, and later they can even find the items.

3 Research Steps

The system development process is shown in Figure 3, which includes three main modules: motion recognition, camera triggering module, and APP connection application. In the motion recognition module, this study aims to use two ways to receive the data of action behaviors, including collecting data by writing programs in Python and receiving data from two sensors simultaneously through the Bluetooth function of the sensor device. The action classification will be done by constructing Random Forest models to predict behavior patterns using Fourier, first order differential, second order differential, and arc-tangent features. The second module receives the classification results of the previous module, triggers the camera's on/off record based on the predicted results, and finally accesses the photos to the database using Amazon S3 technology. The third app-connected application receives the photos stored in the database of the second module, presents the results and subsequent applications through the mobile UI.

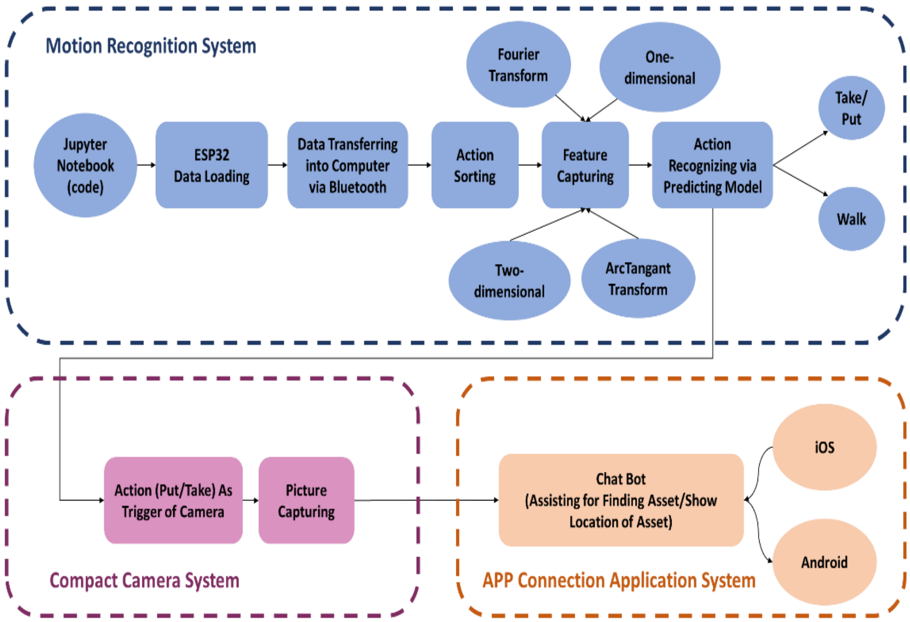

Fig.3. Flow chart of develpment

3.1 Motion recognition system

Feature selection The nine-axis sensor collects the data of the tester's movement, which is usually continuous and contains many movements such as holding, placing, walking, forming a long data consisting of many discrete points. Segmentation is to decompose the time domain signal of each action, and the segmented signal must contain a complete action. If it is too small, the action will be sampled incompletely. If it is too large, more than one action will be sampled, which will affect the accuracy rate in the later stage. Therefore, this study sets the Window Size to 30 at a sampling frequency of 20Hz for each action not exceeding 1.5 seconds, which is most reasonable. Fourier transform is used to convert the motion data from time domain to frequency domain. It is easy to identify the repetitive motion patterns and confirm the relative frequency information by using the frequency domain signal. Equation (1) is the Fourier transformation formula.

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-i2\pi kn/N} = \sum_{n=0}^{N-1} x(n) \cos \cos \left(\frac{2\pi kn}{N}\right) - i \sum_{n=0}^{N-1} x(n) \sin \sin \left(\frac{2\pi kn}{N}\right) \quad (1)$$

Where, $An = x(n)cos\left(\frac{2\pi kn}{N}\right), Bn = x(n)sin\left(\frac{2\pi kn}{N}\right)$ (2)

First order differential, second order differential calculate the extremes of the occurrence point and for the maximum or minimum worth appearing, by these two features can also be known according to the amount of time change and acceleration and this data can be more visually see the change of action (walking, holding and placing) first order differential from the rear minus forward divided by the time to do the approximation, s to do the differential approximation of the primary differential. Equation (3) is formula.

$$f(n) = \frac{(x_{n+1} - x_n)}{(t_{n+1} - t_n)} \qquad f(N) = \frac{(f(N+1) - f(N))}{(t(N+1) - t(N))} \tag{3}$$

Training model approach (Random Forest) The reason of choosing Random Forest model is that can handle high dimensional data (i.e., multiple feature data) and does not require preliminary feature selection. Moreover, the effect between features can be detected during the training process. Random Forest can balance the error when dealing with unbalanced data sets. It also provides an effective way to balance the error of a dataset when there is a classification imbalance. The training process can also give the importance of features and thus select the 3 features needed in the study (removing features with low importance). In this study, 274 data sets were used as training samples, and 36 features were selected for training the model. The five main parameters that are deployed in this model are: the maximum number of features that can be used in a single decision tree as a percentage of the total number of features (max_features is finally set to 0.2), the number of subtrees, usually the larger the value, the better the result, but after a certain range, it becomes volatile and unstable (n_estimators is finally set to 100), the minimum leaf If the actual number is smaller than this sample size, it will choose to merge with other sibling leaves (min_samples_leaf is finally set to 50), the random forest cross-validation method (oob_score is finally set to True) and the evaluation criterion of features when dividing CART trees (criterion finally uses entropy)**.**

APP Technology (Amazon S3) Amazon Simple Storage Service (Amazon S3) is an object storage service that provides a simple web service interface for object storage, allowing researchers to store and retrieve any amount of data anywhere on the web. This technology has a RESTful API that allows easy connection to programs such as Java, PHP, etc. AWS provides corresponding SDKs that allow researchers to easily access files on the cloud.

This study uses this technology as a base to develop an Android and IOS app that allows users to view photos stored in Amazon S3 through the medium of their cell phones to facilitate the subsequent retrieval of items and applications. The development steps are as follows.

Step one. Go to Cognito to get the PoolId.
Step two. Set the services available to the character
Step three . Log in to the AWS console to start using the S3 service and upload the download file via the web console.
Step four. Upload and download files via the AWS SDK.

4 Results and Analysis

Fig. 4 shows the feature curves generated by the real-time of sensor collection action where data generated by the original sensor. The new feature values derived by first order differential, second order differential, etc. We can observe that the curves have two distinct distinctions. In the case of walking motion, because the hand and foot have a fixed frequency of swinging, the data curve of the sensor will show a larger fluctuation in both the hand and foot data. While in the case of picking up/placing motion, because it is usually stopped before placing something, the data curve of the sensor in the foot obviously fluctuates less or even becomes 0. This study recorded and labeled the motion as a training data set and put it into Random Forest Model for prediction. Before putting into the prediction model, this study also add some data transformation, such as Fourier time domain to frequency domain feature changes. As shown in Fig. 5, the Fourier characteristics of gyroscope and accelerometer are shown. In the figure, the change of frequency domain is more obviously to be observe. When the motion is transformed, the change of frequency domain is deeper when picking up/placing and shallower when walking.

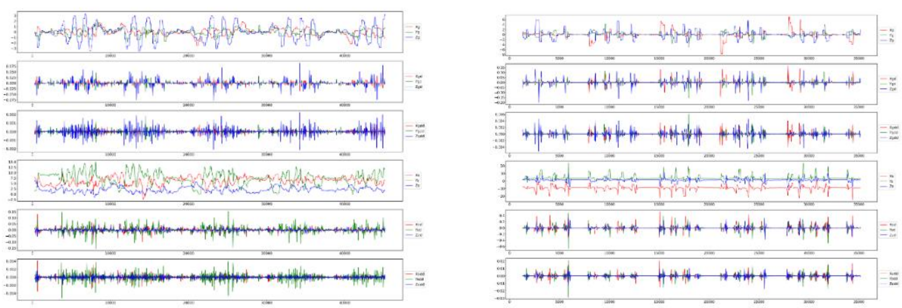

Fig. 4. Data collected by hand (left half) and foot (right half) sensors (Action: walking first, 5 times walking, 5 times picking up/placing)

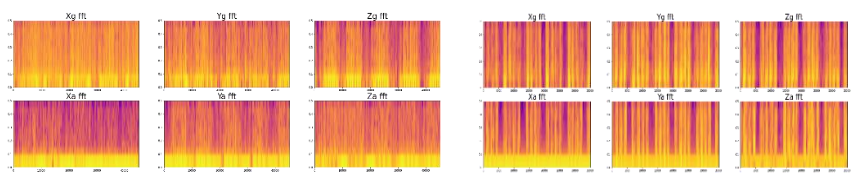

Fig. 5. Results of Fourier conversion of sensor data for hands (left half of 6 pictures) and foot (right half of 6 pictures)
(Action: walk first, 5 walks, 5 picks/places)

Table 1: Accuracy of prediction actions with different number of features

Number of features	6	12	18-1	18-2	24-1	24-2	36
Acc (Take/Put)	83%	75%	87%	86%	93%	92%	89%
Acc (Walk)	93%	88%	97%	94%	98%	98%	97%
F1 Score	82%	84%	87%	88%	94%	93.8%	91%

After the action dataset is labeled and set it as a control dataset (i.e., ground-truth), a separate action dataset is collected and used to predict the results. In the prediction results, the use of different number of features affects the accuracy of the prediction results. In Table 1, six features are mainly predicted by the three-directional gyroscope and three-directional accelerometer values of ESP32 itself. Moreover, the F1 score is 82%. While the 12 features are obtained by adding the Fourier transform of Gyro and Accelerometer to the first six features, and the F1 score is increased slightly to 84%. The 18-1 features are obtained by adding the first 12 features to the Fourier transform of Gyro and Accelerometer. Furthermore, the F1 score is increased slightly to 84%. The 18-1 feature is obtained by adding the first 12 features to the Fourier square root feature of Gyro and Accelerometer, which results in a small increase of 87% in F1 Score. The 18-2 feature is obtained by adding the first 12 features to the Fourier arctangent feature of Gyro and Accelerometer, which results in a small increase of 88% in F1 Score. Then the 18-2 feature is obtained by adding the first 12 features to the Fourier arctangent feature of Gyro and Accelerometer, which results in a small increase of 88% in F1 Score. score is 88%; while the 24-1 feature is obtained by adding the first 18-2 features to the first 18-2 features of Gyro and Accelerometer, and the F1 score is 94%. While the 24-2 feature is obtained by adding the first 18-2 features to the first 18-2 features of Gyro and Accelerometer, and the F1 score is 93%; and the 24-2 feature is obtained by adding the first 18-2 features to the first 18-2 features of Gyro and Accelerometer, and the F1 score is 94%. Moreover, the 24-2 feature is obtained by adding the first 18-2 features to the first 18-2 features of Gyro and Accelerometer. The F1 score is 93% for 24-2 features; and the F1 score is 91% for 36 features by adding the first 18-2 features to the first and second order differential features of Gyro and Accelerometer. From the analysis of the prediction results, it can be observed that the F1 score of the predicted motion has been gradually improved by adding the Fourier, Fourier inverse tangent function, Fourier square root, first and second order differential features of Gyro and Accelerometer to the original data values. However, if only 24-1 or 24-2 features are taken as the number of features compared to all 36 features, the prediction result is relatively the highest (up to 94%). Therefore, in the action prediction system, we use 24-1 features for the action prediction.

In the real-time prediction results, the training of the dataset requires individualized user data to train the model to be accurate. For each subject, this study collected one minute of action dataset individually and collected their data in a cycle of 6 seconds (walking for 5 seconds, picking up/placing for 1 second) for training. However, if the training dataset and the real-time identified users are different, the training prediction result is extremely inaccurate, only 36%. Therefore, the dependency of the dataset is significant when training the model. In addition, this application is known as data dependency in action prediction.

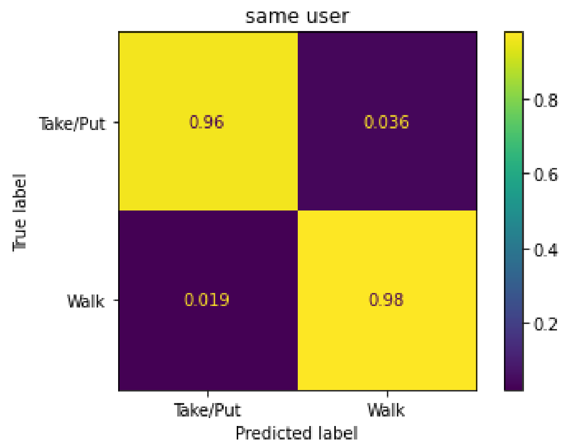

Fig. 6. Predicted results of the training datasets and the test for the same user in real time recognition action

5 Discussion

In conclusion, it can be found in the above results that the accuracy rate of this study is 97% by using 24 features. Compared with the previous study, although the accuracy rate was 98%, which used 80 features. This shows that the selection of key features is significant in the accuracy of prediction. The use of focused features for prediction can make the number of features trained by the model simpler. However, at the same time achieve high prediction results. This study also found that the accuracy of the pick-up/placement prediction is slightly lower than that of the walk prediction. Since each pick-up/placement data collection is less than the walking data set. In addition, there may be a small drop in the smooth or panning dataset during the pre-processing of the dataset, and the intervals and labels given during the action transitions are not quite accurate, which is something we should continue to improve.

In other words, the model in this study needs to import individual training sets for different users to do individual model training, which is the application of data dependency in motion recognition. If popularizing the device, this study will also try to find the average window size of motion recognition among different users to increase the generalizability of the sensor. In terms of model selection, this study currently uses Random Forest model for individual user prediction. In the Random Forest model, finding out a set of parameters that can achieve an accuracy rate of 97%. However, the random forest model does not perform as well as it does in the classification when solving the regression problem. Therefore, random forest does not give a continuous output. When it performs regression, this method cannot make predictions beyond the training set, which may lead to overfitting when modeling data with certain noise. Therefore, we will continue to try to incorporate other features to improve the accuracy of action prediction. We will also try to use different training modules such as SVM and XGBoost to predict actions to validate our model to achieve model stability.

References

1. B. E. Stopschinski, K. Del Tredici, S.-J. Estill-Terpack, E. Ghebremdehin, F. F. Yu, H. Braak, and M. I. Diamond, "Anatomic survey of seeding in Alzheimer's disease brains reveals unexpected patterns," Acta neuropathologica communications, vol. 9, pp. 1-19, (2021).

2. Fu, C. L. , Current status of dementia in Taiwan. Taiwan Journal of Geriatrics and Gerontology, 3(3), 169-181. Fu, Tsung-His and Pau-Ching Lu . Population aging and social policy in Taiwan. In Tsung-His Fu and Rhidian Hughes, eds. Population Aging in East Asia: Challenges and Policy Developments in the 21st Century (pp. 101-116). Taipei: Juliu. (2012)

3. Chen, D.-F., Prevalence of dementia in long-term care facilities in Taiwan. In Lee, M. B., et al. The Complete Dementia Handbook (pp. 53-56). Taipei: Health Culture. (2008)

4. F. Demrozi, G. Pravadelli, A. Bihorac, and P. Rashidi, "Human activity recognition using inertial, physiological and environmental sensors: A comprehensive survey," IEEE Access, vol. 8, pp. 210816-210836, (2020).

5. D. R. Beddiar, B. Nini, M. Sabokrou, and A. Hadid, "Vision-based human activity recognition: a survey," Multimedia Tools and Applications, vol. 79, pp. 30509-30555, (2020).

6. D. Bouchabou, S. M. Nguyen, C. Lohr, B. LeDuc, and I. Kanellos, "A survey of human activity recognition in smart homes based on IoT sensors algorithms: Taxonomies, challenges, and opportunities with deep learning," Sensors, vol. 21, p. 6037, (2021).

7. Yang, Yu-Jin. Multi-sensor System and Real-time Motion Recognition. Graduate Institute of Information Science and Engineering, National Chiao Tung University, pages 1-32,(2017).

8. Shu-Yu Chang. Motion interpretation of stairs, slopes, and flat surfaces using acceleration gauges and gyroscopes National Taiwan Normal University, College of Sports and Recreation, Department of Sports and Athletics, Master's thesis, pages 1-51, 2014

9. Emily J.Huang Jukka-Pekka Onela. Augmented Movelet Method for Activity Classification Using Smartphone Gyroscope and Accelerometer Data *Sensors* , *20*(13),3706; doi.org/10.3390/s20133(2020).