

Realtime Online Daily Living Activity Recognition Using Head-Mounted Display

<https://doi.org/10.3991/ijim.v11i3.6469>

Fais Al Huda

Brawijaya University, Malang, Indonesia
fais.developer@gmail.com

Herman Tolle

Brawijaya University, Malang, Indonesia
emang@ub.ac.id

Rosa Andrie Asmara

State Polytechnics of Malang, Indonesia
rosa_andrie@polinema.ac.id

Abstract—Human activity recognition is one of the popular research fields. The results of this study can be applied to many other fields such as the military, commercialism, and health. With the advent of the wearable head mounted display device mainly like google glass raises the possibility of this research. In this study tries to identify everyday activities are often called the ambient activity. Development of the system is done online using a smartphone and a head mounted display. The system produces an accuracy above 90%, which can be concluded that the system was able to recognize the activities with great accuracy.

Keywords—Android, head mounted display, accelerometer, sensor.

1 Introduction

Research on human activity recognition (HAR) is one of the research areas that are popular recently. Due to the results of this study can be applied to many fields such as military, commercial, and especially the health sector. One of the benefits of the activities recognition in the health field can be used to predict the falling potential of the user [1]. Other studies of HAR is HEMOCS which stands head movement controller system [2]. In this research the activities were recognized is the head movement of the user, the purpose of this study is to provide an alternative for users, especially users who have physical limitations and the elderly when interacting with a computer. Research on HAR also is used to recognize a person's gait that can be used to determine the gender of a user [3]. With so many benefits of HAR resulting in need of research in this field to be multiplied.

In the study of HAR, there are two types of approaches can be used. The first approach uses a visual approach, which in this approach utilizes media such as CCTV cameras and so on. One study in the form of a review for this approach carried out by Xu, in this study conducted an in-depth survey related to recognition activity-based video media and visual with sequential images [4]. The aspects covered include methods, systems, and is also related to the performance evaluation of the system. This approach has the disadvantage that needs large computing power and generally applied in the specific area [5]. The second approach is a wearable sensor-based approach to the user's body such as accelerometer sensor which is paired in the leg or the body of the user. This approach has the disadvantage that is obtrusive, so it will interfere if it is used to identify the daily activities of the user [6].

Various types of sensors attached to the smartphone have made studies of human activity recognition switch to this device. Because almost everyone has a smartphone and it has become part of daily life. By using a smartphone to do the human activity recognition, there are two modes, i.e., offline and online. Offline mode is a mode where the training algorithm is performed in the external environment such as on a personal computer. Weakness in this mode is the results strongly depend on sample data used for initial training thus when used by different user accuracy rate will decline and also since every smartphone has a different sensor capabilities can affect the classification result of the system.

The presence of Google Glass which is one of the wearable devices that type of head mounted display (HMD) is one sign that the days will come wearable device will be popular. To conduct HAR research on the devices like Google Glass can use other cheaper alternative. One such alternative is Google cardboard and its kind. Research on the recognition of human activities with the case of HMD has been done by Huda [7]. The study focused on the activities recognition that can be used as a game controller in augmented reality (AR) or virtual reality (VR).

This study proposes online and real-time recognition of daily activities by utilizing HMD. The algorithm used are backpropagation neural network combined with variable neighborhood search (VNS), by combining these two algorithms thus VNS is expected to help overcome the weaknesses of BPNN that are often trapped in local optima. That expected the recognition accuracy of the system able to get a high degree of accuracy.

2 Proposed Method

Activity recognition approach used in this study using the online mode. And also there are several procedures that necessary steps so that the pattern of the sensor can be classified.

2.1 Online Activity Recognition

By using the online approach mode, all recognition process until the classification is done on smartphones. Based on the study conducted by Shoaib et al.[8], the online

approach is divided into two categories: client-server approach and a local approach. In the first approach, the smartphone acts only as a client while the computation process is done on the server. This approach has the advantage that can make efficient use of high impact on clients. But on this approach depends a lot on the smartphone Internet connection between client and server as the main system. In the second approach, the entire process from data collection to the classification process is done on a smartphone. The overall process is shown in Figure 1.

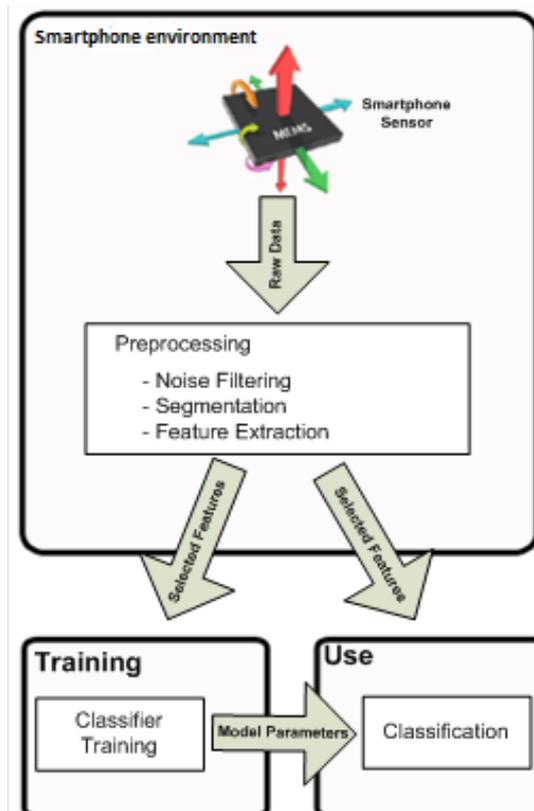


Fig. 1. Online activity recognition

2.2 Data Acquisition

The first preparation for data collection is to provide two main devices; there are smartphones and HMD. The smartphones which used in this research has the following specifications operating system Android 4.4 Kitkat with a processor which has the speed of 1.6 GHz and 1 GB of RAM memory. The process of data collection is done on average for 5 minutes for each activity. Data generated from the recording accelerometer sensor consisting of a three axis value in the form x, y, and z. List the activities that will be recognized are shown in Table 1.

Table 1. List of activity

No	Activity
1	Standing
2	Walking
3	Running
4	Jumping
5	Walking upstairs / downstairs

2.3 Noise Reduction

Data collected from the accelerometer sensor in the previous stage can not be directly used. Because such data has an element of noise coming from the gravitational effects are also caught by these sensors. Noise is a result of data that can reduce the accuracy of the system or increase the workload of the system, though not affect anything. To eliminate such noise can use the functionality of the Android API named `Linear_Acceleration`. This function can be used when the smartphone device must have an accelerometer sensor, magnetometer, and gyroscope.

2.4 Segmentation

The data has been cleared of noise has been able to be processed in the next process of segmentation. This process divides the sensor data in the form of time series into several parts called window. Based on some research done before there are several methods that can be used to process the sensor data segmentation. One study that discusses the segmentation process performed by Kozina [9]. In this study describes two methods of segmentation that is non-overlapping sliding window and sliding window with overlapping. The second method produces a higher degree of accuracy than the first method, but levels of computation required are also greater. Due to the system that will be developed in this study on the mobile platform thus the method used to use the first method for computing level needs lighter and faster process.

To get the response time in realtime then, a system must be able to process and respond to the signals that lead to activity patterns detected less than 0.5 seconds. Accelerometer capabilities on the devices used to have a maximum speed at `SAMPLING_RATE_FASTEST` mode of 94-101 Hz per second. Therefore, the length of the window by using sliding windows with overlapping used on systems with long window 32 which is a set of values axis of the accelerometer sensor, this value is assumed to be less than 0.5 seconds for recording time data so that it can achieve the level of response in realtime.

2.5 Feature Extraction

Once the data is completed in the segmentation of the next process is to look for the important features of the data. Of each value in the form x , y , and z searched the magnitude value according to equation (1).

$$m = \sqrt{x^2 + y^2 + z^2} \quad (1)$$

There are some additional features that are required in addition to the magnitude features. Selection of the features of the existing data in a study is very important because it will affect the accuracy of the results issued by the system. Based on research conducted by Bao, there are some features that are suited to the data obtained from the accelerometer sensor including the mean, standard deviation, and energy [10].

The mean is the average value of a set of values. The formula of the mean is shown in equation (2).

$$x = \frac{x_1 + x_2 + \dots + x_n}{n} \quad (2)$$

Standard deviation is a value that shows the distribution of the data, the higher the value, the higher the standard deviation also spread the values that exist in the data. Standard deviation formulas can be seen in equation (3).

$$\text{Standard deviation} = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2} \quad (3)$$

Whereas energy is the value of the sum of squared values of the magnitude of the data signal and the result of the sum is divided by the value of the window length. The formula of energy calculation shown in equation (4).

$$\text{energy} = \frac{\sum_{i=1}^{|w|} |x_i|^2}{|w|} \quad (4)$$

2.6 Training

Before the system can recognize patterns of data that have been obtained, it takes a training process which has been labeled a class is often called a classification process. Algorithm used is a derivative of a multilayer perceptron namely backpropagation neural network. This algorithm has a weakness often trapped in local minima [11]. Variable neighborhood search (VNS) is a metaheuristic technique for local search, this algorithm works sequentially and systematically to find solutions in neighbor areas, search positions will move when it obtained a better solution [12].

The architecture of backpropagation neural network algorithm consists of three layers, namely the input layer, hidden layer, and output layer. Each layer consists of neurons that represent the calculation of each of the layers. The number of neurons in each layer is different; the input layer consists of 34 neurons, seven neurons in the hidden layer, and three neurons in the output layer. This network architecture is illustrated in Figure 2.

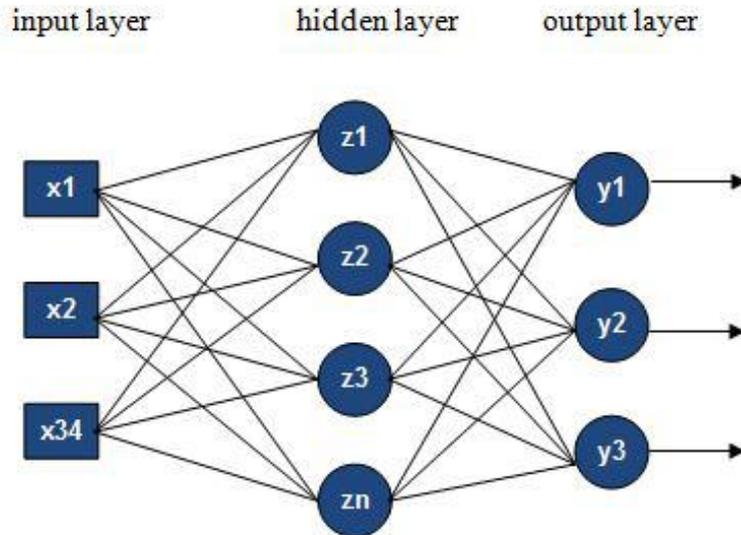


Fig. 2. Backpropagation neural network architecture

2.7 Classification

Once the training process is completed, it will yield model parameters that will be used by the classifier. Since the main algorithms used in this study is the backpropagation neural network then the parameters generated from the results of the training is the best weights for each neuron in the network.

3 Result and Discussion

For this study, we develop an application to perform all the steps described earlier. This application is developed using the Android operating system version 4.4 (Kitkat), but it is also possible to be used by later versions. The user interfaces for collecting activity data based on table 1 shown in Figure 3.

Data recording process begins with choosing which activities will be carried out by the user according to the option selected by activity radio button that has been provided on the user interface.

From the data recording process produce windows consisting of the calculation of the magnitude value of each axis of the accelerometer sensor. Each window is composed of 32 magnitude as well as additional features such as the mean, standard deviation, and energy, where the total of all these features which totaled 35 features form a feature vector that will be used for the algorithms training process. Results of the recording process of each activity are shown using the graph in Figure 4 to figure 8. X-axis represents the value of the sampling while the y-axis represents the magnitude of the acceleration values recorded by the accelerometer

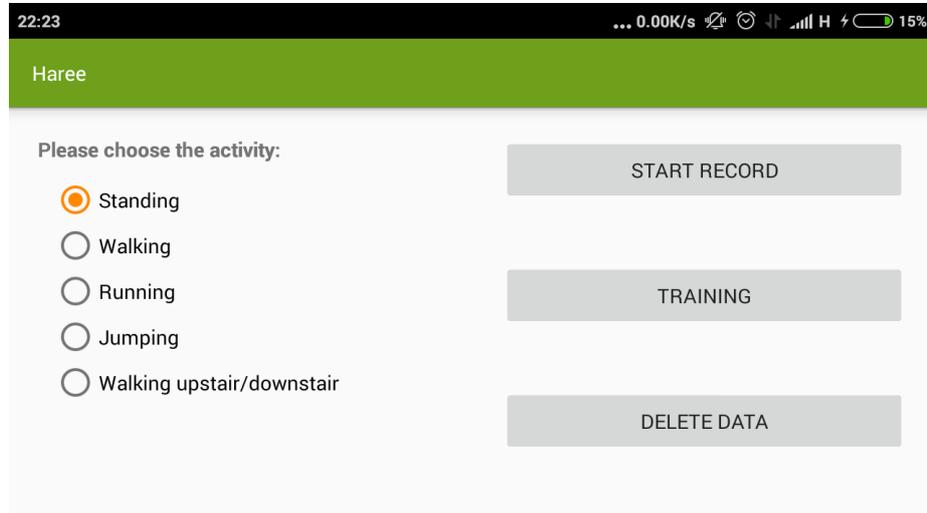


Fig. 3. User interface of data acquisition

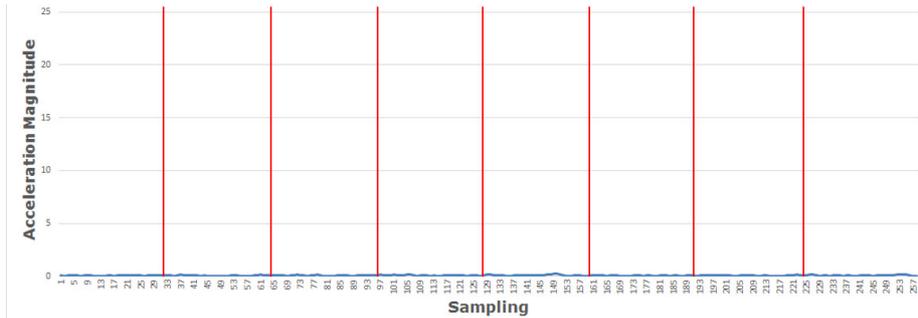


Fig. 4. Standing activity graph

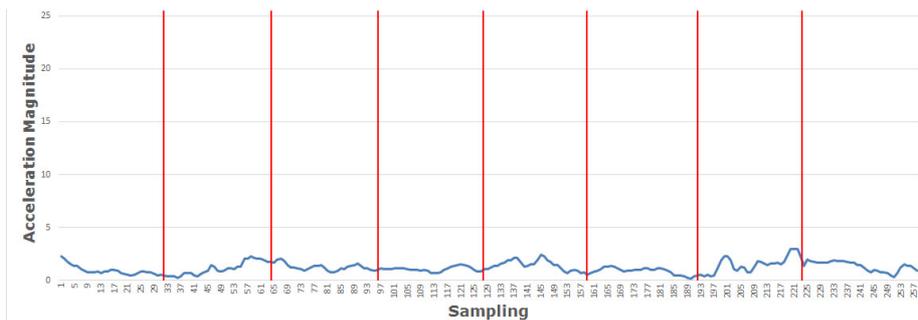


Fig. 5. Walking activity graph

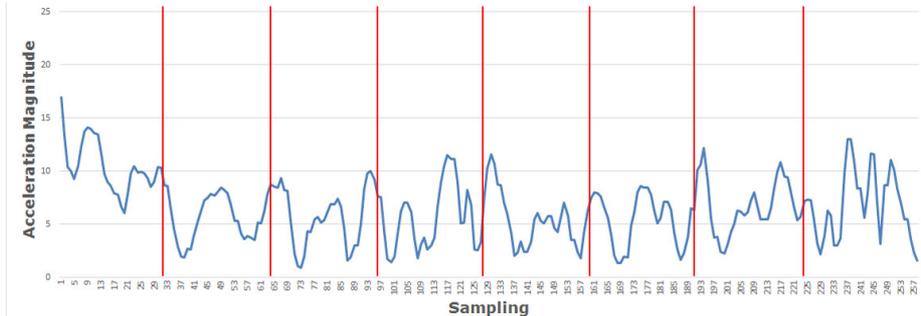


Fig. 6. Running activity graph

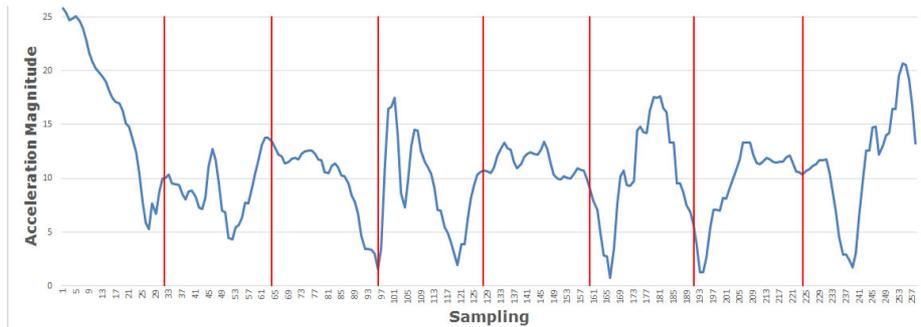


Fig. 7. Jumping activity graph

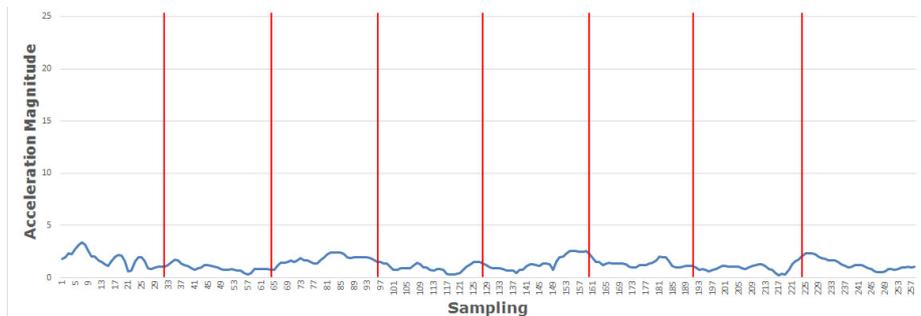


Fig. 8. Walking on stair activity

One of the most important implementation phases that are online training in a smartphone can be seen in Figure 9. On the user interface of the online training process, there are two functions namely the training and the save models. Training function serves to start the training process of the algorithm using data obtained from the activity data recording stage. While the models save function serve to store the system results of the training algorithm model.

One of the methods to measure the accuracy of prediction results from a system that is using the confusion matrix. This method is often used in research in the field of

machine learning to illustrate the performance of an algorithm developed. The columns in this metric represent the class prediction, while each row represents an actual class [13]. Confusion metric to show the accuracy of recognition activity of applications built in this research are shown in Table 2.

In figure 9 shows that the total amount of 2035 data used are derived from data from a total of five activities consist of standing, walking, running, jumping, and walking on stairs. With a total of correct prediction results for 1928 and the predictions wrong by 107.

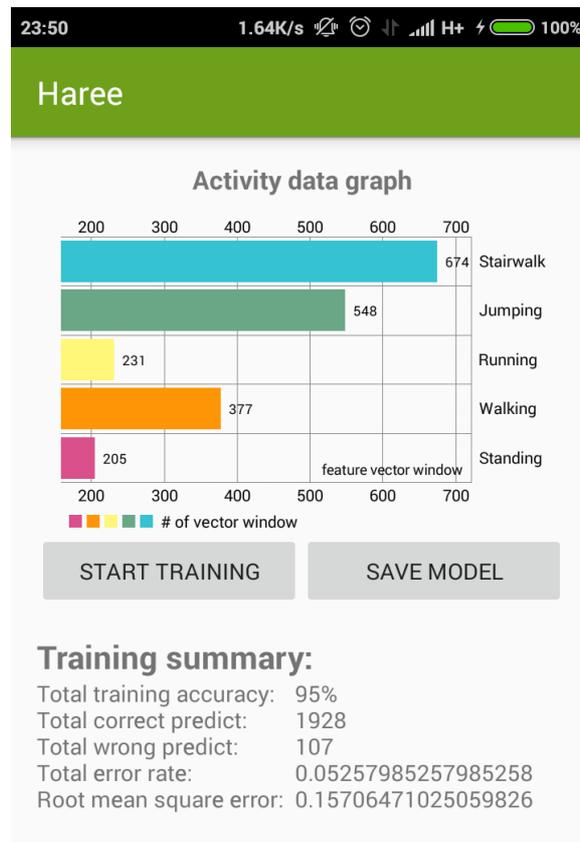


Fig. 9. User interface of online training

Table 2. Confusion matrix from the system

Activity	Standing	Walking	Running	Jumping	W.Stair
Standing	145	60	0	0	0
Walking	1	341	0	0	35
Running	0	0	211	7	13
Jumping	0	0	5	543	0
W.Stair	0	24	4	0	646

From the results shown by confusion metric can be concluded that the system can recognize the activities carried out with good results. With the best results when the system recognizes the jumping activity, because when this activity is performed there is a change in the sensor capture results significantly, making it easier to distinguish the pattern of this activity with others.

The results obtained from the algorithm model training process is used to identify the activity of users that will be displayed on the screen of a smartphone display put in the head mounted display. This screen is the main view of human activity recognition system developed in this research. HMD user interface is shown in Figure 4.

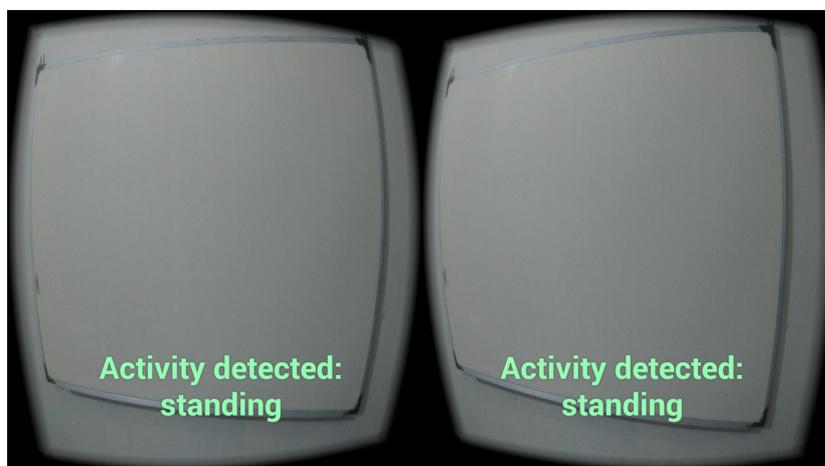


Fig. 10.HMD user interface

4 Conclusion

The recognition of human activity utilizing smartphones and HMD can be performed online. By using backpropagation neural network algorithm that is optimized for the variable neighborhood, search could produce an accuracy above 90%. It can be concluded this system has been getting good results in determining the activities performed by users. But this can be achieved due to the walking on the stairs activities does not differentiate between walking up the stairs and walking down the stairs.

The development of this next research is directed to distinguish between activities walking up stairs and down stairs because in this study still assumes up stairs and down the stairs into the same activity.

5 References

- [1] S. B. Kazi, S. Sikander, and S. Yousafzai. (2014). Fall Detection Using Single Tri-Axial Accelerometer. ASEE 2014 Zo. I Conf.

- [2] H. Tolle and K. Arai. (2016). Design of Head Movement Controller System (HEMOCS) for Control Mobile Application through Head Pose Movement Detection. *Int. J. Interact. Mob. Technol.*, vol. 10, no. 3, p. 24. <https://doi.org/10.3991/ijim.v10i3.5552>
- [3] K. Arai, J. S. City, and R. Asmara. (2014). Human Gait Gender Classification using 3D Discrete Wavelet Transform Feature Extraction. *Thesai.Org*, vol. 3, no. 2, pp. 12–17. <https://doi.org/10.14569/ijarai.2014.030203>
- [4] X. Xu, J. Tang, X. Zhang, X. Liu, H. Zhang, and Y. Qiu. (2013). Exploring Techniques for Vision Based Human Activity Recognition: Methods, Systems, and Evaluation Sensors. vol. 13, no. 2, pp. 1635–50. <https://doi.org/10.3390/s130201635>
- [5] D. Anguita et al.. (2013). Human activity and motion disorder recognition: Towards smarter interactive cognitive environments. *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 19, no. April, pp. 24–26.
- [6] J. Lester, T. Choudhury, and G. Borriello. (2006). A Practical Approach to Recognizing Physical Activities. *Pervasive Comput.*, vol. 3968, pp. 1–16. https://doi.org/10.1007/11748625_1
- [7] F. A. Huda, H. Tolle, K. P. Putra. (2017). Human Activity Recognition using Single Accelerometer on Smartphone for Head-Mounted Display. *IJASCA*, submitted for publication on March.
- [8] M. Shoaib, S. Bosch, O. Incel, H. Scholten, and P. Havinga. (2015). A Survey of Online Activity Recognition Using Mobile Phones,” *Sensors*, vol. 15, no. 1, pp. 2059–2085. <https://doi.org/10.3390/s150102059>
- [9] S. Kozina, M. Lustrek, and M. Gams. (2011). Dynamic signal segmentation for activity recognition. *Proc. Int. Jt. Conf. Artif. Intell.*, pp. 1–12.
- [10] N. Ravi, N. Dandekar, P. Mysore, and M. Littman. (2005). Activity recognition from accelerometer data. *Proc. Natl*, pp. 1541–1546, 2005.
- [11] M. Gori and A. Tesi. (1992). On the problem of local minima in backpropagation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 1. pp. 76–86. <https://doi.org/10.1109/34.107014>
- [12] N. Mladenović and P. Hansen. (1997). Variable neighborhood search. *Comput. Oper. Res.*, vol. 24, no. 11, pp. 1097–1100. [https://doi.org/10.1016/S0305-0548\(97\)00031-2](https://doi.org/10.1016/S0305-0548(97)00031-2)
- [13] Powers, M. W. David. Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness & Correlation. *Journal of Machine Learning Technologies*, vol. 2, no. 1, pp. 37-63

6 Authors

Fais Al Huda is with the Research Group of Multimedia, Game & Mobile Technology, Faculty of Computer Science, Brawijaya University, Malang, INDONESIA (e-mail: fais.developer@gmail.com).

Herman Tolle is with the Research Group of Multimedia, Game & Mobile Technology, Faculty of Computer Science, Brawijaya University, Malang, INDONESIA (e-mail: emang@ub.ac.id, herman.saga@gmail.com).

Rosa Andrie Asmara is a lecturer in the Informatics Management field, Department of Electrical Engineering at State Polytechnics of Malang, INDONESIA (e-mail: rosa_andrie@polinema.ac.id)

Article submitted 29 November 2016. Published as resubmitted by the authors 11 January 2017.