How to cite this paper:

M. N. Shah Zainudin, Raihani Mohamed, Md Nasir Sulaiman, Thinagar-an Perumal, Norwati Mustapha, & Azree Sharel Ahmad Nazri. (2017). Multi-label classification using label combination to recognize human activity based on various sensor positions in Zulikha, J. & N. H. Zakaria (Eds.), Proceedings of the 6th International Conference on Computing & Informatics (pp 669-674). Sintok: School of Computing.

MULTI-LABEL CLASSIFICATION USING LABEL COMBINATION TO RECOGNIZE HUMAN ACTIVITY BASED ON VARIOUS SENSOR POSITIONS

M.N.Shah Zainudin[1,2], Raihani Mohamed[1], Md Nasir Sulaiman[1], Thinagar-an Perumal[1], Norwati Mustapha[1] and Azree Sharel Ahmad Nazri[1]

[1]*Universiti Putra Malaysia, Malaysia,* raihanim@gmail.com, nasir@upm.edu.my, thinagaran@upm.edu.my, norwati@upm.edu.my, azree@upm.edu.my

[2]*Universiti Teknikal Malaysia Melaka, Malaysia,* noorazlan@utem.edu.my

ABSTRACT. Activity recognition using accelerometer sensor shows the good and positive impact in current health style perseverance. The sensor already built in the various smartphones, belts, and tapes for easy usage and applications in order to detect person's activity in real time response. Rampant research has been done to measure the effectiveness of accelerometer location detection of the activity. Hence, a separate method has been used to tackle the issues. This paper proposed multi-label classification (MLC) to look the effectiveness of the sensor location with a correlation of types of activity at the same time the similar activity could be discriminated accurately. The traditional single label works using Random Forest (RF) was used and the accuracy of the model will be compared with MLC using LC (Label Combination) with RF as a base classifier. As a result, MLC outperform the traditional single classifier approach to distinguish the accuracy for both stairs activities such as downstairs and upstairs is highly accurate using proposed method. At the same time, the pocket position successfully considered as the best sensor position to recognize various types of activities.

Keywords: activity recognition, smartphones, accelerometer, multi-label classification, Label Combination, Random Forest

INTRODUCTION

Human activity recognition (HAR) system plays important roles in the area of pervasive computing in the smart intelligent environment. In addition, recognizing daily human activity in the smart environment significantly improves the human life being since the number of diabetic patients among the world populations are drastically increases from time to time. In order to address this issue, people are encouraged to do some physical activities in their daily life. This required people to change their lifestyle to spend more time on exercises and live in a healthier lifestyle. Hence, there are several solutions to undertake this challenge. Researchers claim that recognizing human activities are varies depending on the situations. The numerous of sensor technology provide a good solution in HAR. Generally, there are three types of sensor that are widely used in HAR in smart environment namely; vision, ambient type embedded sensor and wearable sensor. Vision based sensor usually takes place when the camera is used to track and monitor the human activity. However, when the privacy of the human considered as a major concern, this kind of sensor might not a good solution. Furthermore, the conditions of the embedded sensor also affect the recognition of activity. This type of sensor as known as human interaction system since the interaction between the human and the sensor is essential. Motion, temperature, object, item, and humidity sensors are the example sensors that are used in detecting HAR. Unfortunately, the sensors need to be embedded in the fixed location in the house such as the door, shower tap, dining area and bedroom. The easiest way to recognize the HAR with a minimal number of sensors and cost is using a wearable sensor. The most effective sensor used to recognize the human activities reported by researchers are using accelerometer and gyroscope sensors. This micro-machine electromechanical system (MEMs) sensor has advantages for detecting the activity performed. The smaller size and cheaper cost of accelerometer and gyroscope sensors made it easy to own, wear and bring it anywhere. Thus the invention of multi functioning smartphones boosts the usage of activities tracking of HAR (Zainudin, Sulaiman, Mustapha, & Perumal, 2015). This kind of opportunity provides a good solution to the researchers to conduct more research in this area.

In order to recognize the human activity accurately, the selection of the classifier needs to be determined. In machine learning, there are several strategies might be applied in order to perform a classification process. Traditional classification strategy might be a good solution when dealing a multi-class problem. Determination of the human activity in classes is a good solution using this strategy. However, when the positions of the sensor are becoming another concern to know which is the best place to locate the sensor during the activity took place, this strategy might not a good solution since there are more than one class labels appears. Some of the cases reported that the position of the sensor attached significantly affect the performance of the recognition accuracy (Arif, Bilal, & Kattan, 2014; Shoaib, Bosch, Durmaz Incel, Scholten, & Havinga, 2014). In fact, not all activities performed will give the same accuracy result accordingly with the different sensor locations. As an example, running activity contributes to high accuracy when the sensor placed on the human thigh. However, the same positions of the sensor might not effective when the activity involves with hand gesture such as ironing or laundry cleaning. Another problem discovers during the study is the difficulty of distinguishing the similar activity in the different state. The study (Ronao & Cho, 2016) reported that the activity involve such as downstairs and upstairs hard to recognize. Hence, a multi-label problem may take places. There are several outcomes in this paper. First, a multi-label problem is applied to recognize various human activities with different sensor locations using single accelerometer sensor. Second, the result of a multi-label problem has been compared to the traditional multi-class problem. The results achieve high accuracy to tackle both problems. This paper organized as follows. Section 2 describes the materials and methods are used in this work. Section 3 presents the results and discussions that have been reported. Section 4 explains the conclusion for the overall experiment conducted.

MATERIALS AND METHODS

Human Activity Recognition Dataset

In this study, we used a public domain activity recognition dataset from authors (Shoaib et al., 2014). This dataset was utilized since the data collection was performed with different sensors positions on the human body to classify various types of human activities. Our work used Four Samsung Galaxy S2 smartphones for data collection to collect six different human activities; walking, sitting, standing, running, ascending stairs and descending stairs. All four participants were male age between 25 to 30 years old. The smartphones were placed in four different positions on their body; pocket, belt, arm, and wrist. All the activities performed conducted inside the building. Walking and running were performed in the department corridor, office space used for sitting activity and standing activity was collected during the coffee break. For ascending and descending activities, 5-floor stairs were used. Three types of sensors were used during data collection; accelerometer, gyroscope, and magnetometer. However, only accelerometer sensor data was used in this study since our aim is to minimize the number of sensors used in the experiment.

Data Segmentation and Extraction

To extract the features from the signal obtained, there is one process need to undertake in order to divide the raw time series signal data into several numbers of window segments. The common approaches that are widely applied in various applications are sliding window segmentation. In this approach, a raw signal will be divided into several equal sizes of window segment. There are two strategies used in this approach either with overlapping or without overlapping. In the first strategy, every window segment is overlap between two consecutive windows based on a certain percentage determined by the user. Otherwise, if there is no overlapping window between each other then, it is grouped under the second strategy. In this work, 50% overlap percentage was applied with window size 64s. Thus, it will provide 32s overlap between next following windows. This segmented window will be used to extract the features in the next following process. Originally, accelerometer signal record the signal in three different axes for different positions; x, y and z-axis. X record the horizontal movement, y record the vertical movement and z record the forward and backward movements. The aim of feature extraction is to extract several additional features from original data in order to give additional information for the class categories (Zainudin, Mohd Said, & Ismail, 2011). In any classifier models, it is difficult to classify the data with a very minimal number of features or characteristics. Hence, feature extraction may take places. Since time domain features easy and directly computed from the window segment, this approach was applied in this work. Several time domain features were extracted and calculated. The extracted features will be referred as a feature vector and will be adapted as an input for the classifier model. By using these additional extracted features, it will help the classifier to distinguish and learn the pattern of the class category. List of the time domain features used are maximum, minimum, average, standard deviation, variance, skewness and kurtosis for every x, y and z-axis. Meanwhile, the additional one feature namely correlation was extracted for xy, xz and yz-axis.

Traditional Classification Approach

There are various types of classifiers that usually applied in most classification or pattern recognition problem. The classification stage is required to evaluate the set of subsets in order to determine what is the class of those instances belongs. Extracted features from the previous section will be adopted as an input to the classifier. In this study, random forest (RF) was applied to classify physical activities (Arif et al., 2014; Breiman, 2001a). RF is an ensemble learning algorithm for classification, regression and other tasks, which operates by building a number of decision trees at training phase and outputting the class that is the mode of the classes (classification) of the individual trees or mean prediction (regression) of the individual trees (Feng, Mo, & Li, 2015). For validation purpose, 10-fold cross-validation strategy was utilized during this experiment. Default parameter values were used in our experiments in classification stage.

Multi-label Classification (MLC)

Traditional single-label classification is concerned with learning from a set of examples that are associated with a single label λ from a set of disjoint labels L, $|L| > 1$. In multi-label classification (MLC), the examples are associated with a set of labels $Y \subseteq L$. We compare the MLC algorithms such as BR (Binary Relevance) (Madjarov, Kocev, Gjorgjevikj, & Dzeroski, 2012), CC (Classifier Chains) (Jesse Read & Hollmen, 2014), BCC (Bayesian Classifier Chains)(Zaragoza, Sucar, Morales, Bielza, & Larranãga, 2011)(Enrique Sucar et al., 2014), LC (Label Combination) (J. Read, Martino, Olmos, & Luengo, 2015) and RA*k*EL (Random *k* label subsets) (Tsoumakas & Vlahavas, 2007) with random forest (Breiman, 2001b) as base classifier has been used to measure the accuracy of the model. The MLC requires different evaluation measures than traditional single-label classification. Hamming Score (HS), Exact match (EM), accuracy and accuracy per label has been selected to measure the methods.

Label Combination Method (LC). Label Combination previously known as Label Powerset (LC) is one of the fundamental problem transformations method which has been the focus of several works (J. Read et al., 2015). It is also an alternative paradigm to BR (Binary Relevance) method. LC transforms a multi-label problem into a single-label (multi-class) problem by treating all label combinations as atomic labels, i.e. each label set becomes a single class label within a single label problem. Thus, the set of single class labels represents all distinct label subsets in the original multi-label representation. Given a new instance, the single-label classifier of LP outputs the most probable class, which is actually a set of labels. If this classifier can output a probability distribution over all classes, then LP can also rank the labels. To obtain a label ranking we calculate for each label the sum of the probabilities of the classes that contain it. This way LP can solve the complete label correlations task.

Evaluation Metrics. For the purpose of comparison we used five different multi-label evaluation measures (Madjarov et al., 2012)(Zhang & Zhou, 2014)**,** specifically:

$$\text{Accuracy per label } := \frac{1}{L}\sum_{n=1}^{L} Acc_{(n)} = \frac{1}{N}\sum_{i=1}^{N} \delta\left(\hat{y}^{(n)}, y^{(n)}\right), \quad (1)$$

$$\text{Accuracy } := \frac{\mathbf{1}}{\boldsymbol{N}}\sum_{\boldsymbol{i}=\mathbf{1}}^{\boldsymbol{N}} \frac{|\boldsymbol{y}^{(\boldsymbol{n})} \wedge \hat{\boldsymbol{y}}^{(\boldsymbol{n})}|}{|\boldsymbol{y}^{(\boldsymbol{n})} \vee \hat{\boldsymbol{y}}^{(\boldsymbol{n})}|} \quad (2)$$

where $[\boldsymbol{A}]$ is an identity function, returning 1 if condition $\boldsymbol{A}$ is true, whereas $\wedge$ and $\vee$ are the bitwise logical AND and OR operation respectively.

$$\text{Hamming Score} := \frac{\mathbf{1}}{\boldsymbol{NL}}\sum_{\boldsymbol{n}=\mathbf{1}}^{\boldsymbol{N}}\sum_{\boldsymbol{j}=\mathbf{1}}^{\boldsymbol{L}}\left[\boldsymbol{y}_{\boldsymbol{j}}^{(\boldsymbol{n})} = \hat{\boldsymbol{y}}_{\boldsymbol{j}}^{(\boldsymbol{n})}\right], \quad (3)$$

$$\text{Exact match} := \frac{\mathbf{1}}{\boldsymbol{NL}}\sum_{\boldsymbol{n}=\mathbf{1}}^{\boldsymbol{N}}\left[\boldsymbol{y}^{(\boldsymbol{n})} = \hat{\boldsymbol{y}}^{(\boldsymbol{n})}\right], \quad (4)$$

RESULTS AND DISCUSSIONS

In this section, all the experiments conducted have been reported and discussed. As mentioned in the section previously, we conducted several experiments using traditional multi-class problem and multi-labels problem. As a result, Table 1 indicates the classification performance using RF classifier. Four performance indicators involving accuracy, precision, recall and f-measure were used to measure the performance of the classification result. In details, the table shows overall accuracy for different sensor position such as arm, belt, pocket and wrist. Belt position recorded the highest accuracy 95.6% followed by wrist position about 95.2%. Arm and pocket positions achieved 94.1% and 94.4% average accuracy respectively. Table 2 presents the classification performance for every single label of activity. It is clearly can be seen that most of the activities achieved good performance for all sensor positions that is above 95.3%. However, two stairs activities such as downstairs and upstairs recorded slightly lower. For downstairs, wrist and arm positions obtained accuracy 86.2% and 84% respectively. Belt and pocket positions recorded the lowest accuracy about 81%. In comparison with downstairs, 90.8% accuracy obtained for upstairs when the sensor positioned on the belt. Wrist, pocket and arm positions achieved accuracy 85% to 87%. In order to cater this issue, MLC will be exploited in the next experiment.

Table 1. Traditional Classification result on Accuracy using RF classifier

Accelerometer position	Arm	Belt	Pocket	Wrist
Accuracy	0.941	0.956	0.944	0.952
Precision	0.942	0.957	0.945	0.951
Recall	0.941	0.956	0.944	0.952
F-measure	0.941	0.956	0.944	0.951

Table 2. Traditional Classification result on Accuracy Per Label using RF Classifier

Accuracy per label	Arm	Belt	Pocket	Wrist
Downstairs	**0.84**	**0.816**	**0.812**	**0.862**
Running	0.977	0.96	0.957	0.989
Sitting	0.981	0.999	1	0.994
Standing	0.983	0.997	0.998	0.983
Upstairs	**0.851**	**0.908**	**0.863**	**0.87**
Walking	0.953	0.99	0.965	0.956

From the shortcoming results of single label classifier experiments, we use the same datasets to experiment using MLC using RF as a base classifier. Table 3 shows the MLC accuracy per label with its performance evaluation using HS and EM. Overall, LC obtains the highest results with 98.9% HS and 94.6% EM and indicates the highest result accuracy and accuracy per label among the others such as BCC, BR, CC and RA*k*EL. From this, we can conclude that by using MLC, we are able to distinguish the most very similar activities such as downstairs and upstairs. Hence the result shows the improvement of every classifier recorded more than 93.6% for downstairs and above 95.9% for upstairs. At the same time, the position of sensor indicated that the pocked showed 100% accuracy for every type of activities performed. This reveals that by using MLC with LC approach using Random Forest as base classifier significantly improved the performance of the similar activities and showed that the pocket is the most accurate location to recognize HAR.

Table 3. Classification Accuracy using MLC

MLC /Labels		BCC	BR	CC	LC	RA*k*EL
Accuracy per label	downstairs	0.963	0.962	0.936	0.97	0.966
	sitting	0.997	0.997	0.997	0.997	0.997
	standing	0.993	0.993	0.993	0.994	0.994
	running	0.992	0.992	0.992	0.992	0.992
	upstairs	0.96	0.961	0.959	0.965	0.963
	walking	0.976	0.981	0.97	0.978	0.978
	arm	0.996	0.996	0.996	0.996	0.996
	belt	0.999	1	0.999	0.999	0.999
	pocket	1	1	1	1	1
	wrist	0.995	0.995	0.995	0.995	0.995
Accuracy		0.962	0.96	0.962	0.969	0.965
HS		0.987	0.988	0.986	**0.989**	0.988
Exact match		0.927	0.913	0.934	**0.946**	0.931

CONCLUSION

In this work, we present an extensive experimental evaluation of single label and MLC methods to classify human activity. The datasets consist of six different of activities collected from various sensor positions such as arm, belt, pocket and wrist. We minimize the number of sensors used only accelerometer sensor without required any additional sensor attached to human bodies. Previous works mostly focusing on classification of activities using traditional classifier approach. However, by using traditional single classifier approach, we found the difficulties in distinguishing the similar activities such as downstairs and upstairs. Furthermore, another challenging is to determine the most effective sensor positions for classifying various types of activities. Hence, using MLC it proved that the overall accuracy increased at the same time the problem encounter using the prior methods also improved. Pocket position can be considered as the most effective sensor positions for recognizing various types of daily activities.

REFERENCES

Arif, M., Bilal, M., & Kattan, A. (2014). Better Physical Activity Classification using Smartphone Acceleration Sensor. http://doi.org/10.1007/s10916-014-0095-0

Breiman, L. (2001a). Random forests. *Machine Learning*, *45*(1), 5–32. http://doi.org/10.1023/A:1010933404324

Breiman, L. (2001b). Random forests. *Machine Learning*, *45*(1), 5–32. http://doi.org/10.1023/A:1010933404324

Enrique Sucar, L., Bielza, C., Morales, E. F., Hernandez-Leal, P., Zaragoza, J. H., & Larrañaga, P. (2014). Multi-label classification with Bayesian network-based chain classifiers. *Pattern Recognition Letters*, *41*(1), 14–22. http://doi.org/10.1016/j.patrec.2013.11.007

Feng, Z., Mo, L., & Li, M. (2015). A Random Forest-Based Ensemble Method for Activity Recognition, 5074–5077.

Madjarov, G., Kocev, D., Gjorgjevikj, D., & Dzeroski, S. (2012). An extensive experimental comparison of methods for multi-label learning. *Pattern Recognition*, *45*(9), 3084–3104. http://doi.org/10.1016/j.patcog.2012.03.004

Read, J., & Hollmen, J. (2014). A Deep Interpretation of Classifier Chains. In *ADVANCES IN INTELLIGENT DATA ANALYSIS XIII* (Vol. 8819, pp. 251–262).

Read, J., Martino, L., Olmos, P., & Luengo, D. (2015). Scalable Multi-Output Label Prediction: From Classifier Chains to Classifier Trellises. http://doi.org/10.1016/j.patcog.2015.01.004

Ronao, C. A., & Cho, S.-B. (2016). Human activity recognition with smartphone sensors using deep learning neural networks. *Expert Systems with Applications*. http://doi.org/10.1016/j.eswa.2016.04.032

Shoaib, M., Bosch, S., Durmaz Incel, O., Scholten, H., & Havinga, P. J. M. (2014). *Fusion of smartphone motion sensors for physical activity recognition*. *Sensors (Switzerland)* (Vol. 14). http://doi.org/10.3390/s140610146

Tsoumakas, G., & Vlahavas, I. (2007). Random k-labelsets: An Ensemble Method for Multilabel Classification. *European Conference on Machine Learning*, 406–417. http://doi.org/10.1007/978-3-540-74958-5_38

Zainudin, M. N. S., Mohd Said, M., & Ismail, M. . (2011). Feature Extraction on Medical Image Using 2D Gabor Filter. *Applied Mechanics and Materials*, *52–54*, 2128–2132. http://doi.org/10.4028/www.scientific.net/AMM.52-54.2128

Zainudin, M. N. S., Sulaiman, N., Mustapha, N., & Perumal, T. (2015). Activity Recognition based on Accelerometer Sensor using Combinational Classifiers, 68–73.

Zaragoza, J. H., Sucar, L. E., Morales, E. F., Bielza, C., & Larranãga, P. (2011). Bayesian chain classifiers for multidimensional classification. In *IJCAI International Joint Conference on Artificial Intelligence* (pp. 2192–2197). http://doi.org/10.5591/978-1-57735-516-8/IJCAI11-365

Zhang, M., & Zhou, Z. (2014). A review on multi-label learning algorithms. *IEEE Transactions on Knowledge and*, *26*(8), 1819–1837.