



Feature Selection in Recognizing Daily Living Activities Based on Metabolic Equivalent

Nhan D. Nguyen¹, Binh A. Nguyen², Viet Q. Tran², Khoa D. Ta², Manh Hoang², Giao N. Pham³

¹Dept. of Biomedical Engineering, Sungkyunkwan University, Suwon, South Korea, nhannd@skku.edu

²ICT Department, FPT University, Hanoi, Vietnam, binhase04865@fpt.edu.vn, vietqtse06178@fpt.edu.vn, khoatdhe130813@fpt.edu.vn, manhhhe130294@fpt.edu.vn

³Dept. of Computing Fundamentals, FPT University, Hanoi, Vietnam, giaopn@fe.edu.vn

ABSTRACT

Nowadays daily living activity recognition is an important part in many applications. In this paper, we present a method to recognize three groups of activities which are labeled as light intensity, moderate intensity, and high intensity activities. We also investigate the effect of the statistical and physical features to each group. In experimental results, we obtain an accuracy of 89.30%, 66.20%, and 87.89% for light, moderate, and high intensity activities, respectively.

Key words: Activity Recognition, Accelerometer Data, Data Normalization, Feature Selection, Metabolic Equivalent, and Machine Learning.

1. INTRODUCTION

Daily living activity recognition has become a trend in human activity recognition field. Many researchers are working on this field due to its importance in real applications such as health monitoring, ambient intelligent system, and surveillance [1]. In [2], Bao et al. used bi-axial accelerometer placed on five different body locations to classify 20 activities of daily life and the accuracy of classification is 84%. To improve performance of activity recognition, [3] applied correlation based feature selection technique on the extracted features to classify 6 daily activities and obtained 87% of accuracy with top 8 selected features.

In this paper, we proposed a novel method to classify three groups of activities, i.e., light intensity, moderate intensity, and high intensity activities by using physical and statistical features. We first apply a low-pass filter to the raw data. Then, we implement feature selection algorithm on three groups of activities. These groups are created by ordering the ratios of metabolic rate during perform physical activity to a reference metabolic rate (MET) [4]. The MET are divided into three level, i.e., low (MET < 3), medium (3 < MET < 6), and high (MET > 6) which are equivalent to light intensity, moderate intensity, and high intensity activities [5]. The features are then normalized using feature standardization. Finally, K-nearest neighbor algorithm (KNN) is employed to perform

the activity recognition. For the light and high intensity activities we obtained 89.30% and 87.89%, respectively. In moderate intensity activities, the accuracy is 66.20%. To clarify the proposed method, we organize our paper as follow. The proposed method is described in Sec. 2. Experimental results and the evaluation of the proposed solution will be described in Sec. 3. Sec. 4 shows the conclusion.

2. THE PROPOSED METHOD

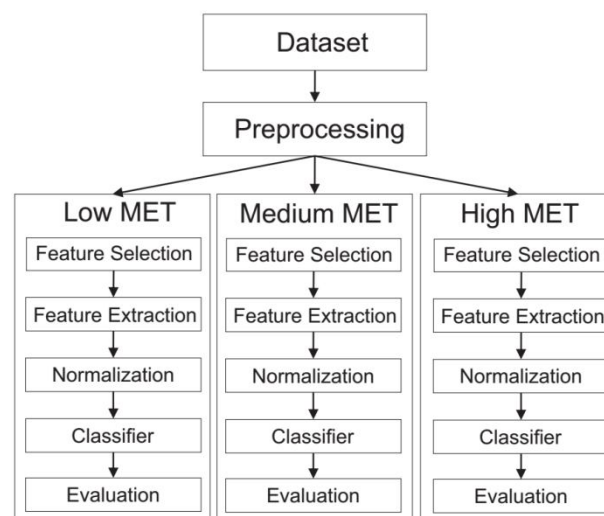


Figure 1: The proposed method

Table 1: Activities and Metabolic Equivalent of Task (MET)

Activity	Intensity [MET]
Sitting	1.3
Lying	1.0
Standing	1.3
Washing dishes	2.5
Vacuuming	3.3
Sweeping	3.3
Walking	3.5
Ascending stairs	5.0
Descending stairs	3.5
Treadmill running	9.0
Bicycling on ergometer 50 W	3.5
Bicycling on ergometer 100 W	6.8
Rope jumping	8.8

Table 2: Statistical and physical features are used in the present work

Statistical Features	Physical Features
Mean	Signal Magnitude Area
Standard Deviation	Root Mean Square
Variance	Intensity of Movement
Mean Absolute Deviation	Sum of Absolute values
Skewness	Average Energy [7]
Kurtosis	Mean and variance of
4 th Moment	Movement Intensity
5 th Moment	

The proposed method is shown in Fig. 1. We used DaLiAc (Daily Living Activities) dataset which includes 13 activities of 19 healthy subjects [6]. Four Shimmer sensors, consist of a 3D accelerometer and a 3D gyroscope, are collect the data from different body location with sampling frequency is 204.8Hz (right hip, chest, right wrist, and left ankle). The MET intensity in the DaLiAc that was considered are show in Tab. 1.

The choice of features that are extracted from raw data is essential to make the machine learning algorithms more efficient and accurate. Tab. 2 shows the list of statistical and physical features we consider in our work. We obtained a total of 272 features including 192 statistical features and 80 physical features. Feature selection technique is used to assess the usefulness and identify the most important features for discriminating different activity groups [8]. In this paper, we use Laplacian Score [9]-[12] which is an unsupervised feature selection algorithm.

In our work, we used K-nearest neighbor algorithm to classify activities .To validate the classification, we used the leave-one-out cross-validation approach which leaves one sample in the dataset as a test set and the remaining samples as the training set. This procedure was repeated for all samples.

3. EXPERIMENTAL RESULTS

Table 1: Classification accuracy of features each MET level.

MET	Stats Features		Physical Features	
	No. Features	Accuracy (%)	No. Features	Accuracy (%)
Low	13	89.23	31	89.30
Medium	40	66.24	22	65.59
High	24	87.81	36	87.61

Table 4: Number of selected features and overall classification accuracy for each MET level.

MET	No. of Selected Features	No. of Physical Features	No. of Stats Features	Accuracy (%)
Low	16	3	13	89.30
Medium	16	9	7	66.20
High	10	7	3	87.89

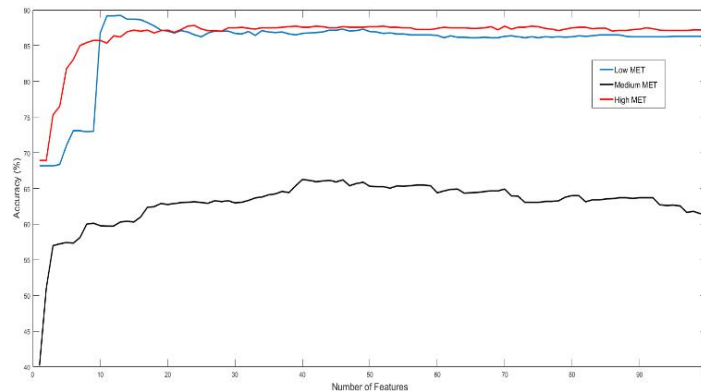


Figure 2: Classification accuracy rates of feature selection with statistical features.

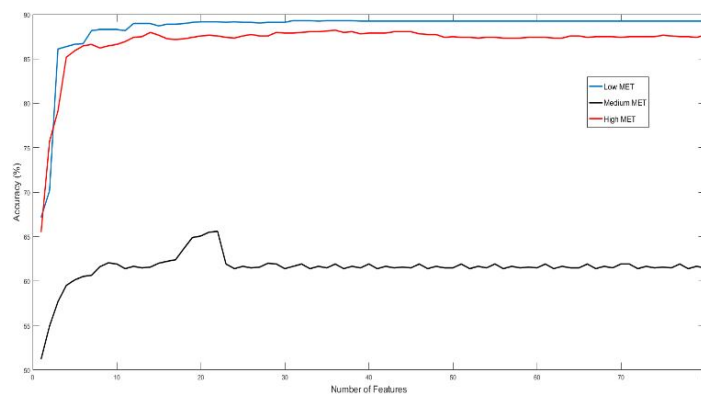


Figure 3: Classification accuracy rates of feature selection with physical features.

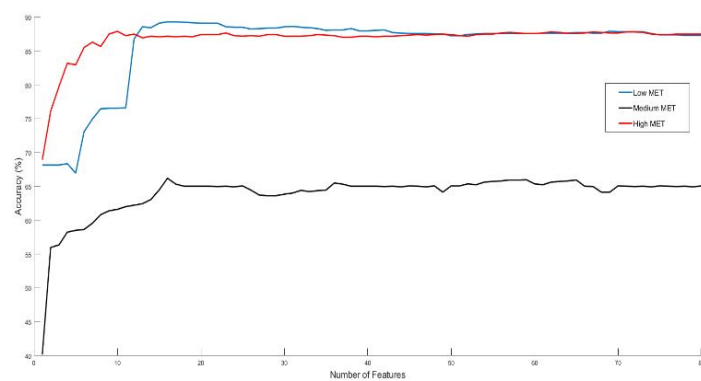


Figure 4: Classification accuracy rates of feature selection with physical and statistical features.

We applied the proposed method to DaLiAc dataset. The KNN algorithm is applied for different kind of features: The statistical features, the physical features and both of these two. Experimental results are shown from Fig. 2 to Fig. 4. Tab. 3 shows the number of statistical and physical features verse accuracy for low, medium, and high MET. To recognize light intensity activities, we can use 13 statistical features to achieve an accuracy rate of 89.23%. A low accuracy rate of 65.59% was obtained by using 22 physical features for moderate intensity activities. In high MET group, 24 statistical and 36 physical features were used to obtain accuracy of 87.81% and 87.61%, respectively.

Tab. 4 shows that to obtain the equivalent accuracy of activity classification, the total number of features when using both physical and statistical features are smaller than when using these kind of features separately which are showed in Tab. 3. The results also shows the importance of physical features of high intensity activities (7 in total 10 features). For light intensity activities, the statistical features are most important (13 in total 16 features).

4. CONCLUSION

In this paper, we proposed a method to recognize three groups of activities which are labeled as light intensity, moderate intensity, and high intensity activities. In the other hand, we focus on the selection of important features depends on the specific group of activities. In the future, we intend to develop a robust activity recognition system by using these results.

ACKNOWLEDGEMENT

This work is supported by Sungkyunkwan University, Suwon, Republic of Korea; and by the FPT University; Hanoi, Vietnam.

REFERENCES

1. Kanna. **Activities of Daily Life (ADL) Recognition using Wrist-Worn Accelerometer**, International Journal of Engineering Technology, Vol. 8, pp. 1406–1413, 2016.
2. L. Bao. **Activity Recognition from User-Annotated Acceleration Data**, Pervasive Computing, pp. 1–17, 2004.
https://doi.org/10.1007/978-3-540-24646-6_1
3. U. Maurer. **Activity Recognition and Monitoring using Multiple Sensors on Different Body Positions**, BSN, pp. 113–116, Cambridge, MA, USA, 2006.
<https://doi.org/10.21236/ADA534437>
4. Royall. **Appendix 1: Translating Scientific Evidence about Total Amount and Intensity of Physical Activity into Guidelines, 2008 Physical Activity Guidelines for Americans**, United States Department of Health and Human Services.
5. B. E. Ainsworth. **Compendium of Physical Activities: Classification of Energy Costs of Human Physical Activities**, Medicine and Science in Sports and Exercise, Vol. 25, pp. 71-80, 1993.
<https://doi.org/10.1249/00005768-199301000-00011>
6. H. Leutheuser. **Hierarchical, Multi-sensor Based Classification of Daily Life Activities: Comparison with State-of-The Art Algorithms using a Benchmark Dataset**, PLoS ONE, Vol. 8, No. 10, pp. 1-11, 2013.
<https://doi.org/10.1371/journal.pone.0075196>
7. S. Gonzalez. **Features and Models for Human Activity Recognition**, Neuro Computing, Vol. 167, pp. 52-60, 2015.
<https://doi.org/10.1016/j.neucom.2015.01.082>
8. M. Zhang. **A Feature Selection-Based Framework for Human Activity Recognition using Wearable Multimodal Sensors**, in Proc. of the 6th International Conference on Body Area Networks, pp. 92-98, 2011.
<https://doi.org/10.4108/icst.bodynets.2011.247018>
9. X. He. **Laplacian Score for Feature Selection**, NIPS, 2015.
10. N. A. Nasharuddin, N. S. Mohd, S. K. Ali. **Multi-feature Vegetable Recognition using Machine Learning Approach on Leaf Images**, International Journal of Advanced Trends in Computer Science and Engineering, Vol. 8, No. 4, pp. 1789-1794, 2019.
<https://doi.org/10.30534/ijatcse/2019/110842019>
11. A. E. Karrar. **A Novel Approach for Semi Supervised Clustering Algorithm**, International Journal of Advanced Trends in Computer Science and Engineering, Vol. 6, No. 1, pp. 1-7, 2017.
12. S. W. Kareem and M. C. Okur. **Bayesian Network Structure Learning Based On Pigeon Inspired Optimization**, International Journal of Advanced Trends in Computer Science and Engineering, Vol. 8, No. 1, pp. 131-137, 2019.