Random Forest for Human Daily Activity Recognition

Nurul Retno Nurwulan¹, Gjergji Selamaj²

¹Department of Industrial Engineering, Sampoerna University
Jl. Raya Pasar Minggu Kav. 46, Jakarta, 12780, Indonesia
²Department of Economics, Universiteti Ismail Qemali
Skelë, Rruga Kosova, Vlorë 9401, Albania

nurul.nurwulan@sampoernauniversity.ac.id

Abstract. Machine learning classifiers are often used to evaluate the predicting accuracy of human activity recognition. This study aimed to evaluate the performance of random forest (RF) compared to other classifiers with considering the time taken to build the models. Human activity daily living data, namely walking, walking upstairs, walking downstairs, sitting, standing, and lying down were collected from smartphone-based accelerometer with sampling frequency of 50Hz. The dataset was evaluated using artificial neural network (ANN), k-nearest neighbors (KNN), linear discriminant analysis (LDA), naïve Bayes (NB), support vector machine (SVM), and random forest (RF). The results of the study showed that RF indeed predicted the activities with the highest accuracy. However, the time taken to build the models using RF was the second-longest after ANN.

1. Introduction

Human activity recognition is a popular topic because of its significance for various applications such as healthcare [1-3], sports science [4], surveillance [5], etc. In the healthcare industry, human activity recognition can be used to develop the correct diagnosis for patients. Whereas in sports science, it can be used to optimize athletes’ training and minimize the potential of injuries. The detection of human movements such as walking and running can also help surveillance and other purposes.

Human activity recognition can be done using the help of a triaxial accelerometer. There are various commercially available triaxial accelerometers. However, it might not comfortable for long-term usage and the motion sensors are sometimes costly. Smartphones with built-in triaxial accelerometer are the more practical and cheaper way to monitor human daily activity recognition [6]. Then the time human signal data collected needs to be analyzed and classified into several activities based on each pattern of the data. It is a challenging problem because there is no clear way to relate the signal data to a specific activity. The common approach is by extracting the features of the time series data and training machine learning models to classify each activity.

Although the machine learning method is considered robust and comprehensive [7], there is no certain classifier to be guaranteed as the best one with high accuracy in recognizing human activity. Past studies evaluated the performance of several classifiers in categorizing the activities accurately, and the results vary between studies. For instance, the K-Nearest Neighbors (KNN) showed a satisfactory result in the study by Bedogni et al [8]. However, in a comparative study, the KNN did not show a good result in classifying the activities [7]. Random forest (RF) was said to outperformed other classifiers, including the KNN, in predicting human activities correctly [7, 9]. Random forest is a decision tree’s classifier with the ability to improve the classification results by combining the bootstrap aggregating method and randomization in the selection of segmenting data nodes [9]. The
disadvantage of decision tree classifier is the time taken in model updating. Thus, the time spent for the analyses might not that worthy regardless of the high accuracy.

The performance evaluation in the past studies was mostly based on the accuracy of predicting the activities without considering the time taken in building the classifier models to classify the activities. Some classifiers might need a lot longer processing time than the other classifiers with slightly lower accuracies. This trade-off should be taken into consideration, especially if the gait signal data are collected in a long timeframe and the data need to be analyzed are very large. With that being said, this study aimed to evaluate the performance of random forest compared to other machine learning classifiers by considering the time taken to build the models.

This paper is organized as follows. Section 2 describes the methodology in obtaining the dataset and a brief description of the theoretical background behind features extraction and machine learning classifiers. Section 3 presents results and discussion. Finally, the conclusion is presented in section 4.

2. Methodology

The dataset used to evaluate the performance of random forest compared to other classifiers was taken from the study by Anguita et al. [6]. The data was collected using the triaxial accelerometer built-in Samsung smartphone. The dataset was collected from thirty subjects aged between nineteen and forty-eight and the smartphone was attached to the waist with the sampling frequency rate was set to 50 Hz. The subjects were asked to do walking, walking upstairs, walking downstairs, sitting, standing, and lying down. The dataset was pre-processed with a fifty percent overlapping window of 2.56s using a median filter and a third-order low pass Butterworth filter with a corner frequency of 20 Hz to remove noise.

2.1. Feature extraction

Two different modes of features can be extracted from human physiological data, namely time- and frequency-domain features. Time-domain refers to the variation of the amplitude of the signal with time. This mode does not require signal pre-processing. Whereas the frequency-domain is more robust, and it needs signal pre-processing such as the Fast Fourier Transform (FFT) [10]. In this study, the features are mean, standard deviation, mean absolute deviation, maximum, minimum, signal magnitude area, energy, interquartile range, and spectral entropy in both the time- and frequency-domain modes.

The mean acceleration data shows the overall effects of the activity. It is best to capture the subtle shifts in static accelerations from time period equal to or greater than the interval parameter [11]. The standard deviation describes the spread of the measurement values from the mean value. The low standard deviation shows the values are mostly close to the average, whereas the high standard deviation shows the values are spread out. This feature is often used to see the stability of the acceleration signal [12]. The mean absolute deviation is the mean of the absolute deviations from the mean acceleration to summarize the dispersion of the acceleration data. The maximum and minimum are the value of the peak of the acceleration data, it can be used to evaluate the variation of the acceleration as a function of time [11]. The signal magnitude area is a measure of the magnitude of a varying quantity. It is used to distinguish the moving activity from static activity [13].

Energy is calculated as the sum of squared discrete FFT to capture the data periodicity to discriminate the static activity from the dynamic activity such as walking [14].

\[
F(k) = \frac{1}{N} \sum_{j=0}^{N-1} b(j) e^{i 2 \pi k j / N}
\]

where, \( k = (0, 1, \ldots, N-1) \)

The interquartile range is a variability measure by dividing the dataset into quartiles, it is the first quartile subtracted from the third quartile. The last feature in this study is the spectral entropy to quantify the level of order in time series data. This feature is able to quantify the gait dynamics due to its sensitivity in quantifying the level of order in time-series data by distinguishing respective activities...
with similar energy values [3, 14]. The Shannon entropy is used as the feature, it utilizes the probability of the signal in discrete states to evaluate the repetitions in a signal data [14].

\[ H = -\sum_{i=0}^{N-1} p_i \log_2 p_i \]  

(2)

where, \( p \) is the probability of the event.

2.2. Classifiers

In this study, we compared the random forest (RF) to other classifiers: artificial neural network (ANN), k-Nearest neighbors (KNN), Fisher’s linear discriminant analysis (LDA), naïve Bayes (NB), and support vector machine (SVM). We used the Weka environment [15] by the University of Waikato to do the machine learning analyses.

The multilayer perceptron (MLP), a class of ANN, was used in this study. An MLP consists of at least three layers of nodes: an input layer, a hidden layer, and an output layer. The hidden and output layers utilize a nonlinear activation function. The MLP uses backpropagation for training, it is a powerful supervised learning technique that can distinguish data that are not linearly separable [16].

The KNN is a fast and simple supervised classifier since it does not require a learning process [17]. The data classification is done by determining the similarity between the training set and new observation that is assigned to the most similar class based on the most votes [18]. The similarity is decided by the Euclidean distance. Despite its simplicity, the KNN has been proven to be an effective classifier for human activity recognition [8].

The LDA is a supervised linear dimensionality reduction method that uses linear combinations to extract lower dimensional features [18]. It is used to ensure the projections of samples from different classes to a line are well-separated. The LDA is able to separate the classes with an obvious distance between mean values and small variances [18].

The NB is a simple supervised method with the assumption of conditional independence of the features [19]. It assumes each pair of features contribute independently regardless of the correlations between the value of the class variable. Although it has shown great results in medical applications due to its ability to do enough data pre-processing, past studies found that the independence assumption of the NB is not suitable for gait signals from accelerometer data [10, 20].

The SVM is a discriminative supervised machine learning method that places a decision hyperplane efficiently in pattern space to distinctly classify the patterns of the data points into different classes. The difference with the LDA is that the SVM selects the hyperplane with maximum margin distance [20]. Maximizing the margin will give more enforcement to classify the data points confidently. However, misclassifications are expected to happen in a few training data points when they cannot be separated perfectly.

The RF is a supervised method that constructs a large number of uncorrelated decision trees as an ensemble at training time with the output is the classification of individual trees [21]. It improves the classification performance by feature bagging which is the modification of the decision tree algorithm to select a random subset of the features [9, 21]. Each individual tree has a class prediction and the class with the most votes will become the model prediction. The RF is an effective method to rank the importance of variables in a classification problem naturally [15]. Unlike the other decision tree classifiers, each tree in the RF can only select a random subset of features that makes it possible to increase the variation among the trees in the model. Thus, the classification will result in higher accuracy considering the low correlation across trees. Although the RF could yield a better accuracy in the classification [7, 9], the large number of trees can make the algorithm too slow and ineffective for real-time predictions.

3. Results and Discussion

This study used both the time- and frequency-domain features in order to see which mode is better for acceleration data. The FFT was used to produce the frequency-domain features. The features were categorized into three different groups: time-domain, frequency-domain, and the combination of both time- and frequency-domain features. Then, these groups were evaluated using ANN, KNN, LDA,
NB, SVM, and RF. In this study, we were not only evaluating the accuracy of the classification but also considering the time taken to build the machine learning models. Table 1 shows the results of the evaluation.

| Table 1. Performance evaluation of the classifiers |
|-------------------------------------------------|
| Accuracy (%)                                    |
| Time-domain                                     |
| ANN     | 84.05 | 83.24 | 83.94 | 82.08 | 83.25 | 84.78 |
| KNN     |       |       |       |       |       |       |
| LDA     | 85.49 | 83.77 | 85.88 | 83.25 | 85.32 | 85.94 |
| NB      |       |       |       |       |       |       |
| SVM     | 85.99 | 85.06 | 85.84 | 82.75 | 85.79 | 87.16 |
| RF      |       |       |       |       |       |       |
| Frequency-domain                               |
| Time-domain                                     |
| ANN     | 10    | 0     | 0.01  | 0.01  | 0.15  | 2.42  |
| KNN     |       |       |       |       |       |       |
| LDA     | 7.76  | 0     | 0.01  | 0.01  | 0.15  | 1.88  |
| NB      |       |       |       |       |       |       |
| SVM     | 20.95 | 0     | 0.09  | 0.02  | 0.2   | 2.75  |
| RF      |       |       |       |       |       |       |
| Combined                                       |

Based on the calculation results, it is evident that the RF indeed outperformed the other classifiers in all data groups with the combined time- and frequency-domain features yielded the highest accuracy. This result is in agreement with previous studies [7, 9]. However, the time taken to build a model was the second-longest after the ANN.

In the time-domain group, only the ANN and RF resulted in an accuracy of 84%. Whereas in the frequency-domain group, the ANN, LDA, SVM, and RF derived the accuracy of 85%. The combined features group of time- and frequency-domain yielded the highest accuracy for all classifiers but the NB. It is interesting that in NB as a classifier, the frequency-domain is the group with the highest accuracy, and the accuracy decreased when the time-domain features were added.

From the evaluation result, we can see that the frequency-domain features give better accuracy than the time-domain features. It is because the frequency-domain features can analyze the non-linear signal data such as the acceleration data. The frequency-domain features can be used as the corrective measurement for time series data with noise disturbance in the system and it is able to capture the slight changes in the system [22].

As for the time taken to build the machine learning models, we can see that the KNN did not need time to build the model because the KNN does not need to train a model since it is an instance-based learning algorithm. The NB only needed 0.01s, 0.01s, and 0.02s to build the model for the time-domain, frequency-domain, and combined group, respectively. This is because the NB does not require the setting of tuning parameters. The third fastest classifier was the LDA with 0.01s, 0.02s, and 0.09s to build the model for the time-domain, frequency-domain, and combined group, respectively. The LDA is a fast classifier with great results in this study because the acceleration data are not linear [16]. The SVM yielded in somewhat similar accuracy as the LDA but it is slower than the LDA. In terms of classification, the LDA selects the hyperplane that separates all data points. While the SVM selects the hyperplane that separates only the points in the frontier between the classes [18]. Thus, the SVM requires a longer time than the LDA.

The time taken to build a model for the RF was 2.42s, 1.88s, and 2.75s for the time-domain, frequency-domain, and combined group, respectively. As we can see here, the frequency-domain features required less time than the time-domain and combined groups. The ANN required the longest time to build the models with 10s, 7.76s, and 20.95s for the time-domain, frequency-domain, and combined group, respectively. The ANN is time-consuming compared to the other classifiers although it is the second-best classifier in terms of accuracy. The training of the ANN is very extensive, and the training process is repeated until a good enough set of models are discovered.

The evaluation provides the trade-offs between accuracy and time taken to build machine learning models. From the evaluation, it is evident that the RF is preferable compared to the ANN since it yielded in higher accuracy but a faster time to build the models. Generally, the RF is the best classifier in terms of accuracy. However, if the data are larger and real-time, it might be better to use faster classifiers such as the LDA and SVM. In this study, the KNN also showed promising results although the accuracy is not as high as the RF.
4. Conclusion

This paper has presented the performance evaluation of random forest (RF) compared to other machine learning techniques, namely the artificial neural network (ANN), k-Nearest neighbors (KNN), Fisher’s linear discriminant analysis (LDA), naïve Bayes (NB), and support vector machine (SVM). In this paper, we compared both the accuracy and the time taken to build the machine learning models to see the trade-offs between the two. Based on the evaluation results, it can be concluded that the RF indeed outperformed the other machine learning techniques in terms of accuracy. However, the time taken to build the model is longer than the KNN, LDA, NB, and SVM. In conclusion, the RF is highly recommended for human daily activity recognition. However, if the data are large, the RF is a time-consuming method.

References

[1] Osmani V, Balasubramaniam S and Botvich D 2008 Human activity recognition in pervasive health-care: Supporting efficient remote collaboration J. Netw. Comput. Appl 31 628-655
[2] Tentori M and Favela J 2008 Activity-aware computing for health 2008 IEEE Pervasive Comput 7 51-57
[3] Nurwulan N R and Jiang B C 2016 Possibility of using entropy method to evaluate the distracting effect of mobile phones on pedestrian Entropy 18 390
[4] Hsu Y L, Yang S C, Chang H C and Lai H C 2018 Human daily and sport activity recognition using a wearable inertial sensor network IEEE Access 6 31715-31728
[5] Shao L, Ji L, Liu Y and Zhang J 2012 Human activity segmentation and recognition via motion and shape analysis Pattern Recognit. Lett 33 438-445
[6] Anguita D, Ghio A, Oneto L, Parra, X and Reyes O J L 2013 A public domain dataset for human activity recognition using smartphones 21th European Symposium on Artificial Neural Networks, Computational Intelligence and machine Learning, ESANN
[7] Nurwulan N R 2020 Performance evaluation of decomposition methods in perturbed walking Revue d'Intelligence Artificielle 34(1) 45-50
[8] Bedogni L, di Felice M and Bononi L 2012 By train or by car? Detecting the user’s motion type through smartphone sensors data Proceedings of the 2012 IFIP Wireless Day (WD), Dublin, Ireland 1-6
[9] Maroco J, Silva D, Rodrigues A, Guerreiro M, Santana I, de Mendonça A 2011 Data mining methods in the prediction of dementia: A real-data comparison of the accuracy, sensitivity and specificity of linear discriminant analysis, logistic regression, neural networks, support vector machines, classification trees and random forests BMC Research Notes 4 299
[10] Nurwulan N and Jiang B C 2020 Window selection impact in human activity recognition International Journal of Innovative Technology and Interdisciplinary Sciences 3(1) 381-394
[11] Rogers M J B, Hrovat K, McPherson K, Moskowitz M E and Reckart T 1997 Accelerometer data analysis and presentation techniques National Aeronautics and Space Administration NASA Technical Reports 19970034695 1-48
[12] Chen C, Grennan K, Badner J, Zhang D, Gershon E and Jin L 2011 Removing batch effects in analysis of expression microarray data: An evaluation of six batch adjustment methods PLoS ONE 6(2) e17238
[13] Karantonis D M, Narayanan M R, Mathie M, Lovell N H and Celler B G 2006 Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring IEEE Trans. Inf. Technol. Biomed 10 156-167
[14] Maekawa T, Yanagisawa Y, Kishino Y, Ishiguro K, Kamei K, Sakurai Y and Okadome T 2010 Object-based activity recognition with heterogeneous sensors on wrist Pervasive 2010 246-264
[15] Weka 3.8.1. 2016 Hamilton, New Zealand: The University of Waikato
[16] Breiman L, Friedman J, Stone C J and Olshen R A 1984 Classification and Regression Trees Boca Raton, FL: CRC Press
[17] Zhang T, Wang J, Xu L and Liu P 2006 Using wearable sensor and NMF algorithm to realize ambulatory fall detection In Advances in Natural Computation 488-491
[18] Duda R O, Hart P E and Stork D G 1999 *Pattern classification* Malden, MA: John Wiley & Sons
[19] Cortes C and Vapnik V 1995 Support-vector networks *Machine Learning* **20**(3) 273–297
[20] Lara O D and Labrador M A 2013 A survey on human activity recognition using wearable sensors *IEEE Communications Surveys and Tutorials* **15**(3) 1192–1209
[21] Ho T K 1995 Random decision forests *3rd International Conference on Document Analysis and Recognition, Montreal* 278-282
[22] Rahmati M T 2016 Comparison of frequency domain and time domain methods for aeromechanical analysis *International Journal of Rotating Machinery* **2016** 1-12