Journal of Industrial Information
Technology and Application Vol.5 No.3 *ISSN(Online) :2586-0852*

Behavior pattern recognition based on convolutional neural network using accelerometer and gyroscope sensors

JunHyuk Kang, Yeseul Lim, Ahyoung Choi*)
Department of Software, Gachon University, Seongnam, Republic of Korea

Abstract. Recently, research on recognizing a human behavior pattern based on accelerometer and gyroscope signals has been done. However, the existing studies were mainly conducted to analyze the extracted features for behavior pattern recognition as inputs. In this study, we propose a model that recognizes and classifies behavior patterns by using raw signals including noise as inputs of the convolutional neural network(CNN) model to which optimal hyperparameters are applied through the grid search algorithm. The behavior patterns were divided into 4-classes. To evaluate the proposed model, 12 data were collected, and it was confirmed that the average accuracy was 93%.

Keywords; wearable devices; human behavior; accelerometer signal; gyroscope signal; convolutional neural network;

1. Introduction

Recently, various studies have been actively done to analyze and recognize behavior patterns using smart devices such as mobile phones and wearable devices. The existing studies used methods of analyzing data after collecting data through sensors built into mobile phones or smart devices [1,2]. Also, research about the effect of sensor placement on behavior pattern recognition has been conducted [3,4]. Extracted feature of an accelerometer signal was used as an input to a deep learning model[5]. Anahita et al. measured accelerometer and gyroscope data of 25 children using a smartwatch [5]. Children wore smartwatches and performed six different activities: running, walking,

---

* Corresponding author: aychoi@gachon.ac.kr
Received: Feb. 10, 2021; Revised: Mar. 6, 2021; Accepted: Sep. 30, 2021

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (No. NRF-2019R1F1A1060731)

standing, sitting, lying down, and climbing stairs. The activities were measured at 10 Hz of 10-minute intervals. They used extracted feature such as Mean, Median, FFT-entropy, and Signal Vector Magnitude etc from the measured data as an input to the deep learning model. Six activities were classified using two deep learning techniques, DNN and RNN. As a result, the RNN model showed an average F1 score of 80%.

Unlike previous studies, recent trend uses the end-to-end model. The raw data are use as an input to the model. Existing studies have used only the accelerometer sensor data or extracted features of the accelerometer sensor and gyroscope sensor data. However, using a single type of sensor data can make accurate recognition of action difficult and using feature extracted data limits deep learning models from learning features that humans cannot detect. If the data-preprocessing is not performed, the model can be trained with better features because the model can directly extract and learn features from raw data. Thus, in this paper, raw data is used as the input of the 1D-CNN model. Although the characteristics of the data for each class are time series, the accuracy of recognition was analyzed through a convolutional neural network model because the characteristics of the data can be categorized into each classes.

2. Proposed method

We collected data with a sampling rate of 100 Hz per second and data at 50 Hz per second. To insert an input of the same size, the two different sampling rates are adjusted through the resampling process to 50Hz per second through the down-sampling process. After that, to obtain an input signal from the raw data, an input cut by 10 seconds is obtained through a sliding window of 10 seconds.

The deep neural network model shows better performance by finding more information than the shallow neural model [7]. In this paper, to train the feature of 3-axis accelerometer signal and 3-axis gyroscope signal uses 1D-CNN. The proposed model is shown in Fig 1. The first hidden layer has 32 filters, batch normalization is used to prevent overfitting and ReLu is applied as the activation function. The second hidden layer has 64 filters, like the first layer batch normalization is used to prevent overfitting and applied ReLu as the activation function. Adaptive average pooling was used to reduce the dimension of the features and after two stages of fully connected layers, the final output was obtained using softmax.

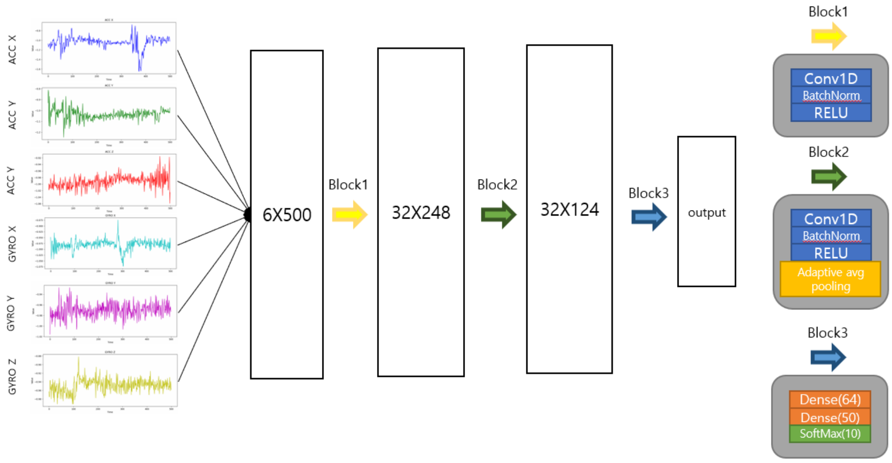

Fig 1. The proposed CNN model

3. Experiments

A. Experimental setup

To obtain the data required for analysis, we propose the following system as shown in Fig 2. The system was developed as an Android app. For data measurements, it runs apps on a smartphone and wearable watch. In a smartphone, select the wearable watch that you want to connect by pressing the "connect to watch" button and connect a smartphone to the wearable watch through Bluetooth. If the connection is successful, the toast message confirms that the connection was successful. In a wearable watch, a sensor manger is created when the app runs. The smartphone mainly operates such as starting measurements, stopping measurements, and changing labels, and the commands selected on the smartphone are delivered to the wearable watch through Bluetooth. By commands from smartphones, data measurement, transfer to database servers, and label changes are made in wearable watches. Our system uses Firebase as its database server and utilizes Cloud firestore and authentication API.

In this work, we classify the behavior into 4 level. Level 1 is low level of bahavior, which includes lying down, sitting, and standing. Level 2 is light level of bahavior, which includes crossing, bending, and finding the stuff out of the box and take it out. Level 3 is locomotion, which includes walking and jumping in place. Level 4 is a vigorous sport, which includes kicking and running. We measured data based on several activities using a wearable watch, fosil sport. After wearing the wearable watch on the participant's left wrist, he performed four stages of bahavior, collecting acceleration, gyro, heart rate, and step count. The collected data include movements to press the start and stop buttons at the beginning and end of the measurement, respectively, so care should be taken when using them in the analysis. Each bahavior was measured for five minutes, taking a total of 50 minutes. In 12 measurements, 3,558 segments were obtained for use as input.

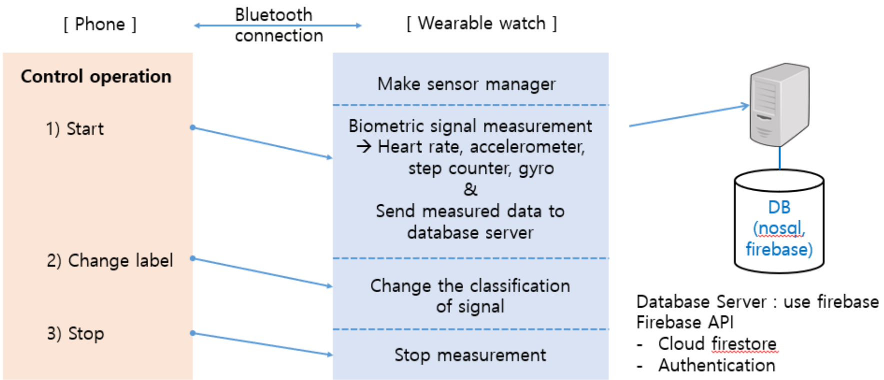

Fig 2. Data collection system architecture

B. *Experimental result*

Table1 shows the results of 4-bahavior classification. The total average accuracy of the 4-bahavior classification is 94.7%. The reason for the classification of Level 4, which shows relatively low performance, found through the confusion matrix. As shonw in Fig.3, a part of Level4 was classified into Level 2 and Level 3. In other words, kicking bahavior (Level 4) was incorrectly classified into bending bahavior (Level 2) and jump in place bahavior (Level 3).

TABLE I. BAHAVIOR CLASSIFICATION RESULT

Class	**Accuracy**
Level 1 Low-level bahavior	100%
Level 2 Light level bahavior	98%
Level 3 Locomotion bahavior	93%
Level 4 Vigorous sports	86%
Total	**94.7%**

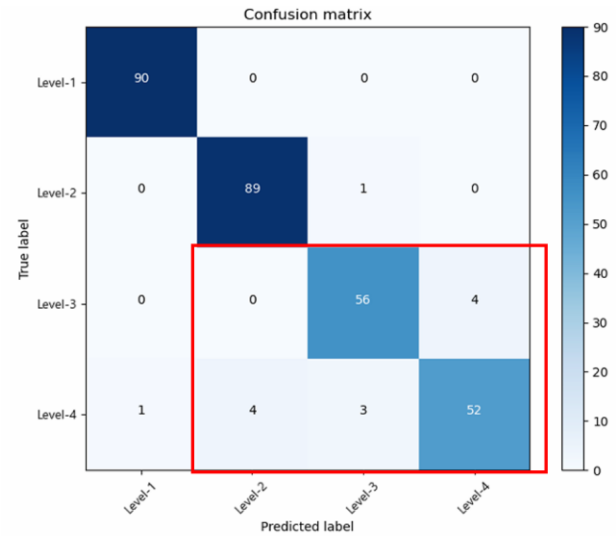

Fig 3. Confusion matrix of behavior classification

4. Conclusion and future works

In this study, we propose human bahavior recognition through end-to-end CNN model. As the proposed model is an end-to-end model, 4 actions were effectively classified using only raw accelerometer and raw gyroscope signal without separate feature extraction. In this study, we observed that the accuracy was greatly improved to 92.6% for 4 behavioral classifications. In future research, we plan to improve the classification performance of bahavior 4, which shows relatively low classification performance, by using heart rate sensor data and step counter data as inputs, and the model will be applied to people of various ages, from children to adults.

References

[1] G. Bieber, J. Voskamp, and B. Urban, “Activity Recognition for Everyday Life on Mobile Phones,” HCII 2009, LNCS 5615, pp. 289–896, 2009.

[2] M. R. Puyua, A. L. Adolph, F. A. Vohra, I. Zakeri and N. F. Butte, “ Prediction of Activity Energy Expenditure Using Accelerometers in Children,” Med Sci Sports Exerc, vol. 36, no.9, pp. 1625–31, Sep 2004.

[3] A. Mannini, M. Rosenberger, W. L.Haskell, A. M. Sabatini, and S. S. Intille, “Activity Recognition in Youth Using Single Accelerometer Placed at Wrist or Ankle,” Med Sci Exerc, vol. 49, no.4, pp. 801–812, Apr 2017.

[4] S. G Tros, Y. Zheng, and W. Wong, “Machine learning for activity recognition : hip versus wrist data,” Physiological Measurement, vol. 35, no.11, pp. 2183–2189, 2014.

[5] A. Hosseini, S. Fazeli, E. Valencia, R. Habre, M. Sarrafzadeh, A. Bui, “Children Activity Recognition: Challenges and Strategies,” International Conference of the IEEE Engineering in Medicine and Biology Society (EMC), IEEE, pp.4331–4334, 2018.

[6] M. Ahmadi, M. O’Neil, M. Fragala-Pinkham, N. Lennon and S. Trost, “Machine learning algorithms for activity recognition in ambulant children and adolescents with cerebral pasly,” Journal of NeuroEngineering and Rehabilitation, 15, 105, 2018.

[7] J. Schmidhuber, “Deep learning in neural networks: An overview,” Neural networks, vol. 61, pp. 85–117, Jan 2015.