Machine Learning and Artificial Intelligence
J.-L. Kim (Ed.)

doi:10.3233/FAIA230780

A New Method for Human Activity Recognition of Photoplethysmography Signals Using Wavelet Scattering Transform

Ziyu WANG [a], Feifei LIU [a,1], Shengxiang XIA [a,2], Shuhua SHI [a], Lin WANG [b], Zheng XU [a], Sen AI [a] and Zhengyong HUANG [a]
[a] *School of Science, Shandong Jianzhu University, Jinan, 250101, China*
[b] *The Fifth People's Hospital of Jinan, Jinan, China*

Abstract. Exercise is an indispensable part of people's lives and is closely related to their health. Human Activity Recognition (HAR), which involves detects and analyzes human body activity, has become the focus of current research. Photoplethysmography (PPG) has advantages such as convenience for detection and low cost, and is widely used in wearable devices becoming an ideal choice for HAR. In this study, we used wavelet scattering transform (WST) to extract features from PPG and then performed activity recognition on it. We achieved excellent classification accuracy of 92.54% and 97.76% respectively in the experiments of three-class and four-class exercise detection. The results showed this method based on wavelet scattering transform and PPG can accurately detect exercise types and provide effective support for HAR.

Keywords. Photoplethysmography, Human Activity Recognition, Wavelet Scattering Transform, Deep Learning, Support Vector Machine

1. Introduction

Human Activity Recognition (HAR) refers to the process of automatically identifying and classifying human activities, has been an important research domain. Humans' physical activity condition is a reflection of their physical health. By collecting and analyzing HAR data, can help people live a healthier lifestyle and prevent various chronic diseases[1]. So HAR has enormous development potential in several fields such as health care and monitorization of patient.

In the past, activity recognition often required professional and cumbersome equipment. But with the development of integrated circuits and data analysis, a large number of sensors have been integrated into wearable devices, which has brought about the rapid popularization of activity detection [2]. Wearable devices can conveniently acquire various physiological and motion signals of the human body, provide variety

[1] Corresponding author, Feifei Liu, School of Science, Shandong Jianzhu University, China. E-mail: liufeifei19@sdjzu.edu.cn
[2] Corresponding author, Shengxiang Xia, School of Science, Shandong Jianzhu University, China. E-mail: xias@sdjzu.edu.cn

service including exercise intensity detection, sleep quality estimation, for users [3]. Among them, photoplethysmography (PPG) sensor is also one of the most common sensors in wearable devices [4]. These raw signals are quickly and simultaneously obtained, making HAR easier.

PPG is a non-invasive method thought use of optical sensors to detect the blood volume changes with every cardia cycle, and is usually used to measure blood pressure, heart rate rhythm, hemoglobin oxygen saturation, etc., [5, 6]. Compared with the traditional cuffed sphygmomanometer, the PPG acquisition device is more convenient to carry and use, and does not need professional medical staff [7]. Compared to using various motion sensors or electrocardiogram (ECG) for activity recognition, PPG signals are easier to obtain and reduce the number of sensors and reduced costs.

In recent years, many studies have attempted to use PPG for HAR. T. Mahmud et al. proposed a multi-stage neural network based on long short-term memory (LSTM) for human activity recognition [8]. The network extracted the time features from various physiological signals in the early stage and fused multimodal features in the later stage. The recognition accuracy can reach 83.20%. However, the recognition accuracy based on PPG was only 72.1%. E. Brophy et al. used the transfer learning method to classify the signal image of PPG, and obtained the accuracy of 90.8%. [9]. But plotting PPG images inevitably requires greater computation, which reducing the detection speed.

However, collecting PPG signals in motion contains a large amount of noise, and motion artifact (MA) is the biggest problem limiting PPG for activity recognition [10]. So, it is very important to extract features from PPG before classification. A preprocessing step was proposed by M. Boukhechba et al. that they decomposed PPG into three types of signals: cardiac, respiration, and MA to predict different ambulatory activities [4]. T. Aydemir et al. used Hilbert transform to extract features, detected three types of exercise, and compared the classification results of various classifiers [11]. Under their three-class experiments of exercise detection, an average classification accuracy rate of 89.39% was obtained.

In this study, we presented using wavelet scattering transform (WST) to extract features from PPG signal for the first time. WST is a new signal processing method, which was proposed by Mallat [12]. It can provide translation invariant and stable time-frequency resolution, and fully preserves the high-frequency information of the signal [13]. WST can simultaneously extract low-frequency and high-frequency features from signals, and the data volume of the extracted features is small, which is more suitable for detection. We extracted scattering features from PPG signals and generate scattering matrix, which were fed into support vector machine (SVM) for classification. The results showed that PPG signals transformed by WST can detect activity accurately. And the proposed method has a simple process, can greatly improves the detection accuracy in both three-class and four-class detection experiments.

2. Methodology

2.1. Database and Preprocessing

The PPG database used in this study was Wrist PPG During Exercise database. It was obtained by D. Jarchi et al. of the University of Manchester in 2017 [14]. There are a total of 8 participants in database, including 3 males and 5 females, aged from 22 to 32 years old. Each participant was asked to perform a minimum of 1 type and a maximum of 4 types of exercise, including walking, running, low-resistance cycling, and high-resistance cycling, with each exercise lasting no more than 10 minutes. The database consists of 19 recordings, length of each recording varies from 4 minutes to 10 minutes, with a sampling rate of 256 Hz. PPG recording was obtained through a sensor placed at the wrist of participants, and the device also synchronously recorded signals from gyroscopes, low-noise accelerometers and wide-range accelerometers. In addition, ECG recordings from the chest were also recorded at the same time. More details are shown in Table 1.

Table 1. Duration of each recording.

Subject ID	Walk	Run	Low-Resistance Bike	High-Resistance Bike
1	9:48	-	9:39	9:48
2	6:39	-	5:41	6:54
3	4:47	5:07	4:54	4:41
4	-	4:52	-	-
5	-	5:08	4:40	-
6	5:36	5:02	4:40	-
8	6:42	4:47	-	-
9	3:40	-	-	-

2.2. Wavelet Scattering Transform

WST is an improved method based on wavelet transform, suitable for feature extraction of non-stationary signals [15]. And it extracts features by convolution operations, complex modulus, and local averaging on the signal. Let $f(t)$ be the original PPG signal, ψ_λ represents Morlet wavelet function and ϕ_J represents the scale function.

Firstly, convolution of PPG signal and scale function can obtain the 0-th order wavelet scattering coefficients:

$$S_0 f(t) = f(t) \star \phi_J \tag{1}$$

By convolution, low-frequency features in signals can be obtained, but it also leads to the loss of high-frequency information. Then high-frequency information in the signal needs to be retrieved through wavelet modulus transformation [15]. PPG signals is convoluted with the complex wavelet functions ψ_λ to obtain the first-order wavelet modulus coefficients U_1 :

$$U_1 f(t) = \left| f(t) \star \psi_{\lambda_{1,i}} \right|, \qquad i = 1 \ldots n \tag{2}$$

Then, averaging the first-order wavelet modulus coefficients with scaling functions ϕ_J , the first-order wavelet scattering coefficient can be obtained:

$$S_{1,i} f(t) = \left| f(t) \star \psi_{\lambda_{1,i}} \right| \star \phi_J, \qquad i = 1 \ldots n \tag{3}$$

Local averaging results in the loss of high-frequency information, we can repeat the above steps to obtain the second-order wavelet modulus coefficients:

$$U_2 f(t) = \left\| f(t) \star \psi_{\lambda_{1,i}} \right| \star \psi_{\lambda_{2,j}} \Big|, \qquad i = 1...n, j = 1...m \tag{4}$$

And second-order wavelet scattering coefficients:

$$S_{2,i,j} f(t) = \left\| f(t) \star \psi_{\lambda_{1,i}} \right| \star \psi_{\lambda_{2,j}} \Big| \star \phi_J, \qquad i = 1...n, j = 1...m \tag{5}$$

With an increase in the scattering order, the energy of the scattering coefficient rapidly decreases, and the energy of the scattering coefficient after second-orders less than 1% [13]. So in this study, we only used a second-order scattering network, as shown in Figure 1. In our second-order scattering network, the first-order Q1 and second-order Q2 were set to 8 and 1 wavelets per octave, respectively. The invariance scale was set to be equal in length to the signal segment length. A scattering coefficients matrix with size of 166×5 was converted was generated from a PPG fragment with a length of 10s and frequency of 256Hz by WST. 166 was the number of scattering paths, and 5 represented the number of time windows. When the segment length is 5s, a scattering matrix with a size of 121×5 can be obtained.

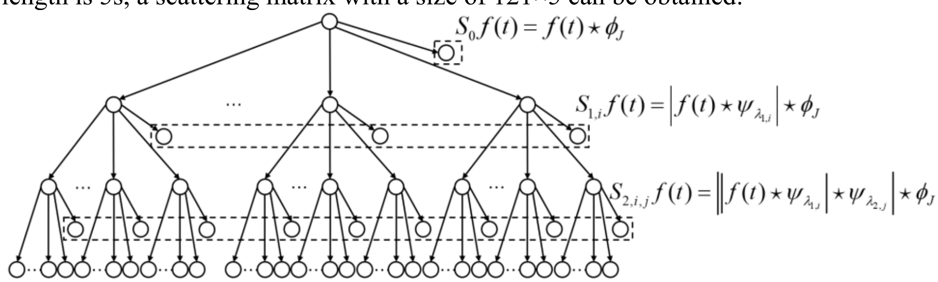

Figure 1. The wavelet scattering network used in this work.

2.3. Experimental Design

In the preprocessing section, we did not perform any filtering or down sampling on the original signal, just segmented the recordings. Due to the length of the original recordings are large, we fragmented the original PPG recordings into 5-second and 10-second PPG samples, respectively. There was no overlap between all samples, and remove the samples with insufficient length. The number of samples were shown in Table 2. Then, WST was used to extract features from PPG samples, generating the feature matrix. The matrixes were transformed into a one-dimensional feature sequence, which were fed into the SVM classifier for classification. The main block diagram of the proposed method is shown in Figure 2.

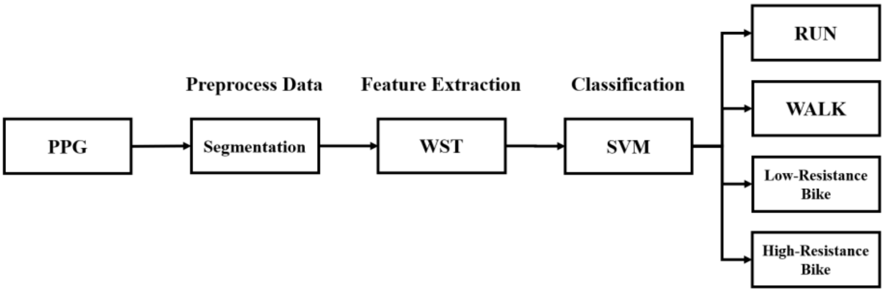

Figure 2. Block diagram representation of the proposed method.

We conducted a four-class classification experiments on four types of exercise recordings to test the feasibility of the proposed method for activity recognition. In addition, we also merged low-resistance bike and high-resistance bike into a bike set for three-class classification experiments.

We divided the sample set into training, validation, and testing sets. In order to test the generalization of the method, we randomly divided 20% of the samples from all samples. These samples not used for model training and validation, and were only used as testing data. Then a 10-fold cross-validation was used for training and testing.

Table 2. Number of samples used in four-class classification.

Exercise Activity	5s		10s	
	Training+Validation	Testing	Training+Validation	Testing
Run	238	59	118	29
Walk	356	88	176	44
High	203	52	101	26
Low	283	70	141	35
Total	1080	269	536	134

In this study, we adopted a multi class SVM based on one-against-one (OAO) and linear kernel to classify PPG samples. Bayesian optimization method and random search were used to optimize the hyperparameters of the classifier.

3. Results and Discussion

The evaluation indexes used in this study are accuracy (*Acc*), recall (*Rec*), precision (*Pre*) and F1-score (*F1*). Their definitions are as follows:

$$Accuracy = \frac{TP_{total} + FN_{total}}{TP_{total} + FP_{total} + TN_{total} + FN_{total}} \tag{6}$$

$$Recall_i = \frac{TP_i}{TP_i + FN_i} \tag{7}$$

$$Precision_i = \frac{TP_i}{TP_i + FP_i} \tag{8}$$

$$F1_i = \frac{2 * Precision_i * Recall_i}{Precision_i + Recall_i} \tag{9}$$

where TP is the number of true positives, it represents that the positive signals are correctly predicted as positive signals; TN is the number of true negatives, indicating that the signals are actually negative and the prediction are also negative; FP is the number of false positives, indicates that the signals are actually negative and incorrectly predicted as positive signals; and FN is the number of false negatives, it is the normal signals that are incorrectly predicted as negative signals.

Table 3. The test results of the proposed approach.

Theme	Activity	5s				10s			
		Acc	*Rec*	*Pre*	*F1*	*Acc*	*Rec*	*Pre*	*F1*
Three-class	Bike	97.03%	100%	100%	1.00	97.76%	100%	100%	1.00
	Run		89.83%	96.36%	0.93		93.10%	96.43%	0.95
	Walk		97.75%	93.55%	0.96		97.73%	95.56%	0.97
	Mean	**97.03%**	**95.86%**	**96.64%**	**0.96**	**97.76%**	**96.94%**	**97.33%**	**0.97**
Four-class	High	91.08%	98.07%	75.00%	0.85	92.54%	84.61%	88.00%	0.86
	Low		75.71%	98.15%	0.85		91.43%	88.89%	0.90
	Run		93.22%	96.49%	0.95		93.10%	96.43%	0.95
	Walk		97.73%	95.56%	0.97		97.73%	95.56%	0.97
	Mean	**91.08%**	**91.18%**	**91.30%**	**0.90**	**92.54%**	**91.72%**	**92.22%**	**0.91**

Table 3 shown the experimental results of the proposed method. The four-class classification experiment on exercise types with 5 seconds of PPG signal samples, can obtained a classification accuracy of 91.08%, a recall of 91.18%, a precision of 91.30%, and an F1 score of 0.90. When using 10s samples, the classification accuracy reached 92.54%, and the recall, precision, and F1 index reached 91.72%, 92.22% and 0.91, respectively. When classifying exercise types into three-class, using samples of 5s and 10s can achieved classification accuracy of 97.03% and 97.76%, respectively.

Through comparison, we can know that the features of low-resistance bike and high-resistance bike are more similar, so prone to misclassification in the four-class classifications, while in the three-class classifications, bike samples were accurately detected and completely distinguished from the other two class. And the 10s PPG sample contains more information than the 5s sample, resulting in more extracted features and higher classification accuracy.

Table 4 shown the results comparison with other studies. These studies used the same database with us and used PPG signals for classification. Through comparison, it can be seen that all evaluation indexes of our method are above 91%. Our approach obtained a accuracy of 92.54%, comparable the results of the study by R.K. Bondugula et al. [16]. But our study has independent test sets, the result is more reliable. Result in this study is superior to the results of other studies, and the classification results are the best.

Table 4. Results comparison with other studies.

Research	Method	Theme	*Acc*	*Rec*	*Pre*	*F1*
Mahmud et al. [8]	Multi-Stage LSTM	Four-class	72.10%	73.10%	71.40%	0.723
Brophy et al. [9]	CNNR	Four-class	90.80%	NR	NR	NR
Almanifi et al. [17]	Continuous wavelet transforms + Ensemble CNN	Four-class	88.91%	NR	NR	NR
Bondugula et al. [16]	Randomized kernels + MLP	Four-class	93.15%	NR	NR	NR
This work	**Wavelet scattering transform + SVM**	**Three-class**	**97.76%**	**96.74%**	**97.33%**	**0.9712**
		Four-class	**92.54%**	**91.72%**	**92.22%**	**0.9195**

NR: Not report

4. Results and Discussion

In this study, a new method of extracting features from PPG using WST to recognize human activity was proposed. Extracting image features from PPG signals by scattering transformation and classified them with the SVM. The experimental results showed that among the four-class classification of exercise types, the classification accuracy, recall, precision and F1 of 91.30%, 91.19%, 91.42%, and 0.9125 can be achieved, respectively. In addition, we also done three-class classification experiments and achieved an accuracy of 98.86%. The results are better than other existing studies, suggest that using WST to extract features from PPG can accurately recognize human activities, and can be widely used in the field of HAR.

Funding

This work was supported by the National Natural Science Foundation of China [grant numbers 61901114], Shandong Province Key Basic Research program [grant numbers ZR2020ZD25], the Development Plan of Youth Innovation Team of University in Shandong Province [grant numbers 2021KJ067] and the Autonomous Innovation Team Foundation for "20 Items of the New University" of Jinan City (No. 202228087).

Acknowledgments

The authors would like to thank the Wrist PPG During Exercise Database for providing the data of PPG.

References

[1] Cero Dinarevic E, Barakovic Husic J, Barakovic S. Step by step towards effective human activity recognition: a balance between energy consumption and latency in health and wellbeing applications. Sensors (Basel). 2019;19(23):5206. Epub 2019/12/01. doi: 10.3390/s19235206. PubMed PMID: 31783705; PubMed Central PMCID: PMCPMC6928889.

[2] Meng Z, Zhang M, Guo C, Fan Q, Zhang H, Gao N, et al. Recent progress in sensing and computing techniques for human activity recognition and motion analysis. Electronics. 2020;9(9):1357. doi: 10.3390/electronics9091357.

[3] Wang Y, Cang S, Yu H. A survey on wearable sensor modality centred human activity recognition in health care. Expert Systems with Applications. 2019;137:167-90. doi: 10.1016/j.eswa.2019.04.057.

[4] Boukhechba M, Cai L, Wu C, Barnes LE. ActiPPG: Using deep neural networks for activity recognition from wrist-worn photoplethysmography (PPG) sensors. Smart Health. 2019;14: 100082. doi: 10.1016/j.smhl.2019.100082.

[5] Alian AA, Shelley KH. Photoplethysmography: analysis of the pulse oximeter waveform. Monitoring Technologies in Acute Care Environments. 2014. p. 165-78.

[6] Shabaan M, Arshid K, Yaqub M, Jinchao F, Zia MS, Bojja GR, et al. Survey: smartphone-based assessment of cardiovascular diseases using ECG and PPG analysis. BMC Med Inform Decis Mak. 2020;20(1):177. Epub 2020/07/31. doi: 10.1186/s12911-020-01199-7. PubMed PMID: 32727453; PubMed Central PMCID: PMCPMC7392662.

[7] Tanc YH, Ozturk M. Hypertension classification using PPG signals. 2022 Medical Technologies Congress (TIPTEKNO)2022. p. 1-4.

[8] Mahmud T, Akash SS, Fattah SA, Zhu WP, Ahmad MO, editors. Human activity recognition from multi-modal wearable sensor data using deep multi-stage LSTM architecture based on temporal feature

aggregation. 2020 IEEE 63rd International Midwest Symposium on Circuits and Systems (MWSCAS); 2020 Aug 9-12.

[9] Brophy E, Muehlhausen W, Smeaton AF, Ward TE. CNNs for Heart rate estimation and human activity recognition in wrist worn sensing applications. 2020 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)2020. p. 1-6.

[10] Keerthiveena B, Esakkirajan S, editors. Denoising of PPG signal by wavelet packet transform. 2017 International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT); 2017 July 6-7.

[11] Aydemir T, Şahin M, Aydemir O. A new method for activity monitoring using photoplethysmography signals recorded by wireless sensor. Journal of Medical and Biological Engineering. 2020;40(6):934-42. doi: 10.1007/s40846-020-00573-0.

[12] Mallat S. Group invariant scattering. Communications on Pure and Applied Mathematics. 2012;65(10):1331-98. doi: doi.org/10.1002/cpa.21413.

[13] Bruna J, Mallat S. Invariant scattering convolution networks. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2013;35(8):1872-86. doi: 10.1109/TPAMI.2012.230.

[14] Jarchi D, Casson A. Description of a database containing wrist PPG signals recorded during physical exercise with both accelerometer and gyroscope measures of motion. Data. 2016;2(1):1. doi: 10.3390/data2010001. (https://physionet.org/content/wrist/1.0.0/)

[15] Andén J, Mallat S. Deep scattering spectrum. IEEE Transactions on Signal Processing. 2014;62(16):4114-28. doi: 10.1109/TSP.2014.2326991.

[16] Bondugula RK, Udgata SK, Sivangi KB. A novel deep learning architecture and MINIROCKET feature extraction method for human activity recognition using ECG, PPG and inertial sensor dataset. Applied Intelligence. 2022;53(11):14400-25. doi: 10.1007/s10489-022-04250-4.

[17] Almanifi ORA, Mohd Khairuddin I, Mohd Razman MA, Musa RM, P.P. Abdul Majeed A. Human activity recognition based on wrist PPG via the ensemble method. ICT Express. 2022;8(4):513-7. doi: 10.1016/j.icte.2022.03.006.