Ministry of Education of the Republic of Belarus
Educational institution
Belarusian State University of
Informatics and Radioelectronics

UDC 004.93'11

WEI
Shishi

RECURRENT NEURAL NETWORK FOR HUMAN ACTIVITY RECOGNITION SYSTEMS USING PHOTOPLETHYSMOGRAPHY AND ACCELEROMETER SMARTPHONE DATA

Abstract
for a Master's Degree
in the Specialty 1-45 80 01 Infocommunication Systems and Networks

____

Supervisor
PhD, Associate professor
Boriskevich Ilya Anatolyevich

____

Minsk 2023

INTRODUCTION

Human activity recognition is a critical area of research that focuses on identifying and categorizing different human activities based on various types of data. In recent times, wearable inertial sensors have emerged as a crucial source of non-visual data for time series analysis in HAR. The growth in wearable devices equipped with these sensors has played a pivotal role in driving significant advancements in HAR research.

In addition to inertial sensors, HAR researchers have been exploring other biosignal sources to enhance recognition performance. One such source is the photoplethysmography sensor, which measures blood flow rate throughout the body. This measurement is governed by the heart's continuous pumping action. While PPG signals were not initially intended for detecting body movements and gestures, our proposed approach aims to extract relevant features from these signals and employ deep learning techniques to predict physical activities. Accurately estimating the PPG signal from a subject's wrist during exercise can be a challenging task due to the severe corruption caused by motion artifacts in the raw PPG signal. MAs are primarily caused by the relative motion between the PPG light source detector and the subject's wrist skin during exercise. Despite these challenges, PPG data has the potential to provide valuable information for HAR systems, particularly when combined with other sensor modalities such as ACC, which have proven to be highly effective. Addressing these challenges through advanced signal processing techniques and algorithm design can lead to more accurate and efficient HAR systems using PPG and ACC data.

The research proposes to accomplish this task using a recurrent neural network, ensuring the required performance in terms of accuracy and low complexity. The purpose of using RNN for human activity detection is that RNN is an artificial neural network that can efficiently process sequence and time series data. In human activity detection, data from PPG and accelerometer sensors need to be processed and analyzed. These data are time-series data that contain information about changes in blood flow and acceleration of body movements. RNN has a memory function that can remember the previous input and process it together with the next input. This property makes RNN very effective in processing sequence and time series data because they can capture temporal relationships and dependencies in the data. In addition, RNN can also adaptively adjust its internal state to adapt to time series data of different lengths and distributions. Therefore, using RNN for human activity recognition can improve the accuracy and robustness of recognition while handling time-series data of different lengths and distributions.

The thesis aims to contribute to the field of human activity recognition by proposing a novel approach that utilizes a recurrent neural network to process and analyze the PPG and accelerometer data. The goal is to develop a highly accurate and reliable machine learning model that can recognize various human activities in real-time, with potential applications in healthcare, sports, and fitness monitoring. The proposed model will be trained and evaluated on a large dataset of PPG and accelerometer measurements collected from diverse populations, and its performance will be compared with other state-of-the-art machine learning algorithms. The thesis will also investigate the effectiveness of different features extraction and selection methods, as well as the impact of various hyperparameters on the model's performance. The findings of the study will provide insights into the potential of using PPG and accelerometer data for human activity recognition, and contribute to the development of more advanced and reliable models for this purpose.

Our proposed model incorporates both PPG and ACC data from a public dataset. PPG sensors measure changes in blood volume in tissue microvascular beds, while accelerometers measure the acceleration of body motion. The proposed model is trained on the PPG dataset and the accelerometer dataset collected from 7 participants performing 3 different activities including stepping, resting, squatting. The model was then evaluated on a separate test set to measure how accurately it recognized different activities.

The innovative approach has the potential to unlock new avenues for HAR research and lead to more accurate, efficient, and versatile activity detection systems. Furthermore, the combination of PPG and inertial sensor data can benefit various applications, such as health monitoring, sports performance analysis, and elderly care, by providing a more comprehensive understanding of human activities and their underlying physiological processes.

GENERAL DESCRIPTION OF WORK

Relevance of the subject

The work corresponds to paragraph 1 «Digital information and communication and interdisciplinary technologies, production based on them» of the State Program of innovative development of the Republic of Belarus for 2021–2025.

The work was carried out in the educational institution Belarusian State University of Informatics and Radioelectronics within the framework of research work 21-2033 "Processing, coding and transmission of information in network-centric systems".

The aim and tasks of the work

The aim of the work is to use recurrent neural network for human activity recognition systems based on photoplethysmography and accelerometer sensor data and estimate RNN model performance in terms of accuracy and low complexity.

To achieve this aim, the following tasks were solved in the dissertation:

1 PPG and ACC Sensor data preprocessing for feature extraction.

2 Construct a neural network model based on RNN.

3 Estimation of the RNN model performance for HAR system.

Personal contribution of the author

The content of the dissertation reflects the personal contribution of the author. It consists in the analytical processing of the dataset, the construction of the neural network layers, the selection of the different layers, and the analytical discussion of the conclusions.

Task setting and discussion of the results were carried out together with the supervisor of Sc., Associate Professor Boriskevich Ilya Anatolyevich.

Testing and implementation of results

The main provisions and results of the dissertation work were reported and discussed at: 58th scientific conference of postgraduates, undergraduates and students, (Minsk, April 18–22, 2022) and International scientific and technical seminar "Technologies of information transmission and processing" (Minsk, March - April 2023).

Author's publications

According to the results of the research presented in the dissertation, 2 author's works was published, including: 2 articles in conference proceedings.

Structure and size of the work

The dissertation work consists of introduction, general description of the work, three chapters with conclusions for each chapter, conclusion, list of author's publications, 1 appendix.

The total amount of the thesis is 78 pages, of which 56 pages of text, 21 figures on 9 pages, 12 tables on 4 pages, a list of used bibliographic sources (42 titles on 4 pages), a list of the author's publications on the subject of the thesis (2 titles on 1 page), 1 appendix on 3 pages, graphic material on 6 pages.

Plagiarism

An examination of the dissertation « Recurrent neural network for human activity recognition systems using photoplethysmography and accelerometer sensor data» by Wei Shishi was carried out for the correctness of the use of borrowed materials using the network resource «Antiplagiat» (access address: https://antiplagiat.ru) in the online mode 13.06.2023. As a result of the verification, the correctness of the use of borrowed materials was established (the originality of the thesis is 92.6 %).

SUMMARY OF WORK

In the first chapter, the introduction of PPG and ACC sensor and the dataset was described in detail, the methods for pre-processing the dataset were presented. In section 1.1, an introduction to the ACC and PPG sensors was given and it is explained that the use of PPG and ACC data for HAR can enhance accuracy and reduce motion artefacts. In section 1.2, research on the use of neural networks for HAR in recent years was presented. In section 1.3, the dataset was described in detail, which consisted of data from seven sets of subjects, five series, each with three types of activity, resting, squatting, stepping, respectively. In section 1.4, the data was pre-processed and the raw data was normalised, after which the following operations were performed on the data: downsampling, data segmentation, and data augmentation operations. Preprocessing of PPG and ACC data was a crucial step in the development of a human activity recognition system. Normalization can help to represent different aspects of the activity being performed in a comparable and consistent way. Sensor data segmentation was also important for organizing the signals into smaller windows of fixed duration. Sensor data downsampling can reduce the amount of data, increase the speed of model training, and reduce the risk of overfitting. Sensor data augmentation can increase the amount of data, improve the robustness and generalization ability of the model. Proper preprocessing of PPG and ACC data can improve the accuracy and reliability of human activity recognition systems, leading to more effective monitoring and analysis of physical activity. The best accuracy was reached 91.79 % with a window of 1200 samples (before downsampling), corresponding to 3 seconds and 50 % overlapping.

In the second chapter, a model for recurrent neural networks was explored, an RNN neural network based on PPG and ACC data for HAR was constructed, and the detailed parameters and structure of each layer in the network framework were provided in detail. In section 2.1, we have discussed the fundamental principles of Recurrent Neural Networks and some essential mathematical formulas that are crucial to understand their functioning. However, it is important to note that RNNs have evolved over the years, and different variants have been developed to address the limitations of the traditional RNNs. These variants include Long Short-Term Memory, Gated Recurrent Unit, and many others. RNN and its variants is essential to comprehend how deep learning models process sequential data. In section 2.2, we have described that three gates to control the flow of information in LSTM, namely input gate, forget gate and output gate. The role of these three gates is to control the input of new information, the retention of old information and the selection of output information. Through the operation of these gates, LSTM can better deal with long-term dependencies in long sequence data, and alleviate the gradient disappearance and gradient explosion problems in traditional RNN. In section 2.3, time feature extraction is presented. Feature extraction is the process of computationally generating feature vectors that characterize the actions of the segmented data pieces. The purpose is to obtain physically and statistically significant features from the original data. In section 2.4, the model based on recurrent neural network is constructed, each layer of RNN model has: Fully Connected layer I, Batch Normalization, LSTM I, Dropout I LSTM II, Dropout II, Fully Connected Layer II, Softmax layer. Each layer of the RNN model is described in detail and its parameters are analyzed. Softmax layer is used for classification.

It was set that Fully Connected layer I included the following description: an input tensor of 676020 rows and 4 columns, a condition of 4 rows and 32 columns weights and 32 biases, and an output tensor of size 1024 rows and 32 columns. LSTM layer I included the following description: an input tensor of 1024 rows and 32 columns, a condition of 32 rows and 128 columns weights and 128 biases, and an output tensor of size 1024 rows and 128 columns. Batch Normalization layer can accelerate neural network training, reduce gradient vanishing problems, and improve model robustness. Dropout layer can prevent neural network overfitting and improve model generalization ability. LSTM layer II included the following description: an input tensor of 1024 rows and 128 columns, a condition of 128 rows and 32 columns weights and 32 biases, and an output tensor of size 1024 rows and 32 columns. Fully Connected Layer II included the following description: an input tensor of 1024 rows and 32 columns, a condition of 32 rows and 3 columns weights and 3 biases, and an output tensor of size 1024 rows and 3 columns. Finally, the activities were classified into three categories through the softmax layer.

In the third chapter, accuracy and loss progress of training epochs and estimation of human physical activity classification accuracy are presented. In section 3.1, we utilized the first five data sets for training and the last two for testing. The research was carried out using the Google Colab platform. To implement the proposed neural network model and other DL–based models, the Python library was used with TensorFlow as the backend. To visualize the data analysis and model evaluation results, Seaborn and Matplotlib were implemented, making it easier to comprehend the models' performance. Scikit–learn was also employed as a valuable resource for producing samples and data during the execution of the experiments. Finally, a high–performing and accurate model was created by utilizing the powerful combination of TensorFlow and Keras for the training and implementation of DL models. Decimation factor is a term used in signal processing to describe the reduction in the number of samples in a signal. It is a crucial concept in the design of digital filters and in the downsampling of signals. In section 3.2, when applying distinct extraction factors to the RNN model for input data, the results varied. After experimenting with various decimation factors and evaluating the final test accuracy, we discovered that a decimation factor of 40 and 100 epochs produced the highest accuracy of 95.36 %. However, increasing the decimation factor caused a larger gap between the training and testing accuracy. To examine this bias, we conducted a cross-validation strategy using one subject at a time, testing seven models in each experiment, with six subjects for training. As reducing the test material increases overfitting, we repeated the test with 50 epochs in addition to the 100 epochs.

CONCLUSION

1 Preprocessing of PPG and ACC data was a crucial step in the development of a human activity recognition system. Normalization can help to represent different aspects of the activity being performed in a comparable and consistent way. Sensor data segmentation was also important for organizing the signals into smaller windows of fixed duration. Sensor data downsampling can reduce the amount of data, increase the speed of model training, and reduce the risk of overfitting. Sensor data augmentation can increase the amount of data, improve the robustness and generalization ability of the model. Proper preprocessing of PPG and ACC data can improve the accuracy and reliability of human activity recognition systems, leading to more effective monitoring and analysis of physical activity. The best accuracy was reached 91.79 % with a window of 1200 samples (before downsampling), corresponding to 3 seconds and 50 % overlapping.

2 The neural network model based on recurrent neural network was constructed for the HAR system, which consists of Fully Connected layer I, Batch Normalization layer, LSTM layer I, Dropout layer I, LSTM layer II, Dropout layer II, Fully Connected Layer II, and Softmax layer. It was set that Fully Connected layer I included the following description: an input tensor of 676020 rows and 4 columns, a condition of 4 rows and 32 columns weights and 32 biases, and an output tensor of size 1024 rows and 32 columns. LSTM layer I included the following description: an input tensor of 1024 rows and 32 columns, a condition of 32 rows

and 128 columns weights and 128 biases, and an output tensor of size 1024 rows and 128 columns. Batch Normalization layer can accelerate neural network training, reduce gradient vanishing problems, and improve model robustness. Dropout layer can prevent neural network overfitting and improve model generalization ability. LSTM layer II included the following description: an input tensor of 1024 rows and 128 columns, a condition of 128 rows and 32 columns weights and 32 biases, and an output tensor of size 1024 rows and 32 columns. Fully Connected Layer II included the following description: an input tensor of 1024 rows and 32 columns, a condition of 32 rows and 3 columns weights and 3 biases, and an output tensor of size 1024 rows and 3 columns. Finally, the activities were classified into three categories through the softmax layer.

3 The first five sets of data are used for the training data set, and the last two sets of data were used for the test set. Different extraction factors were used for the RNN model of the input data, and the obtained results are different. After conducting training and assessing the final test accuracy using various decimation factors, we discovered that the highest test accuracy of 95.36 % was attained when the decimation factor is set to 40 and the epoch is 100. However, increasing the decimation factor led to a greater disparity between the training and testing accuracy. To evaluate the impact of bias, we conducted another test using a one–subject cross–validation strategy. This involved testing seven models in each experiment, with six subjects used for training and one (different each time) for testing. Although a minority of participants may have a negative impact on the average results, the majority of participants exhibited accuracy similar to the previously reported well–performing participants.

LIST OF AUTHOR'S PUBLICATIONS

1–A. S.S Wei. Photoplethysmography and accelerometer sensors signals for recognizing physical activity / S.S Wei // Технологии передачи и обработки информации (Technologies of information transmission and processing): материалы Международного научно-технического семинара, Минск, март – апрель 2023 г. – Minsk: BSUIR, 2023. – С. 123–127.

2–A. S.S Wei. Development of recurrent neural network / S.S Wei // Технологии передачи и обработки информации (Technologies of information transmission and processing): материалы Международного научно-технического семинара, Минск, март – апрель 2023 г. – Minsk: BSUIR, 2023. – С. 154–158.