

A Method for Combining Classical and Deep Machine Learning for Mobile Health and Behavior Monitoring

Martin Gjoreski

Department of Intelligent Systems, Jozef Stefan Institute, Jamova 39, Ljubljana, Slovenia

E-mail: martin.gjoreski@ijs.si

<https://martingjoreski.github.io>

Keywords: machine learning, deep learning, mobile health, behavior monitoring, wearable sensors

Received: March 30, 2021

This paper summarizes the doctoral dissertation of the author, which presents a method for fusing classical and deep machine learning for mobile health and behavior monitoring with wearable sensors.

Povzetek: Prispevek povzame doktorsko disertacijo avtorja, ki temelji na metodi za kombiniranje klasičnega in globokega strojnega učenja za mobilno spremljanje zdravja in obnašanja z nosljivimi senzorji..

1 Introduction

Commercially available smartphones, smart glasses, smartwatches, and smart rings are just a few examples of sensor-packed devices that are enabling the technological revolution currently underway. To further extend the successful applicability of wearable devices in sectors such as mobile health, methods for accurate measurements of psycho-physiological information are required. However, accessing psycho-physiological information using wearable devices remains challenging. One reason is that the relationship between sensor data and human psycho-physiological states is not as unambiguous as the relationship between sensor data and individual physical states is. Thus, we are facing a question: How can we transform wearable sensor data into valuable human health and behavior information? Such information has the potential to improve healthcare, decrease healthcare costs, improve the quality of life and, ultimately, save human lives.

For a decade, deep learning (DL) has dominated the AI world by achieving a breakthrough in several areas such as image processing, natural language processing, and reinforcement learning. Thus, a successful fusion of classical machine learning (ML) and DL methods could lead to beyond state-of-the-art results for mobile health and behavior monitoring.

2 Case studies

The proposed method was applied in seven health and behavior-monitoring domains [1]: stress recognition from physiological sensors, blood pressure estimation from ECG sensors, emotion recognition from physiological sensors and cognitive-load recognition from physiological sensors, chronic heart failure monitoring from heart sounds [2], driver distractions monitoring from physiological and video-based sensors [3], and locomotion recognition from smartphone sensors [4].

3 Method

The proposed method (Figure 1) extracts valuable human health and behavior information from wearable sensor data. The method uses end-to-end learning on sensor data as a standalone approach or in combination with classical ML to produce beyond state-of-the-art performance.

The method uses as input any data collected using wearable sensors from human users. The type of sensors depends on the use-case. Regarding the sensors utilized in the thesis, in the studies on stress and cognitive load monitoring, physiological and acceleration data from a wrist-worn device was used. In the study on emotions, physiological data from wearable sensors was used. In the study on driving distractions, physiological and video-based sensors were used. In the study on locomotion recognition, smartphone sensors were used. In the study on blood pressure estimation, data from chest-worn ECG sensor was used. Finally, in the study on chronic heart failure (CHF) detection, digital stethoscope was used to record heart sounds. Each of these studies had a different hardware setup, while the method is hardware-independent. Regarding the data labels, in the study on CHF, the labels were provided by medical experts. In the rest of the studies, the labels were provided by the users themselves.

The data from the wearable devices is quite often noisy. The usual source of the noise are movement artefacts and sensor misplacement. The filtering strategies include winsorization, detrending, moving average, low-pass, high-pass, band-pass, etc. The sensor data can be transformed into different domains (e.g., time domain and frequency domain), each of them specialized for extracting different types of information from the input data. For example, by combining gyroscope and acceleration data from smartphone data, the acceleration data can be rotated. This produces location-independent acceleration data, which is useful for more robust activity recognition from smartphone sensors. Another simple transformation calculates the acceleration magnitude by combining the sensor from each axis (x-, y- and z-axis).

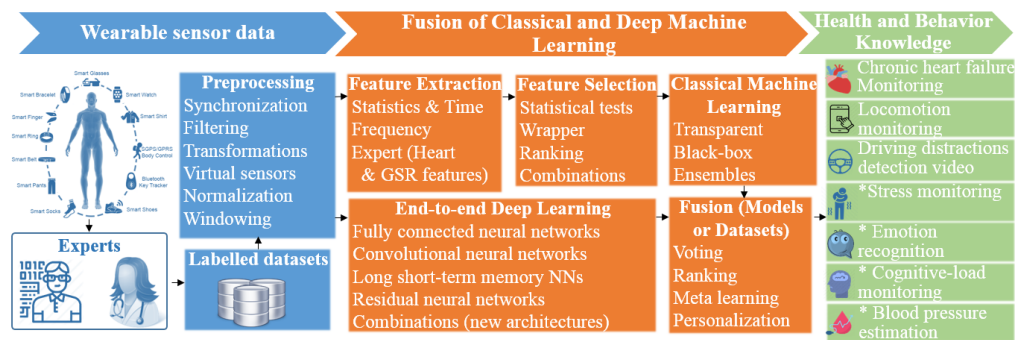


Figure 1: A fusion of classical and deep machine learning for mobile health and behavior monitoring.

Some people have a faster heart rate than other people, some sweat more than others, some walk faster than others, etc. This variability can harm any ML model, especially if there is a small dataset to train the model on. To minimize the variability, different normalization techniques can be employed. The normalization can be done either on the sensor data, or on features. In both cases, the normalization has similar effects, i.e., it scales the values of the variables, and in some cases, it also changes the distribution of the variables.

Informative features should be extracted which are used as input into classical ML algorithms. The feature extraction is an important step as it offers the possibility to encode expert knowledge into the system. In addition to the expert knowledge, for domains where expert knowledge is not well defined, all possible features can be extracted by “borrowing” expert knowledge from similar domains. In the thesis, we experimented with several types of features: statistical features, frequency features, heart-related features and galvanic skin response features.

In the study on monitoring stress, a new feature selection method was proposed by combining ranking and wrapper methods. The method aims to minimize the number of evaluations required by the wrapper method by prioritizing top features ranked by information gain, and by removing low-ranked features which are correlated with the top-ranked features.

The extracted features can be fed into a variety of classical ML algorithms including algorithms that produce comprehensible classifiers (e.g., Decision Trees), black-box classifiers (e.g., SVM) and ensembles (e.g., Extreme Gradient Boosting).

Processed sensor data can be also used to build end-to-end DL models, i.e., models that do not require feature extraction and learn directly from sensor data with the potential to discover new useful patterns in the data, previously unknown to experts. The DL models include existing DL architectures (e.g., Convolutional Neural Networks, Long Short-Term Memory Neural Networks, their combination – ConvLSTMs, Residual Networks – ResNet etc.), and our novel DL architecture Spectro-Temporal Residual Network (STResNet). STResNet is a novel DL architecture for end-to-end learning specialized for multimodal sensor data. More specifically, STResNet can learn from several sensors simultaneously, each of them having a different sampling frequency, and it learns both in the time domain and in the frequency domain.

Regarding the final fusion of the models ML and DL models, depending on the use-case, in some cases, only the highest-ranked model is used to minimize complexity. In other cases, classical meta-learners and voting ensembles are used to maximize accuracy. In third cases, meta-learners that can account for temporal dependencies in the data are used.

4 Conclusion

This paper summarized the dissertation [1] and presented the main idea and findings of the same. The thesis presented: a novel general method that combines expert knowledge, classical machine learning, and deep learning for extracting human physical, physiological, and psychological information from wearable sensor data; a novel deep learning architecture for end-to-end learning (STResNet) specialized for multimodal sensor data; unified application of the method on seven domains; new datasets and publicly available software for mobile health and behavior monitoring with wearable sensors.

References

- [1] M. Gjoreski, A fusion of classical and deep machine learning for mobile health and behavior monitoring with wearable sensors, PhD Thesis, IPS Jožef Stefan, Ljubljana, Slovenia, 2020.
- [2] M. Gjoreski, A. Gradišek, B. Budna, M. Gams, and G. Poglajen, “Machine learning and end-to-end deep learning for the detection of chronic heart failure from heart sounds,” *IEEE Access*, vol. 8, pp. 20313–20324, 2020.
<https://doi.org/10.1109/ACCESS.2020.2968900>
- [3] M. Gjoreski, M. Gams, M. Luštrek, P. Genc, J.-U. Garbas, and T. Hassan, “Machine learning and end-to-end deep learning for monitoring driver distractions from physiological and visual signals,” *IEEE Access*, vol. 8, pp. 70590–70603, 2020.
<https://doi.org/10.1109/ACCESS.2020.2986810>
- [4] M. Gjoreski, V. Janko, G. Slapničar, M. Mlakar, N. Reščič, J. Bizjak, V. Drobnič, M. Marinko, N. Mlakar, M. Luštrek, and M. Gams, “Classical, and deep learning methods for recognizing human activities, and modes of transportation with smartphone sensors,” *Information Fusion*, vol. 62, pp. 47–62, 2020.
<https://doi.org/10.1016/j.inffus.2020.04.004>