НАУЧНО-ТЕХНИЧЕСКИЙ ВЕСТНИК ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, МЕХАНИКИ И ОПТИКИ
ноябрь–декабрь 2023 Том 23 № 6 http://ntv.ifmo.ru/
SCIENTIFIC AND TECHNICAL JOURNAL OF INFORMATION TECHNOLOGIES, MECHANICS AND OPTICS
November–December 2023 Vol. 23 No 6 http://ntv.ifmo.ru/en/
ISSN 2226-1494 (print) ISSN 2500-0373 (online)



doi: 10.17586/2226-1494-2023-23-6-1178-1186

# Personalization of convolutional neural networks within the stress detection task using heart rate variability data

**Maksim O. Dobrokhvalov[1]✉, Anton Yu. Filatov[2]**

[1,2] Saint Petersburg Electrotechnical University “LETI”, Saint Petersburg, 197022, Russian Federation

[1] night1337bot@gmail.com✉, https://orcid.org/0000-0002-0571-5836
[2] aifilatov@etu.ru, https://orcid.org/0000-0003-4298-8523

**Abstract**
Stress detection is an active area of research with important implications for personal, occupational, and social health. Most modern approaches use features computed from multiple sensor modalities, i.e., grouping different types of data from multiple sources for processing. These include electrocardiogram, electrodermal activity, electromyogram, skin temperature, respiration, accelerometer data, etc. Also, traditional machine learning algorithms (decision tree, discriminant analysis, support vector machine, etc.) or fully-connected neural networks are mostly used. Using these methods requires large amounts of data. Researchers are considering different approaches to personalization or generalization of models relative to subjects, namely subject-independent and subject-dependent (initially personal or adapted) models. The aim of the presented work is to develop a method for detecting stress based on heart rate variability data, taking into account the process of personalization of neural networks. The use of a convolutional neural network is proposed. The dependence of accuracy on the length of the input signal is studied. The dependence of accuracy on the data dimensionality reduction layer (one-dimensional convolutional layer, maximizing and averaging pooling) used in the network is also considered. The importance of personalizing models is demonstrated to significantly increase the accuracy of models of specific subjects. It is shown that the proposed method, based on 60 intervals between heartbeats, makes it possible to binary determine whether a person is under stress. Personalization allowed increasing the accuracy from 91.8 % to 98.9 ± 2.6 %. The F1-score value increased from 0.907 to 0.983 ± 0.038. The proposed personalized networks can be used in systems for monitoring the functional state of a person. They can also be used as part of a system that grants or restricts access to private resources based on whether a person is currently at rest.

**Keywords**
stress detection, convolutional neural network, machine learning, heart rate variability, subject-dependent models

**Acknowledgements**
The article was prepared within the project “Methods of hybrid intelligence for building heterogeneous multi-agent systems with self-learning and self-organization” of the development program of St. Petersburg Electrotechnical University “LETI”.

**For citation:** Dobrokhvalov M.O., Filatov A.Yu. Personalization of convolutional neural networks within the stress detection task using heart rate variability data. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2023, vol. 23, no. 6, pp. 1178–1186. doi: 10.17586/2226-1494-2023-23-6-1178-1186

УДК 004.032.26

# Персонализация сверточных нейронных сетей в задаче обнаружения стресса с использованием данных вариабельности сердечного ритма

**Максим Олегович Доброхвалов[1]✉, Антон Юрьевич Филатов[2]**

[1,2] Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» имени В.И. Ульянова (Ленина), Санкт-Петербург, 197022, Российская Федерация

[1] night1337bot@gmail.com✉, https://orcid.org/0000-0002-0571-5836

[2] aifilatov@etu.ru, https://orcid.org/0000-0003-4298-8523

**Аннотация**

**Введение.** Обнаружение стресса является активной областью исследований с важными последствиями для личного, профессионального и социального здоровья человека. Большинство современных подходов используют признаки, вычисленные на основе нескольких сенсорных модальностей, т. е. группируют для обработки различные типы данных, полученные из нескольких источников. К ним относятся электрокардиограмма, кожно-гальваническая реакция, электромиограмма, температура кожи, дыхание, данные акселерометров и др. При этом чаще используются традиционные алгоритмы машинного обучения, такие как решающие деревья, дискриминантный анализ, метод опорных векторов и другие, а также полносвязные нейронные сети. Использование этих методов требует больших объемов данных. Исследователи рассматривают отличающиеся подходы к персонализации или общности моделей относительно субъектов, а именно субъекто-независимые и субъекто-зависимые (изначально персональные или адаптированные) модели. Целью представленной работы является разработка метода детектирования стресса на основе данных вариабельности сердечного ритма с учетом процесса персонализации нейронных сетей. **Метод.** Для решения поставленной задачи предложено применение сверточной нейронной сети. Исследована зависимость точности детектирования от длины входного сигнала. Рассмотрена зависимость точности от используемого в сети слоя уменьшения размерности данных (одномерный сверточный слой, максимизирующий и усредняющий пуллинги). Продемонстрирована важность персонализации моделей, для значительного увеличения точности детектирования для конкретных субъектов. **Основные результаты.** Показано, что предлагаемый метод на основании 60 интервалов между ударами сердца позволяет бинарно определить, находится ли человек в состоянии стресса. Персонализация сверточных нейронных сетей позволила повысить точность с 91,8 до 98,9 ± 2,6 %. Значение F1-меры повысилось с 0,907 до 0,983 ± 0,038. **Обсуждение.** Предложенные персонализированные сети могут применяться в системах мониторинга функционального состояния человека. Также могут быть использованы как часть системы, предоставляющей или ограничивающей доступ к приватным ресурсам на основании того, находится ли человек в состоянии покоя в данный момент.

**Ключевые слова**

детектирование стресса, сверточные нейронные сети, машинное обучение, вариабельность сердечного ритма, субъекто-зависимые модели

**Благодарности**

Работа подготовлена в рамках проекта «Методы гибридного интеллекта для построения гетерогенных многоагентных систем с самообучением и самоорганизацией» программы развития СПбГЭТУ «ЛЭТИ».

**Ссылка для цитирования:** Доброхвалов М.О., Филатов А.Ю. Персонализация сверточных нейронных сетей в задаче обнаружения стресса с использованием данных вариабельности сердечного ритма // Научно-технический вестник информационных технологий, механики и оптики. 2023. Т. 23, № 6. С. 1178–1186 (на англ. яз.). doi: 10.17586/2226-1494-2023-23-6-1178-1186

## Introduction

Stress is the body's response to perceived physical or psychological threats [1] and it is defined as the transition from a calm state to an excited state, triggering a set of physiological responses [2]. Moreover, stress detection is important for many health problems, such as depression, anxiety, heart attacks and strokes [3]. Stress also affects a person's decision-making ability, attention span, learning and problem-solving ability [4]. Therefore, stress detection is an important task.

Various classical machine learning methods as well as neural networks are used in various studies to solve this problem. Also, various input data for stress detection are used from various data sensors, such as electrocardiogram (ECG), electrodermal activity (EDA), etc. This study proposes to use a convolutional neural network which receives a set of RR intervals (Heart Rate Variability (HRV)) as input data. In [5], the authors successfully applied personalization to EDA data. Based on that research, this paper examines the process of personalization of convolutional neural networks with HRV input data. Thus, the aim of the work is to develop a method for stress detection based on HRV data, taking into account the process of personalization of neural networks, as well as the implementation of this method. The proposed approach is competitive with other modern methods. The code used in the work is available[1].

## Related works

Summary of review related works is presented in Table 1. Wearable Stress and Affect Detection (WESAD) is a commonly used dataset in related works (14 of 20) but also some studies use their own data. Among the works reviewed, ECG [2, 3, 6–13], EDA [2–6, 9, 11, 14–19] and

[1] Available at: https://github.com/Nightbot1448/human_stress_detection (accessed: 10.01.2023).

*Table 1.* Summary of reviewed related works

| Paper | Year | Dataset | Subjects | Data | Model | Accuracy, % | Window, s |
|---|---|---|---|---|---|---|---|
| [2] | 2018 | WESAD | 15 | ECG, EDA, BVP, Temp, Resp, EMG, ACC | kNN, DT, RF, **LDA**, AB | 92.83 | 0.25, 5, 60 |
| [3] | 2020 | WESAD | 15 | ECG, EDA, BVP, Temp, Resp, EMG, ACC | kNN, SVM, AB, **FCN** | 95.21 | 1 |
| [4] | 2020 | WESAD | 15 | EDA | **kNN**, SVM, RF | 91.6 | — |
| [5] | 2021 | WESAD | 15 | EDA | CNN | 92.85 | 60 |
| [6] | 2022 | WESAD | 15 | ECG, EDA, BVP, Temp, Resp, EMG, ACC | CNN using GAF | 94.8 | — |
| [7] | 2016 | Other | 42 | ECG | C4.5 tree | 79 | 180 |
| [8] | 2021 | Other | 20 | ECG | CNN | 83.5 | 10 |
| [9] | 2021 | WESAD | 15 | ECG, EDA, BVP, Temp, Resp, EMG, ACC | CNN | 97.75 ± 2.55 | 60 |
| [9] | 2021 | WESAD | 15 | ECG | CNN | 91.75 ± 9.73 | 60 |
| [10] | 2019 | AffectiveROAD, Other | 9, 17 | ECG | FCN | 90.19 | 10, 60 |
| [11] | 2021 | WESAD | 15 | ECG, EDA, BVP, Temp, Resp, EMG, ACC | LR | 85.71 | 60 |
| [12] | 2021 | Other | 27 | ECG | kNN, **SVM**, FCN, RF, GB | 83 | 30 |
| [13] | 2019 | Other | 20 | ECG | CNN | 82.7 | 10 |
| [14] | 2020 | Other | 20, 3 | HR, EDA | CNN | 82.5, 93.8 | — |
| [15] | 2021 | WESAD | 15 | EDA | sTree | 95.8 | 4 |
| [16] | 2018 | Other | 58 | HR, EDA, Resp | FCN | 89.7 | 90 |
| [17] | 2020 | WESAD | 15 | EDA, BVP, ACC, Temp | **RF**, DT, LR | 96.68 ± 3.2 | 0.25 |
| [18] | 2020 | Other | 41 | BVP | **kNN**, LDA, FCN | 82 | 60 |
| [19] | 2021 | WESAD | 15 | EDA | kNN, **SVM**, FCN, RF | 87.5 | 60 |
| [20] | 2022 | WESAD | 15 | BVP | FCN | 99.04 | 300 |
| [21] | 2019 | WESAD | 15 | Temp, BVP, HR | **LDA**, QDA, RF | 87.4 ± 10.4 | 15, 30, 60, 90, 120 |

Blood Volume Pulse (BVP) [2, 3, 6, 9, 11, 17, 18, 20, 21] were most often used as data sources. Also other sources, such as respiration info (Resp) [2, 3, 6, 9, 11, 16], skin temperature (Temp) [2, 3, 6, 9, 11, 17, 21], electromyogram (EMG) [2, 3, 6, 9, 11], accelerometers info (ACC) [2, 3, 6, 9, 11] used in some research. It should be noted that in most cases many data sources are used when applying feature engineering [2, 3, 6, 18, 21]. Time and frequency domain of ECG, BVP and EDA are widely used in studies [2–4, 6, 7, 9, 10, 15, 16]. Some studies [2, 4, 5, 7, 8, 10, 12, 18–20] extract features from only one data source. And also there are few studies that use raw data (sometimes with applying filters but without feature extraction) [5, 9, 13, 17].

Accuracy metrics reported ranged between 79 % [7] and 99.04 % [20]. Half of the studies used neural networks. Convolutional Neural Networks (CNN) were used in 6 papers [5, 6, 8, 9, 13, 14], Fully Connected Networks (FCN) also were used in 7 studies [3, 10, 12, 16, 18–20]. Also different studies used machine learning methods.

Random Forest (RF) was used in [2, 4, 17, 21]. Support Vector Machines (SVM) were utilized in experiments [3, 4, 12, 19]. Linear Discriminant analysis (LDA) was used in [2, 18, 21]. AdaBoost classifier (AB) was utilized in [2, 3]. K-nearest neighbor's classifier (kNN) was used in [2–4, 12, 18, 19]. Tree-based classifiers (like Decision Tree (DT)) were utilized in [2, 3, 7 (C4.5 tree [22]), 15, 17]. Also Logistic Regression (LR) was used in [11, 17]. Indikawati and Winiarti [17] are the only ones who directly used the signal without feature extractions with the classical machine learning methods. Work [23] used convolutional and long short-term memory [24] neural networks for encoding signal with sequent passing to clustering algorithms.

### Materials and Methods

**Data and preprocessing.** Many studies conducted in the field of stress detection use data collected by researchers independently. This study uses the WESAD dataset [2], which has also been used in many studies in recent years [3, 5, 8, 14, 17, 20, 21]. It is a public dataset containing ECG. RR intervals were calculated from the ECG using the heartpy python library[1]. Data with stress

[1] Available at: https://python-heart-rate-analysis-toolkit.readthedocs.io/en/latest/ (accessed: 10.01.2023).

and resting state labels were taken from the dataset. The amusement state was omitted. Next, the RR interval is the interval between neighboring heart beats. The interval is the set of RR intervals used as input data.

**Model.** The convolutional neural network [14] has shown sufficiently high accuracy. Therefore, a 1D convolutional network architecture was chosen. The network architecture is a sequential use of the ConvX block (Fig. 1, *a*) and the dimension reduction layer. The ConvX block consists of a one-dimensional convolutional layer with kernel size 3, a batch normalization layer, and a ReLU activation layer. The network architecture for the interval length (input data) equal to 60 is shown in Fig. 1, *b*. One-dimensional convolution (kernel = 2, stride = 2), max pooling (kernel = 2) and averaging pooling (kernel = 2) were considered as dimensionality reduction blocks. The number of input layers for convolution layers or ConvX blocks is given in parentheses. The first ConvX block parameter (**in**) means that there may or may not be a layer in the input data containing the difference between consecutive RR intervals (numerical derivative). The architecture depends in part on the maximum interval length. The goal was to form an architecture where after each ConvX block a dimensionality reduction layer could be added (except the first and last). Thus, this architecture made it possible to obtain the required data dimension due to convolutional layers and dimensionality reduction layers (without using fully connected layers). The results of various modifications are presented in the following sections.

## Results

This section presents a comparison of different modifications:
— using different interval (input data) lengths,
— choosing layers to reduce dimensionality,
— using numerical derivation.

This section also presents the impact of model personalization for subjects.

**Modifications.** All modifications of the convolutional neural network proposed in the research process were implemented within this study using the PyTorch framework[1]. In all experiments, the CrossEntropy loss function was used, the ASGD optimizer (with default parameters) was used, the number of epochs was 50, and the batch size was 8.

The first study was the choice of interval size. Normal resting heart rates range from 60 to 100 bpm [28]. Therefore, the number of RR intervals is not equivalent to the number of seconds, but they can be mapped. 60 seconds is a widely used interval in review studies. It was decided to consider no more than 60 RR intervals with a step of 15. More than 60 RR intervals were not considered due to too long initialization. The slide of intervals was 5 RR intervals. Fig. 2 shows accuracy for different modifications of models depending on input interval length. It can be seen that for the lengths 15, 30, 45 there is a direct dependence of the accuracy. In the case of input interval length equal to

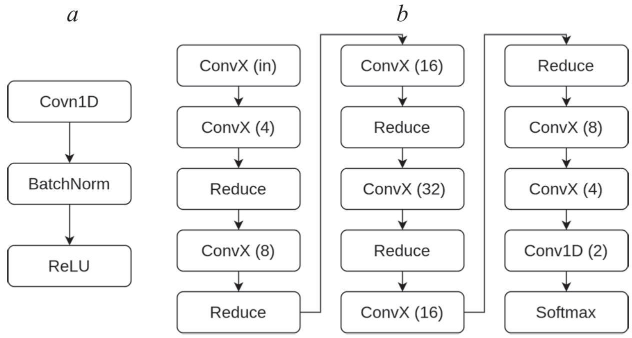


*Fig. 1*. Neural Network: ConvX block (*a*); architecture (*b*)

60, the accuracy for models without numerical derivative is increased. For models with numerical derivation, for the same input interval length, the accuracy decreases regardless of the method of dimensionality reduction. However, the accuracy for the interval length 60 with a max pooling layer is greater (92.16 %) than the accuracy of the other modifications.

As mentioned earlier, one-dimensional convolution layers, max and averaging poolings were considered as dimension reduction methods. Using convolution as a layer for dimension reduction shows the lowest accuracy (Fig. 2). If a numerical derivative was present in the input data, the network with the averaging pooling determined stress with higher accuracy in all cases except when the interval length was 15. If the numerical derivative was not used, then modifications with averaging and max pooling showed greater accuracy depending on the interval length. However the modification with the max pooling showed the highest accuracy (92.16 %) with interval length equal to 60.

The inclusion of an additional layer containing the difference of two consecutive RR intervals to the input data was considered. This value can be treated as a numerical derivative. This difference shows the dynamics of changes in RR intervals, which can be perceived as the rate of change in heart rate. For intervals of length 15, 30, and 45, the accuracy of the networks, whose input was additionally fed by the derivative, is higher than for the corresponding one but without this addition.

In the case of an interval length of 60, the accuracy of modifications without the numerical derivative is slightly higher than that with it. The largest difference between modifications with and without the numerical derivative

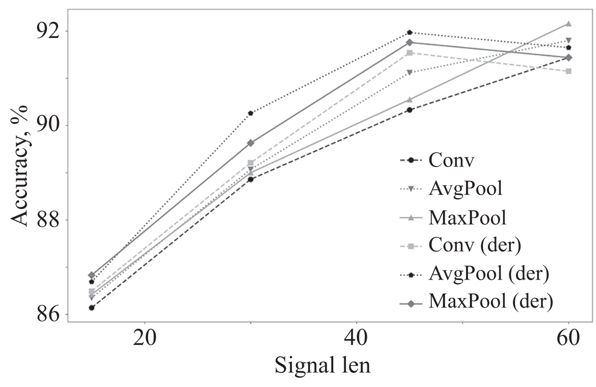


*Fig. 2*. Modifications accuracy

[1] Available at: https://pytorch.org/ (accessed: 13.01.2023).

*Table 2.* The metrics values of the modifications. The first three columns describe the modification type. The following ones are metrics values for this modification

| Signal length | Numerical derivative | Reduce type | Accuracy, % | Balanced accuracy, % | Precision | Recall | F1-score | ROC AUC |
|---|---|---|---|---|---|---|---|---|
| 45 | – | avg | 91.1 | 91.2 | 0.881 | 0.917 | 0.898 | 0.912 |
| 45 | – | conv | 90.3 | 90.2 | 0.883 | 0.892 | 0.888 | 0.902 |
| 45 | – | max | 90.5 | 90.0 | 0.915 | 0.859 | 0.886 | 0.900 |
| 45 | + | avg | 92.0 | 91.7 | 0.915 | 0.896 | 0.905 | 0.917 |
| 45 | + | conv | 91.5 | 91.0 | 0.925 | 0.874 | 0.899 | 0.910 |
| 45 | + | max | 91.8 | 91.4 | 0.919 | 0.886 | 0.902 | 0.914 |
| 60 | – | avg | 91.8 | 91.2 | 0.932 | 0.872 | 0.901 | 0.912 |
| 60 | – | conv | 91.4 | 90.8 | 0.93 | 0.866 | 0.897 | 0.908 |
| 60 | – | max | 92.2 | 91.8 | 0.922 | 0.893 | 0.907 | 0.918 |
| 60 | + | avg | 91.7 | 91.2 | 0.924 | 0.878 | 0.900 | 0.912 |
| 60 | + | conv | 91.2 | 90.8 | 0.908 | 0.883 | 0.895 | 0.908 |
| 60 | + | max | 91.4 | 91.2 | 0.905 | 0.894 | 0.900 | 0.912 |

is 1.2 % (modifications with the max pooling or with the convolutional layer, interval length 45).

**Comparison of proposed modifications.** This section presents a comparison of the metrics of the various proposed modifications (Table 2) when tested using all data (without skipping subjects).

The modification with interval length of 60, max pooling, and without numerical derivative has the highest accuracy, balanced accuracy, F1-score and ROC AUC (Area Under Receiver Operating Characteristic Curve) score. The modification with interval length 60, averaging pooling, and without numerical derivative has the highest precision score. The modification with interval length of 45, without the numerical derivative, and with the averaging pooling has the highest recall. Almost all modifications with signal length 60 differ from the best one by no more than 0.01 on F1-score. Thus, all modifications with signal length 60 were considered for personalization. For all modifications, the maximum accuracy was achieved after the 30th epoch. But 95 % of maximum accuracy had been achieved in the first 10 epochs because accuracy of some subjects reached near 100 %. And in the process of further training the accuracy of the rest of the subjects increased.

**Models personalization.** As stated earlier, each type of these models may have advantages and disadvantages. Subject-dependent models require a large amount of data.

*Table 3.* The accuracy of each subject's personalization, %. The first value in column header is the dimensionality reduction method used. The second is the use (Derivative) or omission (Default) of the numerical derivative

| Subject | Avg, Derivative | Max, Derivative | Conv, Derivative | Avg, Default | Max, Default | Conv, Default |
|---|---|---|---|---|---|---|
| 2 | 95.4 | 95.4 | 100.0 | 95.4 | 98.5 | 100.0 |
| 3 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 |
| 4 | 100.0 | 98.6 | 100.0 | 100.0 | 100.0 | 100.0 |
| 5 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 |
| 6 | 100.0 | 95.2 | 100.0 | 98.8 | 98.8 | 100.0 |
| 7 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 98.8 |
| 8 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 |
| 9 | 90.9 | 89.8 | 76.1 | 84.1 | 83.0 | 85.2 |
| 10 | 96.7 | 92.6 | 96.7 | 98.4 | 94.2 | 100.0 |
| 11 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 |
| 13 | 100.0 | 99.1 | 100.0 | 100.0 | 100.0 | 100.0 |
| 14 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 |
| 15 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 98.0 |
| 16 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 |
| 17 | 100.0 | 100.0 | 99.0 | 100.0 | 98.0 | 100.0 |
| mean | 98.87 | 98.05 | 98.12 | 98.45 | 98.16 | 98.8 |
| std | 2.61 | 3.27 | 6.14 | 4.16 | 4.48 | 3.8 |

*Table 4.* The accuracy of the best model before and after personalization. Other metrics of best model after personalization

| Subject | Accuracy before, % | Accuracy after, % | Accuracy delta, % | Balanced accuracy, % | Precision | Recall | F1-score | ROC AUC |
|---|---|---|---|---|---|---|---|---|
| 2 | 68.2 | 95.4 | 27.3 | 93.2 | 1.0 | 0.864 | 0.927 | 0.932 |
| 3 | 74.3 | 100.0 | 25.7 | 100 | 1.0 | 1.0 | 1.0 | 1.0 |
| 4 | 100.0 | 100.0 | 0.0 | 100 | 1.0 | 1.0 | 1.0 | 1.0 |
| 5 | 56.6 | 100.0 | 43.4 | 100 | 1.0 | 1.0 | 1.0 | 1.0 |
| 6 | 100.0 | 100.0 | 0.0 | 100 | 1.0 | 1.0 | 1.0 | 1.0 |
| 7 | 60.7 | 100.0 | 39.3 | 100 | 1.0 | 1.0 | 1.0 | 1.0 |
| 8 | 90.0 | 100.0 | 10.0 | 100 | 1.0 | 1.0 | 1.0 | 1.0 |
| 9 | 54.6 | 90.9 | 36.4 | 88.6 | 0.926 | 0.806 | 0.862 | 0.886 |
| 10 | 43.0 | 96.7 | 53.7 | 95.8 | 1.0 | 0.917 | 0.957 | 0.958 |
| 11 | 97.3 | 100.0 | 2.7 | 100 | 1.0 | 1.0 | 1.0 | 1.0 |
| 13 | 100.0 | 100.0 | 0.0 | 100 | 1.0 | 1.0 | 1.0 | 1.0 |
| 14 | 97.2 | 100.0 | 2.7 | 100 | 1.0 | 1.0 | 1.0 | 1.0 |
| 15 | 91.8 | 100.0 | 8.2 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 |
| 16 | 100.0 | 100.0 | 0.0 | 100 | 1.0 | 1.0 | 1.0 | 1.0 |
| 17 | 76.0 | 100.0 | 24.0 | 100 | 1.0 | 1.0 | 1.0 | 1.0 |

General models are not taking into account the uniqueness of the subjects. So personalization may be a good solution. Another solution may be subjects grouping by similar patterns of intervals or, in a simpler version, with similar biological traits — gender, age, ethnicity, etc.

As stated previously, a personalization process was performed for all modifications with an interval length of 60. Table 3 presents accuracy after personalization for the modifications. The leave-one-subject-out (LOSO) approach [29] was used for personalization. For each subject, the following actions were performed:

— exclusion of the subject's training data from the total training dataset;
— CNN training;
— testing with the test data of the excluded subject;
— personalization of NN on the subject's training data;
— testing on subject test data.

In the process of personalization, the weights were adjusted not only for the predictor (the last convolutional layer in the network), but also for all other layers. Loss function, optimizer, count of epochs and size of batch was same with model modifications experiments (subsection

*Table 5.* Accuracy of the proposed convolutional neural network and analogs on the WESAD dataset

| Paper | Method | Data | Accuracy, % |
|---|---|---|---|
| [2] | LDA | All Chest | 92.83 |
| [2] | LDA | ECG | 85.44 |
| [3] | MLP | All | 95.21 |
| [4] | kNN | EDA | 91.60 |
| [5] | CNN | EDA | 90.00 |
| [6] | CNN using GAF | All | 94.80 |
| [9] | CNN | All | 97.75 ± 2.55 |
| [9] | CNN | ECG | 91.75 ± 9.73 |
| [11] | Logistic regression-based classifier | HR | 76.38 |
| [15] | sTree | EDA | 95.80 |
| [17] (Subjects only) | RF | Wrist | 96.68 ± 3.2 |
| [19] | SVM | EDA | 87.50 |
| [20] | MLP | BVP | 99.04 |
| [21] | LDA | Skin Temp, BVP, HR | 87.4 ± 10.4 |
| Ours | CNN | ECG | 91.80 |
| Ours (personalized) | CNN | ECG | 98.87 ± 2.61 |

"Modifications"). However, the highest accuracy for most subjects was obtained within 10 epochs. Sizes of subject datasets (amount of intervals) are in range [264, 487] (mean: 373.13, std: 63.20) for training sets and in range [66, 121] (mean: 92.67, std: 15.80) for test sets.

Based on Table 3, it can be concluded that the personalized models of all modifications, on average, give approximately the same result. The modification with numerical derivative and averaging pooling shows the highest accuracy (98.87 %) averaged over users. The difference in accuracy before and after model personalization is presented in Table 4. The table also presents other metrics values of the models after personalization.

Table 5 compares the accuracy of the proposed model with analogues. Based on the table, it can be concluded that the proposed network is competitive with analogs. It can be seen that the accuracy in [20] is higher. However, in that paper, various additional features were calculated as preprocessing, which is additional resource consumption. The proposed method avoids this action.

## Conclusion

This research paper proposes a convolutional neural network for human mental stress detection. The input data were sets of consecutive RR intervals of different lengths. These intervals were calculated from ECG data of the WESAD dataset. Different modifications of convolutional neural networks for the task of stress detection were also considered. The following modifications were considered: input signal length, layer used for dimension reduction (convolutional layer, averaging pooling and max pooling were considered), and use of numerical derivative calculated as difference between consequence RR intervals. This paper proposes initial training of the network on common data, followed by personalization of the model for each individual subject. The best model, which showed the highest average accuracy after personalization, is the network using averaging pooling, numerical derivation, and with an input length of 60. The accuracy of the model at the total given is as 91.7 %. The accuracy obtained after personalization is 98.87 % ± 2.61 %. It should be noted that when learning using LOSO, personalization made it possible to significantly increase the accuracy of the model for some subjects. This is because people, both at rest and under stress, can have different HRV values.

As further plans, the training of the proposed models in the problem of determining 3 classes is considered. Validation using the SWELL dataset is also considered. The use of other input data, such as EDA in combination with HRV, is considered as an area of research.

### References


1. Selye H. *The Stress of Life*. 1956. New York, McGraw-Hill, 324 p.
2. Schmidt P., Reiss A., Duerichen R., Marberger C., Van Laerhoven K. Introducing WESAD, a multimodal dataset for wearable stress and affect detection. *Proc. of the 20th ACM International Conference on Multimodal Interaction*, 2018, pp. 400–408. https://doi.org/10.1145/3242969.3242985
3. Bobade P., Vani M. Stress detection with machine learning and deep learning using multimodal physiological data. *Proc. of the Second International Conference on Inventive Research in Computing Applications (ICIRCA)*, 2020, pp. 51–57. https://doi.org/10.1109/icirca48905.2020.9183244
4. Aqajari S.A.H., Naeini E.K., Mehrabadi M.A., Labbaf S., Rahmani A.M., Dutt N. GSR analysis for stress: Development and validation of an open source tool for noisy naturalistic GSR data. *arXiv*, 2020, arXiv:2005.01834. https://doi.org/10.48550/arXiv.2005.01834
5. Sah R.K., Ghasemzadeh H. Stress classification and personalization: Getting the most out of the least. *arXiv*, 2021, arXiv:2107.05666. https://doi.org/10.48550/arXiv.2107.05666
6. Ghosh S., Kim S., Ijaz M.F., Singh P.K., Mahmud M. Classification of mental stress from wearable physiological sensors using image-encoding-based deep neural network. *Biosensors*, 2022, vol. 12, no. 12, pp. 1153. https://doi.org/10.3390/bios12121153
7. Castaldo R., Xu W., Melillo P., Pecchia L., Santamaria L., James C. Detection of mental stress due to oral academic examination via ultra-short-term HRV analysis. *Proc. of the 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2016, pp. 3805–3808. https://doi.org/10.1109/embc.2016.7591557
8. Dalal S., Khalaf O.I. Prediction of occupation stress by implementing convolutional neural network techniques. *Journal of Cases on Information Technology (JCIT)*, 2021, vol. 23, no. 3, pp. 27–42. https://doi.org/10.4018/jcit.20210701.oa3
9. Lai K., Yanushkevich S.N., Shmerko V.P. Intelligent stress monitoring assistant for first responders. *IEEE Access*, 2021, vol. 9, pp. 25314–25329. https://doi.org/10.1109/access.2021.3057578
10. Cho H.M., Park H., Dong S.Y., Youn I. Ambulatory and laboratory stress detection based on raw electrocardiogram signals using a


### Литература


1. Selye H. The Stress of Life. 1956. New York: McGraw-Hill, 324 p.
2. Schmidt P., Reiss A., Duerichen R., Marberger C., Van Laerhoven K. Introducing WESAD, a multimodal dataset for wearable stress and affect detection // Proc. of the 20th ACM International Conference on Multimodal Interaction. 2018. P. 400–408. https://doi.org/10.1145/3242969.3242985
3. Bobade P., Vani M. Stress detection with machine learning and deep learning using multimodal physiological data // Proc. of the Second International Conference on Inventive Research in Computing Applications (ICIRCA). 2020. P. 51–57. https://doi.org/10.1109/icirca48905.2020.9183244
4. Aqajari S.A.H., Naeini E.K., Mehrabadi M.A., Labbaf S., Rahmani A.M., Dutt N. GSR analysis for stress: Development and validation of an open source tool for noisy naturalistic GSR data // arXiv. 2020. arXiv:2005.01834. https://doi.org/10.48550/arXiv.2005.01834
5. Sah R.K., Ghasemzadeh H. Stress classification and personalization: Getting the most out of the least // arXiv. 2021. arXiv:2107.05666. https://doi.org/10.48550/arXiv.2107.05666
6. Ghosh S., Kim S., Ijaz M.F., Singh P.K., Mahmud M. Classification of mental stress from wearable physiological sensors using image-encoding-based deep neural network // Biosensors. 2022. V. 12. N 12. P. 1153. https://doi.org/10.3390/bios12121153
7. Castaldo R., Xu W., Melillo P., Pecchia L., Santamaria L., James C. Detection of mental stress due to oral academic examination via ultra-short-term HRV analysis // Proc. of the 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). 2016. P. 3805–3808. https://doi.org/10.1109/embc.2016.7591557
8. Dalal S., Khalaf O.I. Prediction of occupation stress by implementing convolutional neural network techniques // Journal of Cases on Information Technology (JCIT). 2021. V. 23. N 3. P. 27–42. https://doi.org/10.4018/jcit.20210701.oa3
9. Lai K., Yanushkevich S.N., Shmerko V.P. Intelligent stress monitoring assistant for first responders // IEEE Access. 2021. V. 9. P. 25314–25329. https://doi.org/10.1109/access.2021.3057578
10. Cho H.M., Park H., Dong S.Y., Youn I. Ambulatory and laboratory stress detection based on raw electrocardiogram signals using a

convolutional neural network. *Sensors*, 2019, vol. 19, no. 20, pp. 4408. https://doi.org/10.3390/s19204408
11. Iqbal T., Redon-Lurbe P., Simpkin A.J., Elahi A., Ganly S., Wijns W., Shahzad A. A sensitivity analysis of biophysiological responses of stress for wearable sensors in connected health. *IEEE Access*, 2021, vol. 9, pp. 93567–93579. https://doi.org/10.1109/access.2021.3082423
12. Dalmeida K.M., Masala G.L. HRV features as viable physiological markers for stress detection using wearable devices. *Sensors*, 2021, vol. 21, no. 8, pp. 2873. https://doi.org/10.3390/s21082873
13. He J., Li K., Liao X., Zhang P., Jiang N. Real-time detection of acute cognitive stress using a convolutional neural network from electrocardiographic signal. *IEEE Access*, 2019, vol. 7, pp. 42710–42717. https://doi.org/10.1109/access.2019.2907076
14. Woodward K., Kanjo E., Brown D.J., McGinnity T.M. On-device transfer learning for personalising psychological stress modelling using a convolutional neural network. *arXiv*, 2020, arXiv:2004.01603. https://doi.org/10.48550/arXiv.2004.01603
15. Liapis A., Faliagka E., Katsanos C., Antonopoulos C., Voros N. Detection of subtle stress episodes during UX evaluation: Assessing the performance of the WESAD bio-signals dataset. *Lecture Notes in Computer Science*, 2021, vol. 12934, pp. 238–247. https://doi.org/10.1007/978-3-030-85613-7_17
16. Al Abdi R.M., Alhitary A.E., Abdul Hay E.W., Al-Bashir A.K. Objective detection of chronic stress using physiological parameters. *Medical & Biological Engineering & Computing*, 2018, vol. 56, no. 12, pp. 2273–2286. https://doi.org/10.1007/s11517-018-1854-8
17. Indikawati F.I., Winiarti S. Stress detection from multimodal wearable sensor data. *IOP Conference Series: Materials Science and Engineering*, 2020, vol. 771, no. 1, pp. 012028. https://doi.org/10.1088/1757-899x/771/1/012028
18. Hantono B.S., Nugroho L.E., Santosa P.I. Mental stress detection via heart rate variability using machine learning. *International Journal on Electrical Engineering and Informatics*, 2020, vol. 12, no. 3, pp. 431–444. https://doi.org/10.15676/ijeei.2020.12.3.3
19. Ninh V.T., Smyth S., Tran M.T., Gurrin C. Analysing the performance of stress detection models on consumer-grade wearable devices. *Frontiers in Artificial Intelligence and Applications*, 2021, vol. 337, pp. 524–537. https://doi.org/10.3233/faia210050
20. Albaladejo-González M., Ruipérez-Valiente J.A., Gómez Mármol F. Evaluating different configurations of machine learning models and their transfer learning capabilities for stress detection using heart rate. *Journal of Ambient Intelligence and Humanized Computing*, 2023, vol. 14, no. 8, pp. 11011–11021. https://doi.org/10.1007/s12652-022-04365-z
21. Siirtola P. Continuous stress detection using the sensors of commercial smartwatch. *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers*, 2019, pp. 1198–1201. https://doi.org/10.1145/3341162.3344831
22. Quinlan J.R. *C4.5: Programs for Machine Learning*. Elsevier, 2014, 312 p.
23. Oskooei A., Chau S.M., Weiss J., Sridhar A., Martínez M.R., Michel B. DeStress: deep learning for unsupervised identification of mental stress in firefighters from heart-rate variability (HRV) data. *Studies in Computational Intelligence*, 2021, vol. 194, pp. 93–105. https://doi.org/10.1007/978-3-030-53352-6_9
24. Hochreiter S., Schmidhuber J. Long short-term memory. *Neural Computation*, 1997, vol. 9, no. 8, pp. 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735
25. Xu Q., Nwe T.L., Guan C. Cluster-based analysis for personalized stress evaluation using physiological signals. *IEEE Journal of Biomedical and Health Informatics*, 2015, vol. 19, no. 1, pp. 275–281. https://doi.org/10.1109/jbhi.2014.2311044
26. Garcia-Ceja E., Brena R. Building personalized activity recognition models with scarce labeled data based on class similarities. *Lecture Notes in Computer Science*, 2015, vol. 9454, pp. 265–276. https://doi.org/10.1007/978-3-319-26401-1_25
27. Lu H., Frauendorfer D., Rabbi M., Mast M.S., Chittaranjan G.T., Campbell A.T., Gatica-Perez D., Choudhury T. Stresssense: Detecting stress in unconstrained acoustic environments using smartphones. *Proc. of the 2012 ACM Conference on Ubiquitous Computing*, 2012, pp. 351–360. https://doi.org/10.1145/2370216.2370270
28. Aladin A.I., Whelton S.P., Al-Mallah M.H., Blaha M.J., Keteyian S.J., Juraschek S.P., Rubin J., Brawner C.A., Michos E.D. Relation of resting heart rate to risk for all-cause mortality by gender after

convolutional neural network // Sensors. 2019. V. 19. N 20. P. 4408. https://doi.org/10.3390/s19204408
11. Iqbal T., Redon-Lurbe P., Simpkin A.J., Elahi A., Ganly S., Wijns W., Shahzad A. A sensitivity analysis of biophysiological responses of stress for wearable sensors in connected health // IEEE Access. 2021. V. 9. P. 93567–93579. https://doi.org/10.1109/access.2021.3082423
12. Dalmeida K.M., Masala G.L. HRV features as viable physiological markers for stress detection using wearable devices // Sensors. 2021. V. 21. N 8. P. 2873. https://doi.org/10.3390/s21082873
13. He J., Li K., Liao X., Zhang P., Jiang N. Real-time detection of acute cognitive stress using a convolutional neural network from electrocardiographic signal // IEEE Access. 2019. V. 7. P. 42710–42717. https://doi.org/10.1109/access.2019.2907076
14. Woodward K., Kanjo E., Brown D.J., McGinnity T.M. On-device transfer learning for personalising psychological stress modelling using a convolutional neural network // arXiv. 2020. arXiv:2004.01603. https://doi.org/10.48550/arXiv.2004.01603
15. Liapis A., Faliagka E., Katsanos C., Antonopoulos C., Voros N. Detection of subtle stress episodes during UX evaluation: Assessing the performance of the WESAD bio-signals dataset // Lecture Notes in Computer Science. 2021. V. 12934. P. 238–247. https://doi.org/10.1007/978-3-030-85613-7_17
16. Al Abdi R.M., Alhitary A.E., Abdul Hay E.W., Al-Bashir A.K. Objective detection of chronic stress using physiological parameters // Medical & Biological Engineering & Computing. 2018. V. 56. N 12. P. 2273–2286. https://doi.org/10.1007/s11517-018-1854-8
17. Indikawati F.I., Winiarti S. Stress detection from multimodal wearable sensor data // IOP Conference Series: Materials Science and Engineering. 2020. V. 771. N 1. P. 012028. https://doi.org/10.1088/1757-899x/771/1/012028
18. Hantono B.S., Nugroho L.E., Santosa P.I. Mental stress detection via heart rate variability using machine learning // International Journal on Electrical Engineering and Informatics. 2020. V. 12. N 3. P. 431–444. https://doi.org/10.15676/ijeei.2020.12.3.3
19. Ninh V.T., Smyth S., Tran M.T., Gurrin C. Analysing the performance of stress detection models on consumer-grade wearable devices // Frontiers in Artificial Intelligence and Applications. 2021. V. 337. P. 524–537. https://doi.org/10.3233/faia210050
20. Albaladejo-González M., Ruipérez-Valiente J.A., Gómez Mármol F. Evaluating different configurations of machine learning models and their transfer learning capabilities for stress detection using heart rate // Journal of Ambient Intelligence and Humanized Computing. 2023. V. 14. N 8. P. 11011–11021. https://doi.org/10.1007/s12652-022-04365-z
21. Siirtola P. Continuous stress detection using the sensors of commercial smartwatch // Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers. 2019. P. 1198–1201. https://doi.org/10.1145/3341162.3344831
22. Quinlan J.R. C4.5: Programs for Machine Learning. Elsevier, 2014. 312 p.
23. Oskooei A., Chau S.M., Weiss J., Sridhar A., Martínez M.R., Michel B. DeStress: deep learning for unsupervised identification of mental stress in firefighters from heart-rate variability (HRV) data // Studies in Computational Intelligence. 2021. V. 194. P. 93–105. https://doi.org/10.1007/978-3-030-53352-6_9
24. Hochreiter S., Schmidhuber J. Long short-term memory // Neural Computation. 1997. V. 9. N 8. P. 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735
25. Xu Q., Nwe T.L., Guan C. Cluster-based analysis for personalized stress evaluation using physiological signals // IEEE Journal of Biomedical and Health Informatics. 2015. V. 19. N 1. P. 275–281. https://doi.org/10.1109/jbhi.2014.2311044
26. Garcia-Ceja E., Brena R. Building personalized activity recognition models with scarce labeled data based on class similarities // Lecture Notes in Computer Science. 2015. V. 9454. P. 265–276. https://doi.org/10.1007/978-3-319-26401-1_25
27. Lu H., Frauendorfer D., Rabbi M., Mast M.S., Chittaranjan G.T., Campbell A.T., Gatica-Perez D., Choudhury T. Stresssense: Detecting stress in unconstrained acoustic environments using smartphones // Proc. of the 2012 ACM Conference on Ubiquitous Computing. 2012. P. 351–360. https://doi.org/10.1145/2370216.2370270
28. Aladin A.I., Whelton S.P., Al-Mallah M.H., Blaha M.J., Keteyian S.J., Juraschek S.P., Rubin J., Brawner C.A., Michos E.D. Relation of resting heart rate to risk for all-cause mortality by gender after

considering exercise capacity (the Henry Ford exercise testing project). *The American Journal of Cardiology*, 2014, vol. 114, no. 11, pp. 1701–1706. https://doi.org/10.1016/j.amjcard.2014.08.042
29. Cawley G.C., Talbot N.L.C. Efficient leave-one-out cross-validation of kernel fisher discriminant classifiers. *Pattern Recognition*, 2003, vol. 36, no. 11, pp. 2585–2592. https://doi.org/10.1016/s0031-3203(03)00136-5

considering exercise capacity (the Henry Ford exercise testing project) // The American Journal of Cardiology. 2014. V. 114. N 11. P. 1701–1706. https://doi.org/10.1016/j.amjcard.2014.08.042
29. Cawley G.C., Talbot N.L.C. Efficient leave-one-out cross-validation of kernel fisher discriminant classifiers // Pattern Recognition. 2003. V. 36. N 11. P. 2585–2592. https://doi.org/10.1016/s0031-3203(03)00136-5

## Authors

**Maksim O. Dobrokhvalov** — PhD Student, Saint Petersburg Electrotechnical University “LETI”, Saint Petersburg, 197022, Russian Federation, https://orcid.org/0000-0002-0571-5836, night1337bot@gmail.com

**Anton Yu. Filatov** — PhD, Associate Professor, Saint Petersburg Electrotechnical University “LETI”, Saint Petersburg, 197022, Russian Federation, sc 57194078312, https://orcid.org/0000-0003-4298-8523, aifilatov@etu.ru

## Авторы

**Доброхвалов Максим Олегович** — аспирант, Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» имени В.И. Ульянова (Ленина), Санкт-Петербург, 197022, Российская Федерация, https://orcid.org/0000-0002-0571-5836, night1337bot@gmail.com

**Филатов Антон Юрьевич** — кандидат технических наук, доцент, Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» имени В.И. Ульянова (Ленина), Санкт-Петербург, 197022, Российская Федерация, sc 57194078312, https://orcid.org/0000-0003-4298-8523, aifilatov@etu.ru

*Received 19.05.2023*
*Approved after reviewing 31.10.2023*
*Accepted 22.11.2023*

*Статья поступила в редакцию 19.05.2023*
*Одобрена после рецензирования 31.10.2023*
*Принята к печати 22.11.2023*