

# Двухветвевые сверточные сети для детектирования стресса на основе биомедицинских данных

М. О. Доброхвалов

Санкт-Петербургский государственный  
электротехнический университет  
«ЛЭТИ» им. В.И. Ульянова (Ленина)

night1337bot@gmail.com

А. Ю. Филатов

Санкт-Петербургский государственный  
электротехнический университет  
«ЛЭТИ» им. В.И. Ульянова (Ленина)

aifilatov@etu.ru

**Аннотация.** Стресс считается одной из основных причин проблем со здоровьем. Объединение физиологических сигналов нескольких модальностей является перспективным методом для более точного определения состояния человека. В этой статье предлагается двухветвевая модель сверточной сети, предназначенная для детектирования стресса. Модель независимо обрабатывает данные сердца (электрокардиограмму или давление пульса крови) и электрическую активность кожи, после чего объединяет результаты ветвей с помощью предиктора. В работе рассматриваются различные предикторы. В некоторых в функцию потерь передается только финальный результат сети, а в одном - учитываются результаты каждой ветви и сети в целом. Экспериментальные результаты с набором данных WESAD показывают, что предложенная модель двухветвевая модель может повысить точность детектирования стресса по сравнению с одномодальными сетями. Также предложенные методы превосходят некоторые современные решения.

**Ключевые слова:** детектирование стресса; сверточные нейронные сети; электрокардиограмма; электрическая активность кожи

## I. ВВЕДЕНИЕ

Стресс – естественная реакция человека в ответ на внешние раздражители [1]. Длительное нахождение в состоянии стресса может приводить к негативному влиянию на внимание, память и принятие решений. Затем приводит к преждевременному старению, когнитивным нарушениям, тревоге, депрессии, сердечно-сосудистым заболеваниям и т. д. [2-3]. Таким образом, детектирование стресса является важной задачей.

В различных исследованиях используются классические методы машинного обучения и также нейронные сети. Электрокардиограмма (ЭКГ, ECG), электрическая активность кожи (ЭАК, EDA), фотоплетизмограмма (PPG), пульс объема крови (BVP), электромиограмма (ЭМГ, EMG), дыхание и др. используются для определения стресса. Таким образом, целью работы является разработка метода определения психологического стресса с помощью сверточных нейронных сетей на основе биомедицинских данных, таких как ECG, BVP, EDA, а также реализация данного метода.

В работе рассматривается использование и двухветвевых сверточных сетей. В двухветвевой архитектуре используются данные ECG/BVP и EDA. Также исследуются три различных способа объединения результатов, полученных из ветвей. Изучается влияние данных EDA. Предлагаемый метод конкурентоспособен по сравнению с другими современными методами.

## II. ОБЗОР СУЩЕСТВУЮЩИХ РЕШЕНИЙ

Часто используемыми наборами данных в исследованиях являются WESAD [4], UBFC-Phys [5].

Среди методов машинного обучения используются решающие деревья (DT) [4], случайный лес (RF) [4, 6–7], метод k-ближайших соседей (kNN) [4–6], метод опорных векторов (SVM) [5–6], линейный дискриминантный анализ (LDA) [4], AdaBoosting (AB) [4], логистическую регрессию (LR) [5] и другие. Liapis A. и др. [8] использовали классификатор STree. Полносвязные нейронные сети использовались в работах [6, 9]. В работе [10] авторы использовали полносвязную нейронную сеть в федеративном обучении. В работе [11] авторы используют многоветвевую сверточную нейронную сеть, в каждую ветвь передается отдельная модальность данных, содержащаяся в датасете WESAD. Авторы работы [12] используют двухветвевую сеть с дополнительным двусторонним механизмом внимания для передачи промежуточного представления между ветвями. Ghosh в работе [13] описывает применение сверточных нейронных сетей, входными данными являются угловые поля Грамма (Gramian Angular Field). Hasanroog и др. [14] демонстрируют модель, объединяющую многослойный перцептрон и полносвязную нейронную сеть. В работе [15] авторы описывают двухветвевую нейронную сеть. Для извлечения особенностей (features) из данных BVP используется сверточная нейронная сеть, для данных EDA – LSTM сеть. Результаты сетей обрабатываются специальными механизмами перекрестного и внутреннего внимания, после чего объединяются с помощью специального предиктора. В статье [16] предлагаются сети гибридного глубокого обучения на основе удаленной фотоплетизмографии (rPPG), использующие LSTM, GRU, одномерные сверточные сети.

### III. ДАННЫЕ

В качестве входных данных рассматриваются ECG, BVP, EDA. Для фильтрации сигналов использовались соответствующие методы из библиотеки NeuroKit2<sup>1</sup>.

Из датасета WESAD используются данные ECG, EDA, снятые с нагрудного датчика, BVP - с запястья. Из датасета UBFC-Phys используются данные BVP с запястья. Данные EDA с запястья из обоих датасетов не используются, так как имеют слишком маленькую частоту дискретизации (4 Гц). В датасете UBFC-Phys представлены данные 56 субъектов, однако в рамках этого исследования используются данные только 15 субъектов. Остальные субъекты были исключены в соответствии с причинами, описанными в публикации авторов датасета. Все сигналы были приведены к общей частоте дискретизации – 256 Гц.

Также сигнал EDA был разделен на фазовую и тоническую составляющие. В датасетах были представлены кратковременные раздражители, поэтому использовалась только фазовая, так как она демонстрирует отклик нервной системы.

### IV. МЕТОД

В данной работе предлагается использование сверточных нейронных сетей для определения стресса. Рассматривается как использование одноветвевой, так и двухветвевой сети.

#### A. Одноветвевая сеть

Как упоминалось ранее, в качестве входного сигнала для одноветвевой сети использовались данные ECG, BVP, фазовая составляющая EDA датасета WESAD и данные BVP из датасета UBFC-Phys.

Архитектура сети была представлена в статье [17]. В рамках текущих исследований использовалась длина входного сигнала равная 15 и 30 секунд. Размер ядра свертки – 64. Исследовалось 3 метода уменьшения размерности (MUP): одномерный сверточный слой (conv), усредняющий (avg) и максимизирующий (max) пулинги. Использовались оптимизаторы Adam [18] и ASGD [19]. Коэффициент обучения оптимизатора – 0.01.

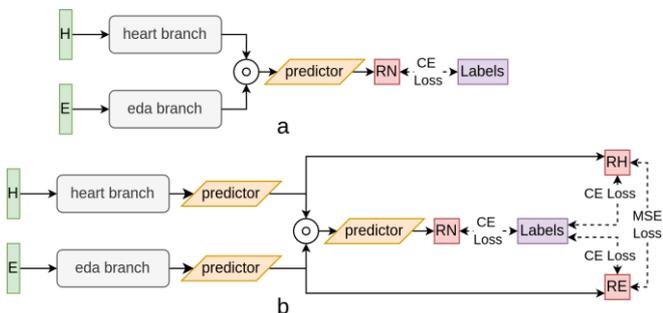


Рис. 1. Архитектура двухветвевой сети. а – способы объединения CL, CC; б – способ объединения LCL

#### B. Двухветвевая сеть

Двухветвевая сеть исследовалась только в рамках датасета WESAD, так как в датасете UBFC-Phys нет двух

источников данных. Архитектура двухветвевой сети представлена на рис. 1. Каждая ветвь представляет собой сеть для обработки одного типа сигнала, описанную ранее. Первым источником сигнала использовались ECG/BVP, в качестве второго сигнала – EDA. В качестве способов объединения результатов ветвей было рассмотрено 3 варианта:

- конкатенация с последующим полносвязным слоем (CL);
- конкатенация с последующим сверточным слоем (CC);
- линейный слой в конце каждой ветви, конкатенация результатов, ещё один полносвязный слой (LCL).

При использовании способов объединения CL и CC результаты ветвей конкатенировать, после чего обрабатывались полносвязным (CL) или однослойным сверточным (CC) слоем. Финальный результат сети передавался в функцию потерь (перекрестную энтропию). При использовании способа объединения LCL результаты каждой ветви обрабатывались полносвязным слоем и передавались в функцию потерь (перекрестную энтропию). Также результаты ветвей передавались в функцию потерь среднеквадратичной ошибки. Выходы ветвей конкатенировались, обрабатывались дополнительным полносвязным слоем и также передавались в функцию потерь (перекрестную энтропию).

### V. РЕЗУЛЬТАТЫ

Эксперименты проводились с использованием процессора Intel® Core™ i5-8300H CPU @ 2.30GHz, 16GB RAM памяти частотой 2667 MHz и графическим процессором NVIDIA GeForce GTX 1050 Mobile 4096MiB VRAM. Использовалась 64-битная операционная система Ubuntu 22.04.1 LTS. Используемые в работе метрики: точность (accuracy), достоверность (precision), полнота (recall), F1 мера.

#### A. Одноветвевая сеть

В таблице I представлены значения метрик лучших по F1-мере одноветвевых сетей на данных датасета WESAD. На основании таблицы можно сделать вывод, что использование данных ЭКГ позволяет получить более высокую точность, чем при использовании данных BVP или фазовой составляющей EDA. Также стоит отметить, что сети с усредняющим пулингом в качестве MUP демонстрировали наилучший или близкий к наилучшему результат. Среди одноветвевых сетей наибольшую точность (0.983) и F1 меру (0.975) имеет модификация, использующая усредняющий пулинг в качестве MUP, в качестве входного сигнала принимающая данные ЭКГ, длительностью 30 секунд, обученная с помощью оптимизатора ASGD. Также стоит отметить, что модификация с тем же входным сигналом, но использующая максимизирующий пулинг и обученная с помощью оптимизатора Adam, имеет такую же точность, а значение F1 меры меньше на 0.0005.

<sup>1</sup> <https://neuropsychology.github.io/NeuroKit/index.html>

ТАБЛИЦА I. РЕЗУЛЬТАТЫ ОДНОВЕТВЕВЫХ СЕТЕЙ НА WESAD

Данные	Длина, с	МУР	Оптимизатор	Точность	Достоверность	Полнота	F1 мера
ECG	15	avg	ASGD	0,980	0,984	0,961	0,972
ECG	30	avg	ASGD	0,983	0,960	0,992	0,976
BVP	15	conv	Adam	0,908	0,860	0,887	0,873
BVP	30	avg	Adam	0,951	0,957	0,902	0,928
EDA	15	avg	ASGD	0,869	0,858	0,758	0,805
EDA	30	avg	ASGD	0,890	0,889	0,787	0,835

ТАБЛИЦА II. РЕЗУЛЬТАТЫ ОДНОВЕТВЕВЫХ СЕТЕЙ НА UBFC-PHYS

Длина, с	Точность	Достоверность	Полнота	F1 мера
15	0,922	0,927	0,958	0,943
30	0,922	0,896	1,000	0,945

ТАБЛИЦА III. РЕЗУЛЬТАТЫ ДВУХВЕТВЕВЫХ СЕТЕЙ НА WESAD

Данные	Длина, с	Предиктор	Точность	Достоверность	Полнота	F1 мера
ECG	15	CC	0,990	0,984	0,988	0,986
ECG	30	CC	0,994	0,984	1,000	0,992
BVP	15	CL	0,955	0,955	0,918	0,936
BVP	30	CC	0,977	0,967	0,967	0,967

ТАБЛИЦА IV. ОПИСАТЕЛЬНЫЕ ХАРАКТЕРИСТИКИ РАЗНОСТИ МЕТРИК ДВУХВЕТВЕВЫХ И ОДНОВЕТВЕВЫХ СЕТЕЙ

Характеристики разности	ECG		BVP	
	Точность	F1 мера	Точность	F1 мера
Среднее	0,012	0,017	0,038	0,054
Стандартное отклонение	0,020	0,031	0,020	0,028
Минимальное значение	-0,008	-0,013	0,017	0,024
Максимальное значение	0,067	0,106	0,085	0,118

Однако у данной модификации значение полноты равно 0,975, что на 0,017 ниже, чем у лучшей модификации. В рамках исследуемой задачи полнота – более приоритетная, чем достоверность.

В таблице II представлены значения метрик лучших по F1-мере одноветвевых сетей на данных датасета UBFC-Phys. Наибольшую точность продемонстрировали модификации, использующие усредняющий пулинг в качестве МУР, обученные с помощью оптимизатора ASGD. У модификации (первая), принимающей в качестве входного сигнала данные, длительностью 30 секунд, значение F1 меры (0.945) на 0.02 больше, чем у той, что принимает сигнал размерностью 15 секунд. При этом полнота у первой модификации больше на 0.042. И как упоминалось ранее, в рамках данной задачи полнота является более приоритетной метрикой.

ТАБЛИЦА V. СРАВНЕНИЕ СПОСОБОВ СЛИЯНИЯ ВЕТВЕЙ

Сигнал	Длина, с	МУР	CC	CL	LCL	Лучший предиктор
ECG	15	avg	0.986	0.969	0.981	CC
ECG	15	conv	0.895	0.915	0.904	CL
ECG	15	max	0.962	0.975	0.969	CL
ECG	30	avg	0.992	0.98	0.976	CC
ECG	30	conv	0.87	0.896	0.944	LCL
ECG	30	max	0.984	0.95	0.967	CC
BVP	15	avg	0.935	0.936	0.926	CL
BVP	15	conv	0.861	0.88	0.872	CL
BVP	15	max	0.914	0.903	0.898	CC
BVP	30	avg	0.967	0.954	0.933	CC
BVP	30	conv	0.86	0.874	0.92	LCL
BVP	30	max	0.95	0.947	0.937	CC

ТАБЛИЦА VI. СРАВНЕНИЕ МЕТОДОВ НА ДАТАСЕТЕ WESAD

Статья	Данные	Точность	F1 мера
[4]	ECG	0,854	0,8131
[4]	Chest	0,928	0,9107
[6]	EDA	0,875	-
[8]	EDA	0,958	0,951
[9]	BVP	0,990	0,990
[10]	EDA, PPG, ACC	0,868	0,97
[11]	ECG	0,978 ± 0,026	0,905 ± 0,129
[11]	All	0,918 ± 0,097	0,977 ± 0,026
[12]	ECG, EDA	0,921	0,9111
[13]	All	0,948	0,95
[15]	BVP, EDA	0,721 ± 0,050	0,670±0,047
Ours	BVP	0,951	0,928
Ours	ECG	0,983	0,976
Ours	BVP, EDA	0,977	0,967
Ours	ECG, EDA	0,994	0,992

ТАБЛИЦА VII. СРАВНЕНИЕ МЕТОДОВ НА ДАТАСЕТЕ UBFC-PHYS

Статья	Данные	Точность	F1 мера
[5]	BVP	0,855	-
[7]	rPPG	0,831	0,851
[14]	BVP	0,820	-
[15]	BVP, EDA	0,818 ± 0,063	0,817±0,063
[16]	rPPG	0,958	0,96
Ours	BVP	0,922	0,945

### В. Двухветвевая сеть

В таблице III представлены значения метрик лучших по F1-мере двухветвевых сетей на данных датасета WESAD. У всех модификаций в таблице МУР – усредняющий пулинг, используемый оптимизатор – ASGD. На основании таблицы можно сделать вывод, что наибольшие значения точности (0.994) и F1 меры (0.992) имеет модификация с усредняющим пулингом, предиктором CC, принимающий в качестве входного сигнала данные ECG и фазовой составляющей EDA длительностью 15 секунд. В таблице IV представлены базовые описательные статистические значения разницы точности и F1 меры между лучшей (по F1 среди модификаций предикторов) мере двухветвевой и одноветвевой модификациями сетей. Соответствие модификаций основывалось на используемом сигнале сердца и его длине, используемых методе уменьшения размерности и оптимизаторе.

На основании таблицы можно сделать вывод, что добавление данных EDA в большинстве случаев позволяет увеличить значение точности и F1-меры. Присутствие отрицательного значения в таблице говорит в разделе с ЭКГ о том, что данные ЭКГ содержат достаточно большое количество особенностей. Однако в случае с данными BVP добавление второй ветви с данными EDA позволило увеличить значения точности на 0.038±0.020 и F1 меры на 0.054±0.028.

Стоит отметить, что 3 из 4 наилучших модификаций в таблице III используют предиктор CC. В таблице V представлено сравнение значений F1 меры различных способов объединения результатов ветвей. В рамках таблицы рассматриваются модификации, обученные с помощью оптимизатора ASGD. Строки с наилучшими модификациями из таблицы III отмечены в таблице V. Можно увидеть, что единственная из наилучших модификаций, использующая предиктор CL, имеет значение F1 меры на 0.001 больше, чем с предиктором

СС. На основании таблицы можно сделать вывод, что половина модификаций (6 из 12) показывает наилучший результат с использованием предиктора СС. Также стоит отметить, что предиктор LCL, имеющий наибольшее значение F1 меры при сравнении различных предикторов, имеет меньшее значение F1 меры, нежели другие модификации с аналогичным сигналом (по длине и типу), но различными способами уменьшения размерности. В случаях, когда наибольшее значение F1 меры имеют модификации с предиктором CL, значение меры не превышает значение модификации СС более чем 0.02. Таким образом, можно заключить, что предиктор СС является лучшей вариацией среди рассматриваемых.

На основании таблиц VI и VII можно увидеть, что предложенные в данной работе методы сопоставимы с современными методами.

## VI. Выводы

В данной работе предложен метод детектирования стресса с использованием сверточных нейронных сетей на основе биомедицинских данных, таких как электрокардиограмма, давление пульса крови, фазовую составляющую электрической активности кожи.

Рассмотрена 1-ветвевая архитектура. На датасете WESAD для данных электрокардиограммы значение метрик точности и F1 меры 0,983 и 0,976 соответственно, для давления пульса крови – 0,951 и 0,928, для электрической активности кожи – 0,890 и 0,835. На датасете UBFC-Phys для данных давления пульса крови значение метрик точности и F1 меры 0,922 и 0,945, соответственно. Также предложена 2-ветвевая архитектура, одна ветвь которой обрабатывает данные электрокардиограммы или давление пульса крови, вторая – фазовую составляющую электрической активности кожи. Рассмотрены различные способы объединения результатов ветвей. На датасете WESAD при использовании данных электрокардиограммы значение метрик точности и F1 меры – 0,994 и 0,992, при использовании данных пульса объема крови – 0,977 и 0,967. Прирост метрик точности и F1 меры при использовании двухветвевой архитектуры по сравнению с соответствующей одноветвевой сетью, использующей только данные электрокардиограммы –  $0,012 \pm 0,017$  и  $0,017 \pm 0,031$ , только данные пульса объема крови –  $0,038 \pm 0,020$  и  $0,054 \pm 0,028$ .

В дальнейших исследованиях планируется изучение механизма перекрестного внимания для коррекции взаимного влияния веток.

## СПИСОК ЛИТЕРАТУРЫ

- [1] Cohen S., Kessler R.C., Gordon L.U. Strategies for measuring stress in studies of psychiatric and physical disorders // *Measuring stress: A guide for health and social scientists*. 1995. Т. 28. С. 3-26.
- [2] Bobade P., Vani M. Stress detection with machine learning and deep learning using multimodal physiological data // 2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA). IEEE, 2020. С. 51-57.
- [3] 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 preprint arXiv:2005.01834*. 2020.
- [4] Schmidt P., Reiss A., Duerichen R., Marberger C., Van Laerhoven K. Introducing wesad, a multimodal dataset for wearable stress and affect detection // *Proceedings of the 20th ACM international conference on multimodal interaction*. 2018. С. 400-408.
- [5] Sabour R.M., Benezeth Y., De Oliveira P., Chappe J., Yang F. Ubfcphys: A multimodal database for psychophysiological studies of social stress // *IEEE Transactions on Affective Computing*. 2021. Т. 14. №. 1. С. 622-636.
- [6] Ninh V.T., Smyth S., Tran M.T., Gurrin C. Analysing the performance of stress detection models on consumer-grade wearable devices // *arXiv preprint arXiv:2203.09669*. 2022.
- [7] Talukdar D., de Deus L.F., Sehgal N., De Deus L.F. Evaluating visual photoplethysmography method // *Cureus*. 2022. Т. 14. №. 7.
- [8] 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 // *Human-Computer Interaction-INTERACT 2021: 18th IFIP TC 13 International Conference, Bari, Italy, August 30-September 3, 2021, Proceedings, Part III 18*. Springer International Publishing, 2021. С. 238-247.
- [9] Albaladejo-González M., Ruipérez-Valiente J.A., Gómez Marmol 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. Т. 14. №. 8. С. 11011-11021.
- [10] Almadhor A., Sampedro G.A., Abisado M., Abbas S., Kim Y.J., Khan M.A., Baili J., Cha J.H. Wrist-based electrodermal activity monitoring for stress detection using federated learning // *Sensors*. 2023. Т. 23. №. 8. С. 3984.
- [11] Lai K., Yanushkevich S.N., Shmerko V.P. Intelligent stress monitoring assistant for first responders // *IEEE Access*. 2021. Т. 9. С. 25314-25329.
- [12] Bhatti A., Behinaein B., Rodenburg D., Hungler P., Etemad A. Attentive cross-modal connections for deep multimodal wearable-based emotion recognition // 2021 9th international conference on affective computing and intelligent interaction workshops and demos (ACIIW). IEEE, 2021. С. 01-05.
- [13] 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. Т. 12. №. 12. С. 1153.
- [14] Hasanpoor Y., Motaman K., Tarvirdizadeh B., Alipour K., Ghamari M. Stress detection using ppg signal and combined deep cnn-mlp network // 2022 29th National and 7th International Iranian Conference on Biomedical Engineering (ICBME). IEEE, 2022. С. 223-228.
- [15] Zhang X., Wei X., Zhou Z., Zhao Q., Zhang S., Yang Y., Li R., Hu B. Dynamic alignment and fusion of multimodal physiological patterns for stress recognition // *IEEE Transactions on Affective Computing*. 2023.
- [16] Fontes L., Machado P., Vinkemeier D., Yahaya S., Bird J.J., Ihianle I.K. Enhancing Stress Detection: A Comprehensive Approach through rPPG Analysis and Deep Learning Techniques // *Sensors*. 2024. Т. 24. №. 4. С. 1096.
- [17] 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
- [18] Kingma D.P., Ba J. Adam. A method for stochastic optimization // *arXiv preprint arXiv:1412.6980*. 2014.
- [19] Polyak B.T., Juditsky A.B. Acceleration of stochastic approximation by averaging // *SIAM journal on control and optimization*. 1992. Т. 30. №. 4. С. 838-855.