

**В.А. Деркачев, В.В. Бахчевников, А.Н. Бакуменко**

**КЛАССИФИКАТОР ИЗОБРАЖЕНИЙ СЕМЯН  
СЕЛЬСКОХОЗЯЙСТВЕННЫХ КУЛЬТУР С ИСПОЛЬЗОВАНИЕМ  
СВЕРТОЧНОЙ НЕЙРОННОЙ СЕТИ**

*В настоящей статье рассматривается создание архитектуры сверточной нейронной сети, классифицирующей изображения сельскохозяйственных культур (в частности пшеницы) для последующего применения в оптическом сепараторе семян (фотосепараторе). Интерес к проектированию нейронных сетей классификации изображений в последнее время сильно возрос, что связано как с развитием теории глубоких нейронных сетей, так и возросшей вычислительной мощностью настольных компьютеров, а также переносом вычислений на графические процессоры. Целью статьи является разработка архитектуры нейронной сети позволяющей осуществить разделение входного потока семян пшеницы на два класса: «хорошие» семена и «плохие» (с изъянами по форме и цвету) семена. Архитектура полученной нейронной сети является сверточной, так как в отличие от полносвязной, данный класс нейронных сетей в определенных пределах невосприимчив к изменению масштаба и угла поворота объектов во входных данных. В работе для формирования обучающей, валидационной и тестовой выборки использовались изображения семян, полученные с использованием бытовой фотокамеры, что негативно сказалось на результатах обучения и тестирования нейронной сети относительно возможного результата применения в реальном фотосепараторе. Архитектура разработанной нейронной сети предварительно оптимизирована для использования на ПЛИС, однако, в рассмотренном случае не осуществлен переход от значений весовых коэффициентов из типа данных с плавающей запятой к целочисленному типу, что может привести к снижению точности работы нейронной сети, при этом позволив значительно уменьшить объем ресурсов ПЛИС. Применение предложенной архитектуры позволяет получить достаточно точную оценку классифицируемых семян пшеницы по верификационным и тестовым наборам данных.*

*Нейронные сети; сверточные нейронные сети; классификация изображений; классификация семян пшеницы.*

**V.A. Derkachev, V.V. Bakhchevnikov, A.N. Bakumenko**

**CLASSIFIER OF IMAGES OF AGRICULTURAL CROPS SEEDS USING  
A CONVOLUTION NEURAL NETWORK**

*This article discusses the creation of a convolutional neural network architecture that classifies images of crops (in particular wheat) for subsequent use in an optical seed separator (photo separator). Interest in the design of neural networks for classifying images has recently increased significantly, which is associated both with the development of the theory of deep neural networks and the increased computing power of desktop computers, as well as the transfer of computing to graphic processors. The aim of the article is to develop the architecture of a neural network that allows the separation of the input flow of wheat seeds into two classes: “good” seeds and “bad” (with defects in shape and color) seeds. The architecture of the resulting neural network is convolutional, because, unlike a fully connected one, this class of neural networks is within certain limits immune to changes in the scale and angle of rotation of objects in the input data. In the work, for the formation of training, validation and test samples, seed images obtained using a household camera were used, which negatively affected the results of training and testing the neural network regarding the possible result of application in a real photo separator. The architecture of the developed neural network is preliminarily optimized for use on FPGAs, however, in the considered case, the transition from the values of weighting factors from the data type from a floating point to an integer type has not been made, which can lead to a decrease in the accuracy of the neural network, while significantly reducing the amount of resources FPGA. Application of the proposed architecture allows one to obtain a fairly accurate estimate of classified wheat seeds from verification and test data sets.*

*Neural networks; convolutional neural networks; image classification; wheat seed classification.*

**Введение.** Основной проблемой классификации изображений является выявление особенностей, с помощью которых возможно устойчивое определение класса изображений. Одним из методов классификации изображений является корреляционный метод [1]. Так же в качестве примера классического классификатора можно упомянуть статистически оптимальные байесовские классификаторы. В таких классификаторах все наблюдаемые объекты, описываются с применением специального набора признаков с использованием байесовского правила классификации [2, 3]. Классические подходы требуют создание признаков вручную для каждого класса объектов, что весьма трудозатратно, особенно при разработке классификаторов сложных объектов с большим количеством признаков. Помимо трудностей ручного задания признаков существует проблема малой скорости подстройки классификаторов под конкретные задачи. Данная проблема связана с наличием большого числа вариантов признаков объектов, которые необходимо учесть для правильной классификации объектов. Перечисленные проблемы увеличивают финансово-временные затраты, связанные с разработкой традиционных классификаторов под каждый из возможных классов объектов.

Использование нейронных сетей для классификации изображений является многообещающим направлением относительно классических методов обладая следующими преимуществами: возможность работать с большими или неточными данными, выделяя особенности, которые не очевидны для человека.

Развитие сверточных нейронных сетей со времени работ Ю. Лекуна [4] в последнее десятилетие сильно продвинулось в реализуемой точности и сложности моделей [5, 6], что значительно расширило область применения этой технологии.

Целью данной работы является разработка нейросетевой архитектуры классификатора изображений семян сельскохозяйственных культур (в данном случае, пшеницы) для дальнейшего применения в оптическом сортировочном устройстве (фотосепараторе). Из-за различий в ориентации семян и их размера во входных оптических изображениях предлагается глубокая сверточная нейронная сеть в качестве архитектуры классификатора. Данная архитектура построения нейронной сети уменьшает негативные эффекты, возникающие из-за возможного низкого качества входных изображений, их ориентации и масштаба.

**Архитектура сверточной нейронной сети.** В этой статье была использована сверточная нейронная сеть. Структура нейронной сети подразумевает наличие различного типа слоев: входного слоя, нескольких скрытых слоев и выходного слоя. Охарактеризуем используемые виды слоев:

- ◆ **Входной слой.** В этом слое каждому нейрону соответствует отдельная точка изображения с учетом количества цветовых каналов.

- ◆ **Сверточный слой.** Осуществляет двумерную свертку значений из предыдущего уровня, используя набор ядер свертки (фильтров). Обычно в каждом слое содержится определенный набор различных ядер свертки, что необходимо для выделения нескольких признаков в одном слое. Значения коэффициентов фильтров перед обучением устанавливаются случайным образом или с использованием специальных алгоритмов (обычно алгоритмов Ксавье [7] или Хе [8]), что важно для возможности достижения глобального минимума в процессе обучения. Далее каждое из ядер свертки обучается для выделения определенных признаков из массива данных, полученных из предыдущего слоя. На выходе сверточного слоя, в случае отсутствия дополнения краевыми точками, размерность массива данных уменьшается в соответствии со следующим правилом:  $(n - k + 1)$ , где  $n$  – является размерность массива данных на входе,  $k$  – размерность используемого ядра свертки.

- ◆ **Слой субдискретизации.** В данном слое происходит уменьшение размерности данных (обобщение) путем объединения значений входной матрицы данных в рамках определенной области в одно значение. Этот слой позволяет увеличить

устойчивость нейронной сети к изменению масштаба изображений, а также уменьшает число требуемых вычислений в последующих слоях нейронной сети. Существует несколько возможных вариантов субдискретизации: по максимуму [9, 10] (самый популярный вариант в настоящее время), по среднему значению [11] и промежуточные варианты [12]. В данной архитектуре применена субдискретизация по максимальному значению.

◆ Полносвязный слой. Нейроны в данном слое связаны с каждым из предыдущего и осуществляют конечное разделение на классы.

В случае классификатора семян сельскохозяйственных культур можно предложить следующую архитектуру сети:

1. Входное изображение размерностью  $3 \times 91 \times 91$  (3 канала цвета при разрешении  $91 \times 91$ );

2. Сверточный слой с ядрами свертки  $5 \times 5$  в количестве 10 штук. Так как изображение не дополняется по краям дополнительно, выходная карта свертки имеет размерность  $10 \times 87 \times 87$ , функция активации ReLU;

3. Слой субдискретизации с окном  $2 \times 2$  и шагом 2. Выходная размерность слоя составила  $10 \times 43 \times 43$ ;

4. Сверточный слой с ядрами свертки  $5 \times 5$  в количестве 10 штук. Так как изображение не дополняется по краям дополнительно, выходная карта свертки имеет размерность  $10 \times 39 \times 39$ , функция активации ReLU;

5. Слой субдискретизации с окном  $2 \times 2$  и шагом 2. Выходная размерность слоя составила  $10 \times 19 \times 19$ ;

6. Сверточный слой с ядрами свертки  $3 \times 3$  в количестве 20 штук. Так как изображение не дополняется по краям дополнительно, выходная карта свертки имеет размерность  $20 \times 17 \times 17$ , функция активации ReLU;

7. Сверточный слой с ядрами свертки  $3 \times 3$  в количестве 20 штук. Так как изображение не дополняется по краям дополнительно, выходная карта свертки имеет размерность  $20 \times 15 \times 15$ , функция активации ReLU;

8. Слой субдискретизации с окном  $2 \times 2$  и шагом 2. Выходная размерность слоя составила  $20 \times 7 \times 7$ ;

9. Сверточный слой с ядрами свертки  $3 \times 3$  в количестве 20 штук. Так как изображение не дополняется по краям дополнительно, выходная карта свертки имеет размерность  $20 \times 5 \times 5$ , функция активации ReLU;

10. Сверточный слой с ядрами свертки  $3 \times 3$  в количестве 20 штук. Так как изображение не дополняется по краям дополнительно, выходная карта свертки имеет размерность  $20 \times 3 \times 3$ , функция активации ReLU;

11. Полносвязный слой с числом нейронов 128, функция активации ReLU;

12. Полносвязный слой с числом нейронов 32, функция активации ReLU;

13. Выходной слой с двумя нейронами, соответствующими двум классифицируемыми классам, функция активации гиперболический тангенс.

Полученная архитектура имеет 43 312 весовых коэффициентов. На рис. 1 схематично показана итоговая архитектура нейронной сети.

**Параметры модели и обучения.** Одним из важнейших параметров нейросети является функция активации нейронов. В данной работе используется два вида функции активации: в большинстве слоев функция «линейный выпрямитель» ReLU (rectified linear unit) и гиперболический тангенс для выходного слоя. Выбор функции активации ReLU в качестве основной был сделан исходя из следующих соображений: данная функция ускоряет обучение нейронной сети [13], приводит к лучшим решениям [14, 15], чем обычные сигмоидоподобные функции, при этом в отличие от гиперболического тангенса и сигмоиды имеет малую вычислительную сложность, что принципиально важно для дальнейшей реализации разрабатываемой



В обоих классах были выделены следующие подклассы: обучающий подкласс, состоящий из 400 изображений, верификационный подкласс, состоящий из 68 изображений и проверочный, состоящий из 50 изображений. Всего получаем следующий набор данных: тренировочные изображения – 800 штук, верификационные изображения – 136 штук и тестовые изображения – 100 штук. Все используемые данные для каждого этапа обучения берутся в случайном порядке, за исключением тестовых данных, которые сохраняют исходную последовательность для корректного последующего сравнения результатов обучения нейронной сети.

Обучение и тестирование нейронной сети произведено в Tensorflow версии 2.1 установленной в системе Ubuntu 18.04 LTS. Для обучения нейронной сети применялась система с графическим процессором Nvidia 2070 Super, что позволило осуществлять обучение нейронной сети за 25-27 секунд. Графики точности нейронной сети и функции потерь в процессе обучения показаны на рис. 4.

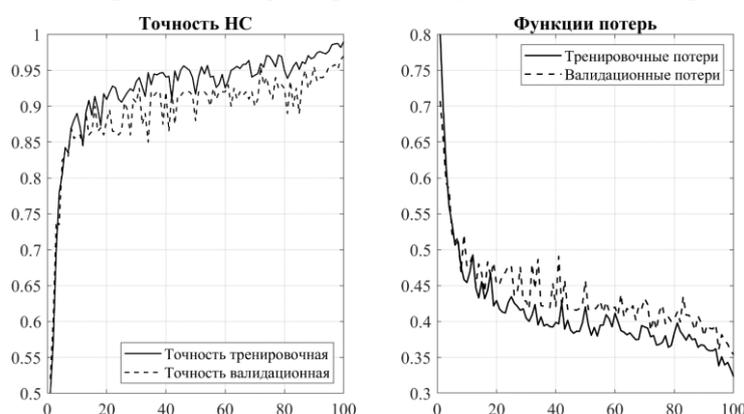


Рис. 4. Графики процесса обучения нейронной сети

Результаты обучения (100 этапов по 8 изображений) показаны в табл. 1.

Таблица 1

Параметр	Потери тренировочные	Потери валидационные	Точность тренировочная	Точность валидационная	Точность тестовая
Значение	0,3233	0,3540	0,9900	0,9600	<b>0,9700</b>

**Выводы.** Исходя из результатов показанных в табл. 1, точность классификации на основе предложенной архитектуры глубокой сверточной нейронной сети на тестовой выборке достигла 97 %. Однако, следует учесть тот факт, что обучающая и тестовая выборки достаточно сильно ограничены в объеме (800 и 100 соответственно). Данные ограничения, возможно, могут привести к некоторой ошибке оценки правильности классификации изображений. Помимо малого объема данной выборки следует отметить, что при реальном применении данного алгоритма в фотосепараторе ожидается увеличение вероятности правильной классификации из-за отсутствия теней на изображении с семенами, а также отсутствием фона. Результаты получены с использованием типа данных с плавающей запятой, для применения результатов на ПЛИС желательно осуществить переход к целочисленному типу данных (INT8) [21], так как полученная нейронная сеть обладает значительным количеством весовых коэффициентов (43 312), что может затруднить реализацию даже на самых больших ПЛИС настоящего времени.

Научные исследования выполнены в рамках проекта «Создание высокотехнологического производства аппаратно-программных комплексов обработки сельхозсырья на основе СВЧ излучения» (Соглашение с Минобрнауки РФ № 075-11-2019-083 от 20.12.2019 г., Договор ЮФУ № 18 от 20.09.2019 г., номер работы в ЮФУ № ХД/19-25-РТ).

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Гонсалес Р., Вудс Р. Цифровая обработка изображений: пер. с англ. – М.: Техносфера, 2005. – 1073 с.
2. Anderson T.W., Goodman L.A. Statistical inference about Markov chains // The Annals of Mathematical Statistics. – 1957. – P. 89-110.
3. Вагис А.А. Процедура распознавания на байесовских сетях // Компьютерная математика. – 2010. – № 2. – С. 124-130.
4. LeCun Y. et al. Gradient-based learning applied to document recognition // Proceedings of the IEEE. – 1998. – Vol. 86, Issue 11. – P. 2278-2324.
5. Gu J. et al. Recent advances in convolutional neural networks // Pattern Recognition. – 2018. – Vol. 77. – P. 354-377.
6. Srinivas S. et al. A taxonomy of deep convolutional neural nets for computer vision // Frontiers in Robotics and AI. – 2016. – Vol. 2. – P. 36.
7. Glorot X., Bengio Y. Understanding the difficulty of training deep feedforward neural networks // Proceedings of the thirteenth international conference on artificial intelligence and statistics. – 2010. – P. 249-256.
8. He K. et al. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification // Proceedings of the IEEE international conference on computer vision. – 2015. – P. 1026-1034.
9. LeCun Y. et al. Backpropagation applied to handwritten zip code recognition // Neural computation. – 1989. – Vol. 1, Issue 4. – P. 541-551.
10. Boureau Y.L. et al. Learning mid-level features for recognition // 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. – IEEE, 2010. – P. 2559-2566.
11. Scherer D., Müller A., Behnke S. Evaluation of pooling operations in convolutional architectures for object recognition // International conference on artificial neural networks. – Springer, Berlin, Heidelberg, 2010. – P. 92-101.
12. Boureau Y.L., Ponce J., LeCun Y. A theoretical analysis of feature pooling in visual recognition // Proceedings of the 27th international conference on machine learning (ICML-10). – 2010. – P. 111-118.
13. Krizhevsky A., Sutskever I., Hinton G.E. Imagenet classification with deep convolutional neural networks // Advances in neural information processing systems. – 2012. – P. 1097-1105.
14. Maas A.L., Hannun A.Y., Ng A.Y. Rectifier nonlinearities improve neural network acoustic models // Proc. icml. – 2013. – Vol. 30, Issue No. 1. – P. 3.
15. Zeiler M.D. et al. On rectified linear units for speech processing // 2013 IEEE International Conference on Acoustics, Speech and Signal Processing. – IEEE, 2013. – P. 3517-3521.
16. Kalman B.L., Kwasny S.C. Why tanh: choosing a sigmoidal function // [Proceedings 1992] IJCNN International Joint Conference on Neural Networks. – IEEE, 1992. – Vol. 4. – P. 578-581.
17. Marra S., Iachino M.A., Morabito F.C. Tanh-like activation function implementation for high-performance digital neural systems // 2006 Ph. D. Research in Microelectronics and Electronics. – IEEE, 2006. – P. 237-240.
18. Cortes C., Mohri M., Rostamizadeh A. L2 regularization for learning kernels // arXiv preprint arXiv:1205.2653. – 2012.
19. Тихонов А.Н. О некорректных задачах линейной алгебры и устойчивом методе их решения // ДАН СССР. – 1965. – Т. 163, № 3. – С. 591-594.
20. Kingma D.P., Ba J. Adam: A method for stochastic optimization // arXiv preprint arXiv:1412.6980. – 2014.
21. Wang J. et al. Design flow of accelerating hybrid extremely low bit-width neural network in embedded FPGA // 2018 28th International Conference on Field Programmable Logic and Applications (FPL). – IEEE, 2018. – P. 163-1636.

## REFERENCES

1. *Gonzalez R., Vuds R.* TSifrovaya obrabotka izobrazheniy [Digital image processing]: transl. from engl. Moscow: Tekhnosfera, 2005, 1073 p.
2. *Anderson T.W., Goodman L.A.* Statistical inference about Markov chains, *The Annals of Mathematical Statistics*, 1957, pp. 89-110.
3. *Vagis A.A.* Protsedura raspoznavaniya na bayesovskikh setyakh, *Komp'yuternaya matematika*, 2010, No. 2, pp. 124-130.
4. *LeCun Y. et al.* Gradient-based learning applied to document recognition, *Proceedings of the IEEE*, 1998, Vol. 86, Issue 11, pp. 2278-2324.
5. *Gu J. et al.* Recent advances in convolutional neural networks, *Pattern Recognition*, 2018, Vol. 77, pp. 354-377.
6. *Srinivas S. et al.* A taxonomy of deep convolutional neural nets for computer vision, *Frontiers in Robotics and AI*, 2016, Vol. 2, pp. 36.
7. *Glorot X., Bengio Y.* Understanding the difficulty of training deep feedforward neural networks, *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, 2010, pp. 249-256.
8. *He K. et al.* Delving deep into rectifiers: Surpassing human-level performance on imagenet classification, *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 1026-1034.
9. *LeCun Y. et al.* Backpropagation applied to handwritten zip code recognition, *Neural computation*, 1989, Vol. 1, Issue 4, pp. 541-551.
10. *Boureau Y.L. et al.* Learning mid-level features for recognition, *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. IEEE, 2010, pp. 2559-2566.
11. *Scherer D., Müller A., Behnke S.* Evaluation of pooling operations in convolutional architectures for object recognition, *International conference on artificial neural networks*. Springer, Berlin, Heidelberg, 2010, pp. 92-101.
12. *Boureau Y.L., Ponce J., LeCun Y.* A theoretical analysis of feature pooling in visual recognition, *Proceedings of the 27th international conference on machine learning (ICML-10)*, 2010, pp. 111-118.
13. *Krizhevsky A., Sutskever I., Hinton G.E.* Imagenet classification with deep convolutional neural networks, *Advances in neural information processing systems*, 2012, pp. 1097-1105.
14. *Maas A.L., Hannun A.Y., Ng A.Y.* Rectifier nonlinearities improve neural network acoustic models, *Proc. icml.*, 2013, Vol. 30, Issue No. 1, pp. 3.
15. *Zeiler M.D. et al.* On rectified linear units for speech processing, *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, 2013, pp. 3517-3521.
16. *Kalman B.L., Kwasy S.C.* Why tanh: choosing a sigmoidal function, *[Proceedings 1992] IJCNN International Joint Conference on Neural Networks*. IEEE, 1992, Vol. 4, pp. 578-581.
17. *Marra S., Iachino M.A., Morabito F.C.* Tanh-like activation function implementation for high-performance digital neural systems, *2006 Ph. D. Research in Microelectronics and Electronics*. IEEE, 2006, pp. 237-240.
18. *Cortes C., Mohri M., Rostamizadeh A.* L2 regularization for learning kernels, *arXiv preprint arXiv:1205.2653*, 2012.
19. *Tikhonov A.N.* O nekorrektnykh zadachakh lineynoy algebry i ustoychivom metode ikh resheniya [On ill – posed linear algebra problems and a stable method for solving them], *DAN SSSR [Reports of the Academy of Sciences of USSR]*, 1965, Vol. 163, No. 3, pp. 591-594.
20. *Kingma D.P., Ba J.* Adam: A method for stochastic optimization, *arXiv preprint arXiv:1412.6980*, 2014.
21. *Wang J. et al.* Design flow of accelerating hybrid extremely low bit-width neural network in embedded FPGA, *2018 28th International Conference on Field Programmable Logic and Applications (FPL)*. IEEE, 2018, pp. 163-1636.

Статью рекомендовал к опубликованию д.т.н. А.В. Ковалев.

**Деркачев Владимир Александрович** – Южный федеральный университет; e-mail: vderkachev@sfedu.ru; 347922, г. Таганрог, ул. Шевченко, 2; тел.: +79614154733; кафедра радиотехнических и телекоммуникационных систем; ассистент.

**Бахчевников Валентин Владимирович** – e-mail: bahchevnikov@sfedu.ru; 347928, г. Таганрог, пер. Некрасовский; 44; тел.: +79518289271; кафедра радиотехнических и телекоммуникационных систем; ассистент.

**Бакуменко Алексей Николаевич** – e-mail: baku@sfedu.ru; 347900, г. Таганрог, ул. Петровская, 81; тел.: +79886031853; Инжиниринговый центр приборостроения радио и микроэлектроники; инженер.

**Derkachev Vladimir Aleksandrovich** – Southern Federal University; e-mail: vderkachev@sfedu.ru; 2, Shevchenko street, Taganrog, 347922, Russia; phone +79614154733; the department of radio engineering & telecommunication systems; assistant.

**Bakhchevnikov Valentin Vladimirovich** – e-mail: bahchevnikov@sfedu.ru; 44, Nekrasovsky lane, Taganrog, 347928, Russia; phone: +79518289271; the department of radio engineering & telecommunication systems; assistant.

**Bakumenko Alexey Nikolaevich** – e-mail: baku@sfedu.ru; 81, Petrovskaya street, Taganrog, 347900, Russia; phone: +79886031853; Engineering center of instrument making, radio- and microelectronics; engineer.

УДК 0004.912

DOI 10.18522/2311-3103-2020-4-39-48

**Е.М. Герасименко, В.В. Степенко**

### **ИНТЕЛЛЕКТУАЛЬНЫЙ МЕТОД ИЗВЛЕЧЕНИЯ ЗНАНИЙ НА ОСНОВЕ ОПРЕДЕЛЕНИЯ ТОНАЛЬНОСТИ ОТЗЫВОВ\***

*В этой работе исследуется влияние возраста и пола при анализе тональности отзывов, поскольку эти данные могут помочь ретейлерам электронной коммерции увеличить продажи, ориентируясь на определенные демографические группы, а также увеличить удовлетворение потребностей людей разных возрастных и гендерных групп. Используемый набор данных сформирован путем сбора отзывов о книгах. Был создан вопросник, содержащий информацию о предпочтениях книжных носителей (мнения пользователей об электронных книгах, книгах в мягкой и твердой обложках, изображениях и аудиокнигах), а также данные о возрастной группе и гендерной принадлежности. Помимо этого, вопросник также содержит информацию о положительном либо отрицательном мнении касательно предпочтений, которая послужила основой достоверности для классификаторов. В результате, было получено 900 анкет, которые были разделены на группы по половому признаку и возрасту. Каждая конкретная группа данных была разделена на обучающую и тестовую. Были проанализированы сегментированные данные на предмет настроений в зависимости от каждой возрастной группы и пола. Возрастная группа «старше 50 лет» продемонстрировала лучшие результаты по сравнению со всеми другими возрастными группами во всех классификаторах; данные в женской группе показали более высокую точность по сравнению с данными из групп без информации о гендерной принадлежности. Высокие результаты, показанные этими группами, показывают, что подходы к анализу тональности способны предсказать настроения в этих группах лучше, чем в других. Анализ тональности проводился с использованием различных подходов машинного обучения (ML), включая максимальную энтропию, метод опорных векторов, сверточную нейронную сеть и долгую краткосрочную память.*

*Анализ тональности; машинное обучение; сверточная нейронная сеть; долгая краткосрочная память.*

---

\* Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 19-07-00099 а.