**ИСПОЛЬЗОВАНИЕ ГЛУБОКОГО ОБУЧЕНИЯ
ПРИ ДИАГНОСТИРОВАНИИ ПНЕВМОНИИ ПО РЕНТГЕНОВСКИМ СНИМКАМ**

А. С. РАСКОПИНА, В. В. БОЖЕНКО, Т. М. ТАТАРНИКОВА*

Санкт-Петербургский государственный университет аэрокосмического приборостроения,
Санкт-Петербург, Россия,
* tm-tatarn@yandex.ru

Аннотация. С развитием методов машинного обучения для распознавания образов открываются новые возможности в области медицинской диагностики. Уровень точности и надежности, достигнутый нейронными сетями, позволяет снизить риск ложных срабатываний и ошибок. Для задачи диагностирования пневмонии по рентгеновским изображениям проведено сравнение различных алгоритмов машинного обучения, таких как метод опорных векторов, k -ближайших соседей, сверточные нейронные сети. Обсуждаются преимущества использования данных методов в медицинской диагностике. Алгоритмы машинного обучения доведены до программной реализации, для каждого из них экспериментальным путем выбраны параметры обучения. Для сравнения методов использовалась стандартная метрика — точность (accuarcy), также методы сравнивались по времени обучения. Эксперименты проведены на реальных рентгеновских снимках больных пневмонией пациентов. Результаты экспериментов показали, что глубокие нейронные сети демонстрируют лучшую точность, по сравнению с традиционными методами машинного обучения, что подтверждает эффективность их потенциального использования для диагностики и лечения данного заболевания.

Ключевые слова: рентгеновское изображение, нейронные сети, машинное обучение, диагностика, метод опорных векторов, k -ближайших соседей, сверточные нейронные сети, глубокое обучение

Ссылка для цитирования: Раскопина А. С., Боженко В. В., Татарникова Т. М. Использование глубокого обучения при диагностировании пневмонии по рентгеновским снимкам // Изв. вузов. Приборостроение. 2024. Т. 67, № 4. С. 315—320. DOI: 10.17586/0021-3454-2024-67-4-315-320.

USING DEEP LEARNING IN PNEUMONIA DIAGNOSIS FROM X-RAYS PATTERNS

A. S. Raskopina, V. V. Bozhenko, T. M. Tatarnikova*

St. Petersburg State University of Aerospace Instrumentation, St. Petersburg, Russia
* tm-tatarn@yandex.ru

Abstract. With the development of neural networks, new effective solutions in the field of medical diagnosis are opening up. The level of accuracy and reliability achieved by neural networks reduces the risk of false positives and diagnostic errors. In the task of diagnosing pneumonia from X-ray images, various machine learning algorithms such as the support vector machine (SVM), K-nearest neighbors (KNN), convolutional neural networks (CNN) are compared. The advantages of these methods in the task of medical diagnostics are discussed. Machine learning algorithms are brought to software implementation, and for each of them training parameters are selected experimentally. To compare the methods, a standard metric, accuracy, was used, and the methods are also compared by training time. The corresponding experiments are conducted on real data from X-ray images of patients with pneumonia. The experimental results demonstrate better accuracy of deep neural networks compared to traditional machine learning methods, which confirms the effectiveness of their potential use for the diagnosis and treatment of this disease.

Keywords: neural networks, machine learning, diagnostics, support vector method, k-nearest neighbors, convolutional neural networks, deep learning

For citation: Raskopina A. S., Bozhenko V. V., Tatarnikova T. M. Using deep learning in pneumonia diagnosis from x-rays patterns. *Journal of Instrument Engineering*. 2024. Vol. 67, N 4. P. 315—320 (in Russian). DOI: 10.17586/0021-3454-2024-67-4-315-320.

Применение искусственного интеллекта (ИИ) в медицине является одним из направлений, которое поддерживает врачей-специалистов в принятии решений. Безусловно, и раньше

технологии ИИ, например экспертные системы и алгоритмы классификации, применялись в постановке диагноза, но в последние десять лет задача распознавания образов получила наибольшую популярность благодаря развитию моделей глубокого обучения. Эти модели позволяют компьютерам анализировать и интерпретировать медицинские изображения, такие как рентгеновские, КТ- и МРТ-снимки [1—3].

Преимущества от применения технологий распознавания образов очевидны:

— повышение оперативности диагностики — медицинские изображения анализируются в режиме реального времени, что сокращает время ожидания результатов и, соответственно, время начала лечения;

— автоматизированный анализ изображений снижает риск ошибок, которые могут возникнуть при интерпретации снимков человеком;

— точность диагностики повышается, что особенно важно при выявлении редких или с трудом поддающихся анализу заболеваний.

Исследование функций и особенностей работы современных специализированных систем анализа и обработки медицинских изображений показало, что эти системы обладают рядом недостатков. Основным недостатком является то, что большинство систем содержит широкий набор методов анализа и обработки изображений без привязки метода к цели диагностирования [4, 5]. В связи с этим выявлены следующие проблемы: невозможно гарантированно выбрать оптимальный метод или их комбинацию для обработки изображений, а осуществить перебор всех имеющихся в распоряжении исследователя методов для достижения наилучшего результата обработки не всегда возможно.

Поэтому задачу анализа и обработки медицинских изображений можно решать в два этапа: во-первых, накопление размеченной врачами-специалистами обучающей выборки и, во-вторых, выбор модели машинного обучения с требуемыми показателями точности результата и ее построение [6, 7].

Одним из эффективных и приоритетных в медицинской практике способов диагностики является рентгенография: анализ и выявление признаков заболевания по флюорографическим снимкам. Это технически простой и наиболее доступный для пациентов метод предварительной диагностики заболевания, в том числе отслеживания заболевания на ранней стадии.

В настоящей статье решается задача распознавания рентгеновских снимков для диагностирования пневмонии с применением модели глубокого обучения.

Несмотря на все преимущества метода рентгенографии, диагностика заболевания по флюорографическим снимкам является нетривиальной задачей. Сложность задачи заключается, как правило, в низком разрешении изображений и наличии различного рода шумов, таких как засветки, посторонние предметы и т.д. Дополнительные трудности вносит наличие на снимках естественных анатомических структур, например, костей, других перекрывающих органов, которые зачастую скрывают аномалии легочной ткани. Выделение и последующая фильтрация подобного рода структур с учетом разметки, выполненной врачом-рентгенологом, является первым важным этапом решения задачи диагностирования заболевания. Этот этап не является частью настоящей работы — нами использована готовая обучающая выборка из открытого источника, являющаяся площадкой для соревнований по машинному обучению.* Выборка включает 5863 рентгеновских снимка грудной клетки, которые были заранее промаркированы как „здоровые легкие“ и „диагноз пневмония“.

Существует несколько методов классификации, которые можно применять, в том числе для распознавания образов. В настоящей работе выбраны методы опорных векторов, k-ближайших соседей и сверточные нейронные сети.

* <https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia>.

Метод опорных векторов (SVM) при решении задач классификации обладает определенными преимуществами: меньшее время прогнозирования, а точность классификатора обнаружения цели может быть гарантирована глобальным оптимальным решением. Недостаток метода заключается в линейном росте времени распознавания с увеличением объема данных [8]. Для сравнительного анализа с другими методами распознавания образов метод SVM выбран по нескольким причинам:

- 1) исходные данные — это рентгеновские изображения, которые могут быть представлены высокоразмерными векторами;
- 2) SVM обладает хорошей способностью к обобщению, следовательно, он может давать хорошие результаты на новых данных, не участвовавших в обучении;
- 3) метод достаточно устойчив к шуму в данных.

Метод *k*-ближайших соседей (KNN) — объекту присваивается класс, который наиболее распространен среди *k* соседей данного элемента, классы которых уже известны [9]. Причины, по которым этот алгоритм выбран для решения задачи диагностирования пневмонии:

- 1) алгоритм прост в интерпретации;
- 2) обучение проходит без учителя: предварительно модель не обучается, следовательно, сокращается время решения задачи [10].

Известный недостаток заключается в том, что метод не может правильно предсказать ни одно решение, если оно не принадлежит ни к одному из predetermined классов в наборе обучающих данных.

Сверточные нейронные сети (CNN) представляют собой многослойный стек из нейронов, благодаря чему CNN можно собирать как конструктор, добавляя слой за слоем. Каждый уровень CNN преобразует изображение на своем уровне абстракции [11]. CNN обычно включает сверточные слои, слои объединения и полносвязные слои. Слой свертки CONV является главным слоем CNN. Его назначение — выделить признаки на входном изображении и сформировать карту признаков. Первый сверточный уровень может изучать базовые элементы, такие как ребра, а второй — шаблоны, состоящие из базовых элементов, проанализированных на предыдущем уровне. Это позволяет CNN эффективно изучать все более сложные и абстрактные визуальные концепции. Сверточный слой сопровождается слоями объединения POOL. Объединяющие слои упрощают информацию, собираемую сверточным слоем, и создают сжатую версию информации, содержащейся в них. Полносвязный слой FC объединяет все нейроны предыдущего и последующего слоев и помогает сопоставить данные на входе с меткой класса, как в нейронной сети прямого распространения [12]. В качестве функции активации наиболее часто используются ReLU, Sigmoid, Tanh, LeakyReLU. Таким образом, использование CNN в задаче диагностирования по рентгеновским снимкам обосновано их способностью эффективно выделять признаки из изображений, обучаться на больших объемах данных, а также автоматизировать процесс [13].

В табл. 1 представлена общая структура построенной нейронной сети. В этой сети применяются пять слоев для обработки входных данных: сверточный, пулинговый и три полносвязных.

Таблица 1

Общая структура сверточной нейронной сети

Слой	Выходная форма данных	Число обучаемых параметров
CONV	(None, 148, 148, 32)	896
POOL	(None, 74, 74, 32)	0
FC	(None, 175232)	0
dense_2 (Dense)	(None, 64)	11214912
dense_3 (Dense)	(None, 1)	65

Сверточный слой CONV применяется для выделения различных признаков изображения на входных данных. Размерность выходных данных этого слоя составляет (None, 148, 148, 32), где значение „None“ указывает на то, что модель может принимать пакеты данных

любого размера. После прохождения через этот слой изображение будет иметь размер 148×148 пикселей, а также 32 канала. Каждый из этих 32 каналов представляет собой набор определенных признаков, выделенных изображением в процессе свертки. Общее число параметров на этом слое 896, что отражает количество настраиваемых коэффициентов и смещений в сверточном фильтре, который используется для выделения этих признаков.

Затем данные проходят через пулинговый слой POOL, который снижает размерность данных вдвое. Пулинговый слой не имеет обучаемых параметров, поэтому число параметров на этом слое равно нулю.

Далее данные подаются на полносвязный слой FC, который преобразовывает их в одномерный вектор размером 175 232. Этот слой не имеет обучаемых параметров, поэтому число параметров также равно нулю.

Последние два полносвязных слоя — dense_2 и dense_3 (Dense). Эти слои принимают на вход выходные сигналы всех нейронов предыдущего слоя, умножают их на свои весовые коэффициенты и применяют активационную функцию.

Слой dense_2 использует функцию активации ReLU и имеет 64 нейрона и 11 214 912 параметров, что указывает на количество обучаемых весовых коэффициентов и смещений между нейронами. А слой dense_3 использует функцию активации Sigmoid и имеет 1 нейрон и 65 параметров, включая веса и смещения. Модель после прохождения выходного слоя dense_3 выдает единственное численное значение на основе входных данных.

Точность классификатора оценивалась метрикой ассигасу, показывающей долю правильных решений:

$$\text{accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}},$$

где TP (истинно-положительное решение) — результат предсказания моделью положительного класса объекта совпал с реальной меткой класса; FP (ложноположительное решение) — ошибка 1-го рода, модель ошибочно отнесла объект, который на самом деле относится к отрицательному классу, к положительному; FN (ложноотрицательное решение) — ошибка 2-го рода, модель ошибочно классифицировала отрицательный объект как положительный; TN (истинно-отрицательное решение) — модель верно классифицировала объект как отрицательный.

Все методы классификации доведены до программной реализации. Параметры обучения выбраны экспериментальным путем:

— для обучения SVM использован класс SVC, который осуществляет поддержку мультиклассов по схеме „один против одного“; ядерная функция — радиально-базисная; степень полинома 3;

— при обучении методом KNN лучшие результаты получились для пяти ближайших соседей;

— в модели обучения CNN использованы функция активации ReLU и алгоритм оптимизации Adam, который позволил построить архитектуру нейронной сети, удовлетворяющей максимально возможному значению метрики ассигасу.

Результаты, которые получены в ходе эксперимента, приведены в табл. 2. Из них можно сделать вывод, что с задачей распознавания лучше других методов справилась сверточная нейронная сеть, худшие результаты получены методом опорных векторов. Причиной этому может быть небольшой объем данных для обучения. Также время обучения у сверточной нейронной сети ожидаемо оказалось самым продолжительным.

Таблица 2

Результаты эксперимента		
Модель	Точность (тестовая выборка), о.е.	Время обучения, с
SVM	0,74	489
KNN	0,63	17
CNN	0,89	987

Результаты проведенного исследования показали, что глубокие нейронные сети обеспечивают более высокую точность, чем традиционные методы машинного обучения, в задаче выявления пневмонии на рентгеновских снимках.

СПИСОК ЛИТЕРАТУРЫ

1. Zhang Zh. Introduction to machine learning: K-nearest neighbors // *Annals of Translational Medicine*. 2016. Vol. 4, N 11. P. 218—225. DOI: 10.21037/atm.2016.03.37.
2. Bozhenko V. V. and Tatarnikova T. M. Application of Data Preprocessing in Medical Research // *2023 Wave Electronics and its Application in Information and Telecommunication Systems (WECONF)*. St. Petersburg, 2023. P. 1—4. DOI: 10.1109/WECONF57201.2023.10148004.
3. Razzak M. I., Imran M., Xu G. Big data analytics for preventive medicine // *Neural Computing & Applications*. 2020. Vol. 32. P. 4417—4451.
4. Васильченко В. А., Бурковский В. Л., Данилов А. Д. Алгоритмизация процесса распознавания состояний физиологических объектов на основе специальных рентгеновских изображений // *Компьютерная оптика*. 2019. Т. 43, № 2. С. 296—303.
5. Максимова Е. И., Хаустов П. А. Алгоритм обнаружения образований в легких человека на снимках компьютерного томографа с использованием искусственной нейронной сети // *Фундаментальные исследования*. 2016. № 4–2. С. 290—294.
6. Комков А. А., Мазаев В. П., Рязанова С. В., Самочатов Д. Н., Базаева Е. В. Основные направления развития искусственного интеллекта в медицине // *Научное обозрение. Медицинские науки*. 2020. № 5. С. 33—40.
7. Jaapar R. M. Q. R., Mansor M. A. Convolutional Neural Network Model in Machine Learning Methods and Computer Vision for Image Recognition: A Review // *J. of Applied Sciences Research*. 2018. Vol. 14, N 6. P. 23—27. DOI: 10.22587/jasr.2018.14.6.5.
8. Awad M., Khanna R. Support Vector Machines for Classification // *Efficient Learning Machines*. 2015. P. 39—66. DOI:10.1007/978-1-4302-5990-9_3.
9. Tatarnikova T. M. and Mokretsov N. S. Wireless Sensor Network Clustering Model // *XXVI Intern. Conf. on Soft Computing and Measurements (SCM)*. St. Petersburg, 2023. P. 240—243. DOI: 10.1109/SCM58628.2023.10159119.
10. Paramasivam K., Sindha M. M. R., Balakrishnan S. B. KNN-Based Machine Learning Classifier Used on Deep Learned Spatial Motion Features for Human Action Recognition // *Entropy*. 2023. Vol. 25, N 844. P. 1—15. DOI: 10.3390/e25060844.
11. Alzubaidi L., Zhang J., Humaidi A. J., Al-Dujaili A., Duan Y., Al-Shamma O., Santamaria J., Fadhel M. A., Al-Amidie M., and Farhan L. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions // *J. of Big Data*. 2021. Vol. 8, N 53. P. 1—74. DOI: 10.1186/s40537-021-00444-8.
12. Le Cun Y., Bengio Y., Hinton G. Deep learning // *Nature*. 2015. Vol. 521, N 7553. P. 436—444. DOI:10.1038/nature14539
13. Akcay S., Kundegorski M. E., Willcocks C. G. and Breckon T. P. Using Deep Convolutional Neural Network Architectures for Object Classification and Detection within X-Ray Baggage Security Imagery // *IEEE Transact. on Information Forensics and Security*. 2018. Vol. 13, N 9. P. 2203—2215.

Сведения об авторах

- Анастасия Сергеевна Раскопина** — студентка; Санкт-Петербургский государственный университет аэрокосмического приборостроения, кафедра прикладной информатики; E-mail: raskopina.anastasia@yandex.ru
- Виктория Вячеславовна Боженко** — Санкт-Петербургский государственный университет аэрокосмического приборостроения, кафедра прикладной информатики; старший преподаватель; E-mail: vibozenko@yandex.ru
- Татьяна Михайловна Татарникова** — д-р техн. наук, профессор; Санкт-Петербургский государственный университет аэрокосмического приборостроения, кафедра прикладной информатики; профессор; E-mail: tm-tatarn@yandex.ru

Поступила в редакцию 04.12.23; одобрена после рецензирования 12.12.23; принята к публикации 08.02.24.

REFERENCES

1. Zhang Zh. *Annals of Translational Medicine*, 2016, no. 11(4), pp. 218–225, DOI: 10.21037/atm.2016.03.37.
2. Bozhenko V.V. and Tatarnikova T.M. *Wave Electronics and its Application in Information and Telecommunication Systems (WECONF)*, St. Petersburg, 2023, pp. 1–4, DOI: 10.1109/WECONF57201.2023.10148004.
3. Razzak M.I., Imran M., Xu G. *Neural Computing & Applications*, 2020, vol. 32, pp. 4417–4451.
4. Vasilchenko V.A., Burkovsky V.L., Danilov A.D. *Computer optics*, 2019, no. 2(43), pp. 296–303. (in Russ.)
5. Maksimova E.I., Khaustov P.A. *Fundamental research*, 2016, no. 4–2, pp. 290–294. (in Russ.)
6. Komkov A.A., Mazaev V.P., Ryazanova S.V., Samochatov D.N., Bazaeva E.V. *Science review. Medical Sciences*, 2020, no. 5, pp. 33–40, URL: <https://science-medicine.ru/ru/article/view?id=1141>. (in Russ.)
7. Jaapar R.M.Q.R., Mansor M.A. *Journal of Applied Sciences Research*, 2018, no. 6(14), pp. 23–27, DOI: 10.22587/jasr.2018.14.6.5.
8. Awad M., Khanna R. *Efficient Learning Machines*, 2015, pp. 39–66, DOI:10.1007/978-1-4302-5990-9_3.
9. Tatarnikova T.M. and Mokretsov N.S. *XXVI International Conference on Soft Computing and Measurements (SCM)*, St. Petersburg, 2023, pp. 240–243, DOI: 10.1109/SCM58628.2023.10159119.
10. Paramasivam K., Sindha M.M.R., Balakrishnan S.B. *Entropy*, 2023, no. 844(25), pp. 1–15, DOI: 10.3390/e25060844.
11. Alzubaidi L., Zhang J., Humaidi A.J., Al-Dujaili A., Duan Y., Al-Shamma O., Santamaría J., Fadhel M.A., Al-Amidie M., and Farhan L. *Journal of Big Data*, 2021, no. 53(8), pp. 1–74, DOI:10.1186/s40537-021-00444-8.
12. LeCun Y., Bengio Y., Hinton G. *Nature*, 2015, no. 7553(521), pp. 436–444, DOI:10.1038/nature14539.
13. Akcay S., Kundegorski M.E., Willcocks C.G., and Breckon T.P. *IEEE Transactions on Information Forensics and Security*, 2018, no. 9(13), pp. 2203–2215.

Data on authors

- Anastasia S. Raskopina** — Student; St. Petersburg State University of Aerospace Instrumentation, Department of Applied Informatics; E-mail: raskopina.anastasia@yandex.ru
- Viktoriya V. Bozhenko** — St. Petersburg State University of Aerospace Instrumentation, Department of Applied Informatics; Senior Lecturer; E-mail: vibozenko@yandex.ru
- Tatiana M. Tatarnikova** — Dr. Sci., Professor; St. Petersburg State University of Aerospace Instrumentation, Department of Applied Informatics; Professor; E-mail: tm-tatarn@yandex.ru

Received 04.12.23; approved after reviewing 12.12.23; accepted for publication 08.02.24.