



Применение нейронных сетей для распознавания хитиной зооид на изображениях

Р. Р. Якупов¹, К. Б. Гусманова²

¹ Институт геологии УФИЦ РАН, Уфа; jakupov@ufaras.ru

² Уфимский университет науки и технологий, Уфа; kamilla@ufagen.ru

Палеонтологическое определение микрофауны с помощью автоматизированного распознавания изображений представляет собой инновационное приложение существующих программных методов анализа и классификации на основе технологий компьютерного зрения и машинного обучения. Разработка программного обеспечения, способного распознавать хитиной зооиды на снимках, упростит и ускорит обработку больших массивов данных по микрофоссилиям. Использование нейронных сетей для анализа изображений возможно и для других групп органических остатков.

Хитиной зооиды имеют ряд преимуществ, позволяющих поэтапно оценить применимость технологии автоматизированного распознавания изображений для биостратиграфических задач по сравнению с другими группами микрофоссилий. Искусственная палеонтологическая классификация хитиной зооиды построена на четких морфологических признаках и поддается формализации. На первом этапе решения задач распознавания проводилось построение функции классификации, прогнозирующей, к какому классу принадлежит ископаемое по входному вектору признаков: «хитиной зооиды» либо «нехитиной зооиды». Разработанная модель алгоритма распознавания хитиной зооиды показала высокую степень точности (более 98 %).

Ключевые слова: хитиной зооиды, распознавание изображений, нейронные сети, машинное обучение, палеонтология

Application of neural networks for chitinozoans recognition in images

R. R. Yakupov¹, K. B. Gusmanova²

¹ Institute of Geology UFRC RAS, Ufa

² Ufa University of Science and Technology, Ufa

Paleontological identification of microfauna using automated image recognition represents an innovative practical application of existing software methods for analysis and classification based on computer vision and machine learning technologies. The development of software capable of recognition of Chitinozoans in images will simplify and speed up the processing of large amounts of microfossil data. The use of neural networks for image analysis is also possible for other groups of paleoorganisms.

Chitinozoans have a number of advantages that allow a step-by-step assessment of the applicability of automated image recognition technology for biostratigraphic problems compared to other groups of microfossils. The artificial paleontological classification of Chitinozoa is based on clear morphological characteristics and can be formalized. At the first stage of solving recognition problems, a classification function was constructed that predicts class, the fossil belongs to, based on the input feature vector – either “chitinozoa” or “non-chitinozoa”. The developed model of the Chitinozoans recognition algorithm showed a high degree of accuracy (more than 98 %).

Keywords: chitinozoan, image recognition, neural networks, machine learning, paleontology

Введение

Хитиной зооиды — вымершая группа микроорганизмов неизвестного систематического положения, используется для определения возраста вмещающих осадочных отложений. Они представляют собой мелкие органостенные микрофоссилии от 40 до 1500 мкм с радиальной симметрией. В биостратиграфии ордовикских и силурийских морских отложений хитиной зооиды используют наравне с конодонтами и граптолитами.

Исследование микрофоссилий начинается с дезинтеграции породы и выделения нерастворимого органического остатка. Из полученного остатка достаются единичные микрофоссилии, в дальнейшем проводится их изучение с помощью световой и электронной микроскопии. В результате накапливается большое количество фотографий, требующих индивидуальной

обработки в графических редакторах и последующего анализа. Совместно с хитиной зооидами встречаются и другие группы микро- и макрофоссилий. Требуется ручной отбор нужного материала для фотографирования, определение и палеонтологическое описание везикул (единичных организмов) хитиной зооиды на основе полученных изображений микропроблематики.

Несмотря на то, что хитиной зооиды являются признанной группой для определения возраста отложений силура и ордовика, они считаются проблематичными — неизвестна их точная биологическая принадлежность к царству животных или растений (Paris, Nolvak, 1999). Искусственная биологическая классификация хитиной зооиды построена на морфологических признаках — строении пробки/оперкулула, форме везикулы и морфологии стенки (Paris, 2006). Несомненным плюсом для использования хитиной зооиды в качестве модельной груп-

Для цитирования: Якупов Р. Р., Гусманова К. Б. Применение нейронных сетей для распознавания хитиной зооиды на изображениях // Вестник геонаук. 2024. 9(357). С. 35–39. DOI: 10.19110/geov.2024.9.5

For citation: Yakupov R. R., Gusmanova K. B. Application of neural networks for chitinozoans recognition in images. Vestnik of Geosciences, 2024, 9(357), pp. 35–39, doi: 10.19110/geov.2024.9.5

пы органики на данный момент является их классификационная иерархия, где виды объединяются в 56 родов, которые относятся всего к трем семействам и объединяются в два отряда (Paris, Nolvak, 1999). Возможно последовательное решение задачи определения изображений от семейства и подсемейства к роду и виду — при видовом описании палеонтологи следуют данному алгоритму. Еще одним аргументом в пользу использования хитинозой является то, что даже поврежденные и частично разрушенные везикулы из-за радиальной симметрии могут виртуально достраиваться до полной формы, если сохранилась некоторая часть или фрагменты везикулы по оси симметрии. Для конодонтов, граптолитов и других важных для био-стратиграфии ордовика и силура групп, не имеющих такой особенности, подобный подход (достраивание формы) невозможен. Решение данной проблемы является нетривиальной задачей и требует отдельного этапа работ.

Современные программные и технические возможности позволяют нейронной сети анализировать и классифицировать изображения в разных областях человеческой деятельности, например в медицине (Гусев, 2017; Gardner et al., 2023), в литологии (Журавлев, Груздев, 2024). Опубликованы работы по распознаванию изображений палеонтологических объектов: фораминифер (Gorur et al., 2023), радиолярий (Ruikar, Agrawal, 2019), спор и пыльцы (White, 2020) и раковинной фауны (Marmo et al., 2006). Автоматизация подготовки, обработки и распознавания изображений палеонтологических объектов способствует существенному сокращению рутинных ручных этапов обработки информации. Увеличение производительности труда критически важно на фоне востребованности узких профильных специалистов-палеонтологов, а инновационные подходы открывают новые горизонты для изучения прошлого Земли и применения новейших методов в фундаментальных науках.

Цель данной работы заключается в создании системы для автоматизации процесса классификации хитинозой на основе их морфологических признаков.

Материалы и методы

Разработка программного обеспечения, способного автоматизировать процесс распознавания хитинозой на изображениях, проходила в несколько этапов. На первом этапе набирался обучающий и тестовый набор данных, включающий 2955 СЭМ-изображений (СЭМ — сканирующий электронный микроскоп) хитинозой. Дополнительно было подготовлено 3484 фотографии других микрофоссилий. В ходе эксперимента было выяснено, что изображений для проверки и обучения нейронной сети недостаточно из-за сильного дисбаланса классов (недостаточно изображений в каждой категории и их количество отличается в два-три раза). Набор данных, необходимый для обучения нейронной сети, расширили, используя существующие изображения, для чего в часть изображений были внесены изменения: некоторые были повернуты на определенный угол, часть изображений кадрировалась по случайным параметрам. Созданные новые изображения вошли в ту же базу данных. Вышеописанные действия называются «аугментацией данных». Набор данных дополнили также изображениями из других источников, до 5000 в каждой категории.

После объединения данных были удалены изображения, содержащие ложные признаки, в частности, среди изображений хитинозой присутствовали фотографии с белым фоном, которые модель приняла за признак, так как у «нехитинозой» белого фона не было. Были убраны рамки, линейки и другие элементы, мешающие распознаванию. Все изображения были приведены к градациям серого и лишены цветовой информации (рис. 1).

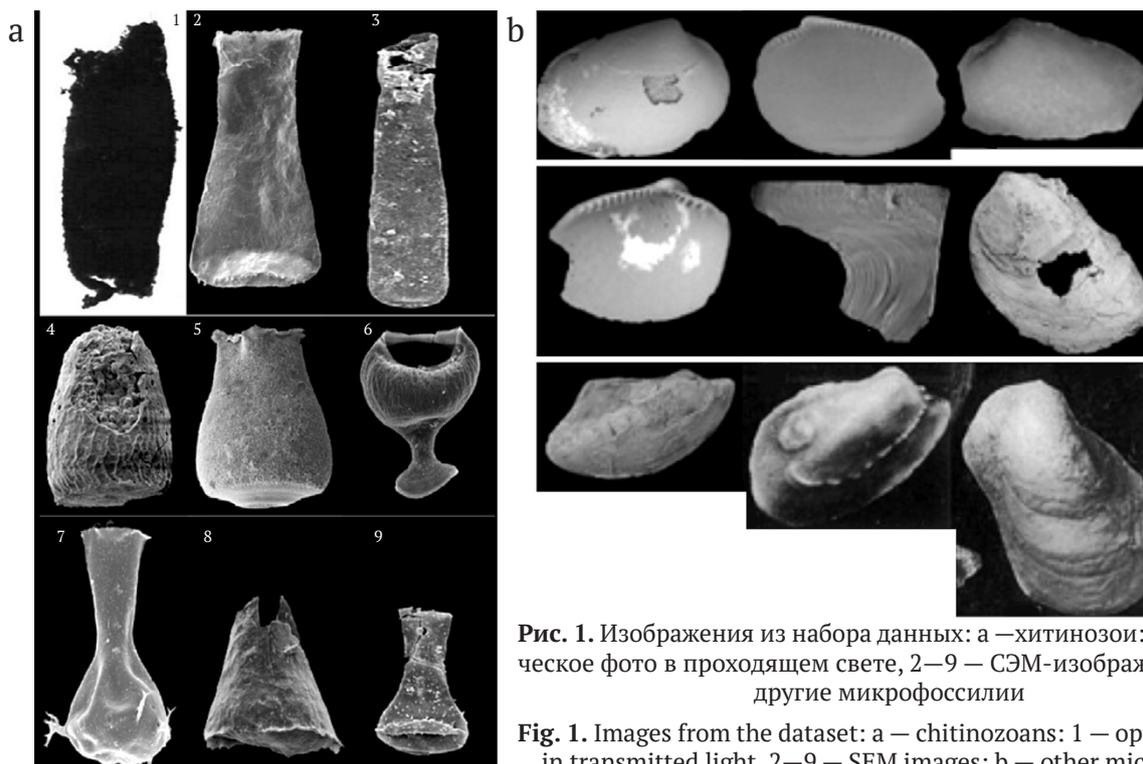


Рис. 1. Изображения из набора данных: а — хитинозои: 1 — оптическое фото в проходящем свете, 2–9 — СЭМ-изображения; б — другие микрофоссилии

Fig. 1. Images from the dataset: a — chitinozoans: 1 — optical photo in transmitted light, 2–9 — SEM images; b — other microfossils



Математическая постановка задачи выражалась в решении задачи бинарной классификации и построении функции классификации, предсказывающей, к какому классу принадлежит изображение ископаемого по входному вектору признаков — «хитинозоа» либо «нехитинозоа».

Для системы распознавания изображений была выбрана архитектура свёрточной нейронной сети (CNN) (рис. 2).

Выбор свёрточной нейронной сети для распознавания изображений обусловлен её способностью эффективно обрабатывать визуальные данные и сравнительно высокой точностью в классификации изображений (Tan, Quoc, 2021).

Была выбрана модель EfficientNetV2B0, так как она принадлежит к семейству EfficientNetV2, улучшающему исходную серию EfficientNet. EfficientNetV2 — семейство более мелких и быстрых нейронных сетей для распознавания изображений. По сравнению с Effi-

cientNet и более поздними разработками, EfficientNetV2 обучается до 11 раз быстрее, будучи при этом в 6.8 раза меньше (Tan, Quoc, 2021).

Stem-блок — начальная часть модели, включающая несколько сверток и активирований, оптимизированных для быстрого извлечения признаков (рис. 3).

Сетевые стадии включают последовательность блоков MBConv и Fused-MBConv. Каждый блок состоит из свёрточного слоя, за которым следует активация, и завершается блок нейронной сети слоем пакетной нормализации.

В отличие от обычных MBConv, Fused-MBConv-блоки совмещают свёртки и нормализацию в одном шаге, что помогает увеличить эффективность модели.

В конце сети используется глобальное усреднение для снижения размерности признаков перед передачей в окончательный полносвязный слой для предсказания классов.

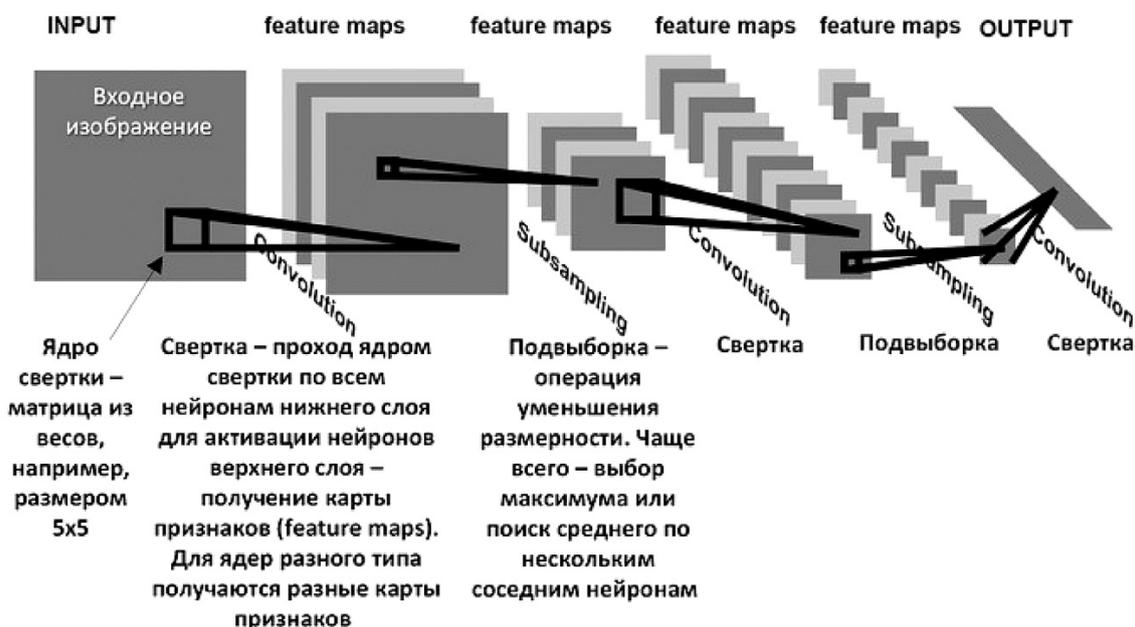


Рис. 2. Архитектура свёрточной нейросети в общем виде (LeCun et al., 1995)

Fig. 2. General architecture of a convolutional neural network (LeCun et al., 1995)

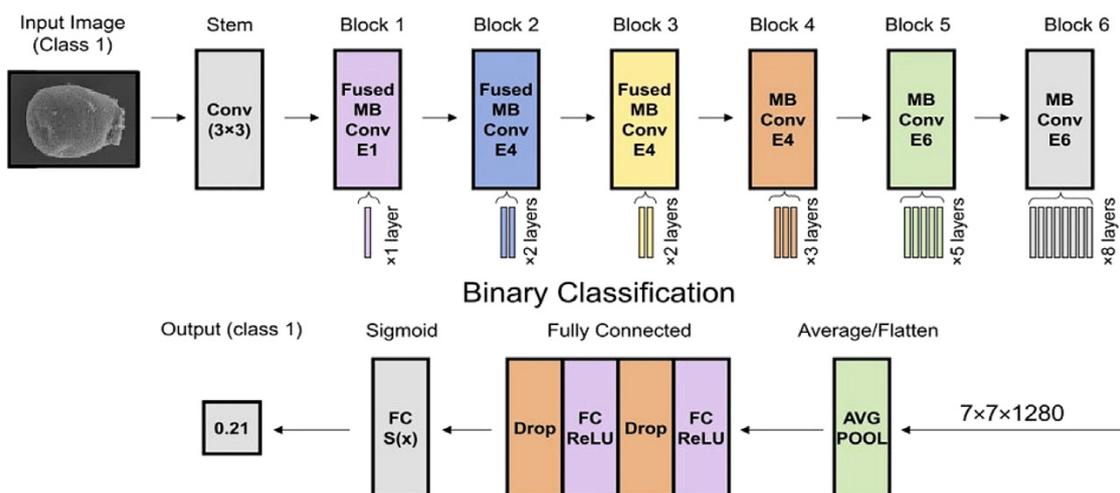


Рис. 3. Архитектура нейронной сети EfficientNetV2B0

Fig. 3. Architecture of the neural network EfficientNetV2B0

Таблица 1. Конфигурация EfficientNetV2B0

Table 1. EfficientNetV2B0 configuration

Этап // Stage	Блоки / Blocks	К. выхода / Outputs
Stem	Conv3x3	32
Стадия 1 / Stage 1	Fused-MBConv	16
Стадия 2 / Stage 2	Fused-MBConv	32
Стадия 3 / Stage 3	Fused-MBConv	48
Стадия 4 / Stage 4	MBConv	96
Стадия 5 / Stage 4	MBConv	112
Стадия 6 / Stage 6	MBConv	192
Классификация / Classification	Conv1x1, Global Pooling, FC	1280

Таблица 2. Отчет классификации оценочных данных

Table 2. Validation data classification report

Название класса Name of class	Количество изображений в классе Number of images in class	Распознано верно Recognized correctly	Распознано неверно Recognized incorrectly	Точность Accuracy
Хитинозои Chitinozoans	737	729	8	99 %
Нехитинозои Non-chitinozoans	876	862	14	98.4 %

Таблица 3. Точность распознавания по классам на тестовом наборе данных

Table 3. Recognition accuracy by classes on the test data set

Название класса / Name of class	Количество изображений в классе Number of images in class	Точность / Accuracy
Хитинозои / Chitinozoans	239	98.9 %
Нехитинозои / Non-chitinozoans	294	98.4 %

Для реализации обучения и распознавания использовались: IDE Visual Studio Code 1.90.2, язык программирования Python 3.11.6, платформа Kaggle Code. Исходные размеры используемых изображений не являются важными для распознавания, так как перед подачей на вход нейросети они приводятся к единому размеру. Количество извлекаемых характеристик или признаков соответствует количеству выходных фильтров каждого блока модели. В данной модели их от 16 до 1280 (табл. 1)

По результатам обработки и классификации изображений из проверочных и оценочных данных была проведена оценка эффективности (табл. 2).

По отчету о классификации данных можно сделать вывод, что модель чуть хуже классифицирует хитинозои по сравнению с нехитинозоями. Это может указывать на то, что класс изображений хитинозой визуально менее разнообразен, чем класс нехитинозой, но так как в каждом классе достаточно большое количество изображений, ошибка очень мала, а точность модели составляет 98 %.

На конечном этапе проводилась оценка точности работы системы с использованием тестового набора данных, не участвовавших в обучении нейронной сети (табл. 3).

Высокая точность достигается благодаря большому количеству параметров модели нейронной сети и достаточному количеству изображений для обучения — 4826.

Ошибочные определения хитинозой как нехитинозой связаны с ограничением самих нейросетей и зависимостью от выбора обучающих данных (рис. 4).

Данные изображения хитинозой (на рис. 4) не были определены как хитинозои, так как на них при-

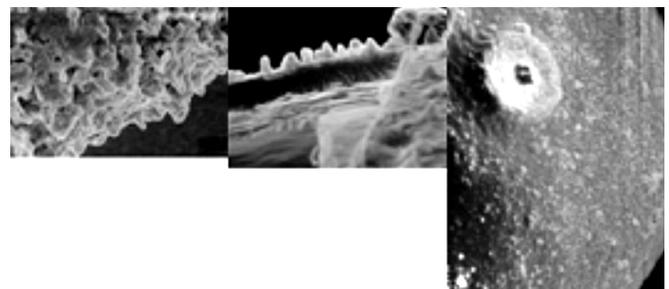


Рис. 4. Хитинозои, определенные как нехитинозои

Fig. 4. Chitinozoans defined as non-chitinozoans

сутствует только часть исходного изображения. Из-за ограничений нейронной сети в формате входных данных она работает с изображениями строго размером 224 на 224 пикселя. Изображения меньшего размера дополняются цветами соседних пикселей, которые на большинстве изображений представлены чёрным или серым. Изображения большего размера обрезаются, из-за чего нейронная сеть ошибочно относит их к другому классу. Данная проблема может быть решена путём уменьшения разрешения исходного изображения.

Выводы

Современные нейронные сети применяются для распознавания изображений в разных областях фундаментальных и прикладных наук (Гусев, 2017; Журавлев, Груздев, 2024; Gorur et al., 2023; Marmo et al., 2006; Ruikar, Agrawal, 2019; White, 2020) и впервые были использованы для распознавания изображений хитинозой.



Создание инновационного решения, которое интегрирует передовые методы компьютерного зрения и машинного обучения для автоматизированного распознавания изображений микрофоссилий на примере хитинозой, на первом этапе показало хорошие результаты. Эффективность разработанных методов была подтверждена на тестовом наборе данных. Тестирование показало высокую точность алгоритма распознавания, равную 98 %.

Продолжение работ по созданию системы анализа изображений и определения хитинозой включает следующие этапы.

После решения задачи определения объекта и отнесения его к классу хитинозой потребуются отнесение изображений выбранных хитинозой к одному из трех семейств. Следующим этапом будет отнесение изображения к подсемейству, а затем и к роду. Финальный этап применения нейросетей предполагает автоматизированное определение изображения до вида и решение задачи анализа изображений неполных (поврежденных, деформированных) везикул хитинозой. Для данных преобразований потребуются применение не только сверточной сети (CNN), но и генеративно-состязательной сети (GAN) для воссоздания отсутствующих фрагментов везикулы.

Авторы выражают глубокую признательность Г. Р. Шахматовой за разработку методики и ценные замечания к статье.

Публикация выполнена по государственной программе FMRS-2022-0010 «Фундаментальные исследования. Основные события фанерозоя: палеонтология, стратиграфия, корреляция».

Литература/ References

- Гусев А. В. Перспективы нейронных сетей и глубокого машинного обучения в создании решений для здравоохранения // Искусственный интеллект в здравоохранении. 2017. № 3. С. 92—105. URL: https://www.researchgate.net/publication/337923586_Perspektivy_nejronnyh_setej_i_glubokogo_masinnogo_obucenia_v_sozdanii_resenij_dla_zdravoohranenia
- Gusev A. V. Prospects of neural networks and deep machine learning in creating solutions for healthcare. Artificial Intelligence in healthcare. 2017, No. 3, pp. 92—105. (in Russian)
- Журавлев А. В., Груздев Д. А. Анализ изображений шлифов карбонатных обломочных пород с помощью систем

на основе искусственного интеллекта // Вестник геонаук. 2024. 6 (354). С. 28—31. DOI: 10.19110/geov.2024.6.3

Zhuravlev A. V., Gruzdev D. A. Analysis of images of thin sections of carbonate clastic rocks using artificial intelligence-based systems. Vestnik of Geosciences, 2024, 6 (354), pp. 28—31. DOI: 10.19110/geov.2024.6.3 (in Russian)

EfficientNetV2. URL: <https://github.com/google/automl/tree/master/efficientnetv2> (дата обращения: 11.08.2024)

Gardner K., Joshi R., Kashem M., Pham T., Lu Q., Li W. Label free identification of different cancer cells using deep learning-based image analysis // AIP Publishing 2023, Volume 1, Issue 2 <https://doi.org/10.1063/5.0141730>

Gorur K., Ozer C. K., Ozer I., Karaca A. C., Cetin O., Kocak I. Species-Level Microfossil Prediction for Globotruncana genus Using Machine Learning Models // Arabian Journal for Science and Engineering (2023) 48: P. 1315—1332. <https://doi.org/10.1007/s13369-022-06822-5>

LeCun Y., Jackel L., Bottou L., Brunot A., Cortes C., Denker J., Drucker H., Guyon I., Müller U., Sackinger E., Simard P., Vapnik V. Comparison of Learning Algorithms for Handwritten Digit Recognition // International Conference on Artificial Neural Networks, 1995. P. 53—60.

Marmo R., Amodio S., Cantoni V. 2006 Microfossils shape classification using a set of width values // 18th International Conference on Pattern Recognition, 2006. ICPR 2006 Volume: 1 DOI: 10.1109/ICPR.2006.797

Paris F. (2006). Chitinozoans: A fascinating and mysterious microfossil-group. Technical-card and direction for use. 81 p. <https://doi.org/10.13140/RG.2.1.2462.7682>.

Paris F., & Nölvak, J. (1999). Biological interpretation and paleobiodiversity of a cryptic fossil group: The “chitinozoan animal”. Geobios, 32, P. 315—324.

Ruikar T., Agrawal V. L. A Review: Classification of Fossilized Radiolarian image using computational Intelligence Techniques // Journal of Emerging Technologies and Innovative Research February 2019, Volume 6, Issue 2 P. 345—349

Tan M., Quoc V. Le EfficientNetV2: Smaller Models and Faster Training / Proceedings of the 38 th International Conference on Machine Learning, PMLR 139, 2021. — URL: <https://arxiv.org/pdf/2104.00298>

White A., Deep learning in deep time // PNAS — 2020 117 (47) P. 29268—29270: <https://www.pnas.org/doi/full/10.1073/pnas.2020870117>

Поступила в редакцию / Received 16.08.2024