

АНАЛИЗ ГРАФИЧЕСКОЙ ИНФОРМАЦИИ МЕТОДАМИ МАШИННОГО ОБУЧЕНИЯ

Немковский Г.Б., руководитель департамента разработки и внедрения ООО «ВЕСТТРЕЙД ЛТД»; ассистент кафедры информационных технологий и обработки медицинских данных Центра цифровой медицины ФГАОУ ВО «Первый МГМУ им. И.М. Сеченова» Минздрава России (Сеченовский университет)

Ашапкина М.С., канд. техн. наук, ассистент кафедры информационных технологий и обработки медицинских данных Центра цифровой медицины ФГАОУ ВО «Первый МГМУ им. И.М. Сеченова» Минздрава России (Сеченовский университет)

Ефимов С.А., студент ФГАОУ ВО «Первый МГМУ им. И.М. Сеченова» Минздрава России (Сеченовский университет)

Деревянко Е.А., студент ФГАОУ ВО «Первый МГМУ им. И.М. Сеченова» Минздрава России (Сеченовский университет)

Современные технологии компьютерного зрения и искусственного интеллекта находят широкое применение в различных областях, таких как медицина, транспорт, системы видеонаблюдения и безопасность. В данной работе рассматривается анализ методов обработки визуальной информации, используемых в машинном обучении. Особое внимание уделяется задачам обнаружения объектов, распознавания лиц, анализа эмоций, определения положения тела человека, выявления аномалий и оптического распознавания символов. Обсуждаются ключевые применения этих технологий в здравоохранении, безопасности, транспорте и других отраслях. Целью работы являются изучение и классификация методов обработки визуальных данных, а также определение основных задач, возникающих в этом процессе. В данной статье предпринята попытка классификации существующих инструментов с открытым исходным кодом.

Ключевые слова: компьютерное зрение, машинное обучение, анализ графической информации, обнаружение объектов.

Цитирование: Немковский Г.Б., Ашапкина М.С., Ефимов С.А., Деревянко Е.А. Анализ графической информации методами машинного обучения // Информационно-экономические аспекты стандартизации и технического регулирования. 2025. № 2 (83). С. 66–74.

ВВЕДЕНИЕ

С развитием вычислительной техники стало возможным применение технологий компьютерного зрения и глубокого обучения для анализа статических и динамических графических данных. Эти технологии нашли широкое применение в различных предметных областях, включая медицину, системы видеонаблюдения и транспорт. С использованием современных методов анализа стало возможным решение задач обнаружения и классификации объектов, распознавания лиц, анализа эмоций и многих других задач.

В современных транспортных системах компьютерное зрение играет ключевую роль в распознавании номерных знаков автомобилей. Интеллектуальные технологии задействованы на нескольких этапах обработки видеопотока: с их помощью улучшаются освещение и качество изображения, выделяются контуры номера и осуществляется распознавание символов [1].

Анализ медицинских изображений – это обширная область исследований, охватывающая процессы обработки,

интерпретации и изучения медицинских снимков [2]. Алгоритмы глубокого обучения стали все чаще применяться для улучшения диагностики, лечения и мониторинга различных медицинских состояний [3]. При анализе медицинских изображений алгоритмы глубокого обучения способны автоматически выявлять и классифицировать аномалии на разнообразных снимках, таких как рентгеновские, магнитно-резонансная томография (МРТ), компьютерная томография и ультразвуковые изображения [4].

Целью данной работы являются изучение и анализ методов обработки визуальных данных, используемых в машинном обучении (ML), а также в определении ключевых задач, возникающих в процессе обработки визуальной информации.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Задача анализа статических и динамических изображений встречается в финансовом секторе, здравоохранении и безопасности, но не ограничивается ими. Для получения и анализа информации из изображений и видео используется сочетание технологий нейронных сетей, компью-

терного зрения и обработки изображений. Эти технологии предоставляют возможность частичной, а в отдельных случаях и полной автоматизации выполнения задачи анализа и реализации системы поддержки принятия решений.

В таблице представлены стандартные задачи графического анализа и области их применения.

Стандартные задачи графического анализа и области их применения

№	НАИМЕНОВАНИЕ	ОПИСАНИЕ	ОБЛАСТЬ ПРИМЕНЕНИЯ	ИНСТРУМЕНТЫ
1	Обнаружение и отслеживание объектов (<i>Object Detection and Tracking</i>)	Идентификация и локализация объектов на изображениях и видео, а также их дальнейшее отслеживание в последовательности кадров	■ безопасность (видеонаблюдение, обнаружение подозрительных действий) [5]; ■ транспорт (отслеживание пешеходов и транспортных средств для обеспечения безопасности на дорогах) [6, 7]; ■ здравоохранение (диагностика заболеваний по различным снимкам) и др. [3, 4, 8]	YOLO (You Only Look Once) Detectron2 (Faster R-CNN) Keras PyTorch
2	Определение положения тела человека (<i>HPE, Human Pose Estimation</i>)	Идентификация и классификация узлов человеческого тела. Определение координат ключевых точек (рук, головы, туловища) и формирование скелетоподобного, описание человеческого тела	■ спорт (анализ движений спортсменов для улучшения техники) [9]; ■ здравоохранение (реабилитация пациентов) [10]; ■ AR/VR (создание интерактивных приложений, использующих данные о позе) [11]	OpenPose HRNet Keras PyTorch
3	Распознавание лиц (<i>Facial Recognition</i>)	Идентификация и верификация личности человека на основе его лица. Анализ изображения для выявления уникальных черт лица	■ безопасность (системы видеонаблюдения для идентификации подозреваемых) [12, 13]; ■ мобильные устройства и IoT (авторизация в системе с помощью распознавания лиц) [13]; ■ социальные сети (разметка пользователей на фотографиях) [14]	FaceNet DeepFace Dlib OpenCV
4	Распознавание эмоций (<i>Emotion Recognition</i>)	Определение эмоционального состояния человека на основе выражения лица или иных биометрических данных. Включает анализ лица, жестов и голоса	■ маркетинг (анализ реакции потребителей) [15]; ■ здравоохранение (исследование эмоционального состояния пациентов) [16]	FER (Facial Emotion Recognition) DeepFace
5	Выявление аномалий (<i>Anomaly Detection</i>)	Обнаружение редких событий или наблюдений, отличающихся от большинства данных	■ банкинг (антифрод-системы) [17]; ■ производство (мониторинг состояния с помощью визуальных данных для предотвращения поломок оборудования) [18]	Keras
6	Оптическое распознавание символов (<i>OCR, Optical Character Recognition</i>)	Преобразование документов различного типа (сканированные бумаги или изображения текста) в редактируемые данные	■ бизнес (автоматизация ввода данных с бумажных носителей) [19]; ■ юридические услуги (архивирование документов в электронном формате) [19]; ■ образование (цифровизация учебных материалов) [20]	Tesseract OCR OCRmyPDF EasyOCR
7	Генерация объектов (<i>Object Generation</i>)	Создание новых графических объектов на основе существующих данных	■ искусство и дизайн (генерация графических объектов) [21]; ■ игровая индустрия (генерация дизайна персонажей и окружения) [21]; ■ медицина (генерация медицинских изображений) [22]	StyleGAN & StyleGAN2 TensorFlow-GAN PyTorch GAN Zoo
8	Сегментация изображений (<i>Image Segmentation</i>)	Разделение изображения на несколько сегментов для упрощения анализа. Выделение объектов и границ внутри изображения	■ здравоохранение (выделение органов или опухолей на медицинских снимках) [23, 24]; ■ транспорт (определение дорожной разметки и препятствий для автономных транспортных средств) [25]	Detectron2 (Mask R-CNN) U-Net Keras
9	Классификация сцен (<i>Scene Classification</i>)	Определение содержания изображения на уровне сцен	■ мультимедийные приложения (упорядочивание изображений по категориям через их анализ) [26, 27]; ■ поиск изображений в интернете [28]; ■ анализ видеоконтента для рекомендаций [29]	VGG16/VGG19 (Keras) Caffe ResNet (TensorFlow) PyTorch Keras

В машинном обучении можно выделить ряд подходов, каждый из которых подходит для анализа графической информации:

■ Сверточные нейронные сети (CNN, Convolutional Neural Networks)

CNN – это широко используемый в задачах машинного обучения класс глубоких нейронных сетей, специально разработанных для обработки данных с сетчатой структурой, к которым относятся и изображения. CNN эффективно привлекают пространственные и временные признаки из видеопотока, что делает их особенно полезными для задач классификации и распознавания объектов в видео. Отличительной особенностью CNN являются так называемые «сверточные слои» (Convolutional Layers), обрабатывающие входные данные. Каждый нейрон сверточного слоя получает входные данные от небольших областей, называемых рецептивным полем. Это позволяет системе изучать локальные закономерности в данных [30–32].

■ Рекуррентные нейронные сети и сети с долгой краткосрочной памятью (RNN & LSTM, Recurrent Neural Networks & Long Short-Term Memory)

Рекуррентные нейронные сети (RNN) – это особый вид нейронных сетей, отличающийся своей способностью обрабатывать последовательные данные (тексты, временные ряды и аудио). Основным элементом RNN является рекуррентный блок, обеспечивающий передачу информации от одного временного шага к следующему через систему обратных связей в сети [33]. Отличительной особенностью рекуррентных сетей являются циклы обратной связи, позволяющие сетям обрабатывать последовательные данные, учитывая контекст и зависимости между элементами последовательности.

Необходимо отметить, что RNN представляют собой динамическую систему, то есть состояние сети в текущий момент времени зависит не только от входных сигналов, но и от предыдущего состояния [34]. При этом у рекуррентных сетей есть свои недостатки: проблема затухания градиента и ограниченная способность обрабатывать долгосрочные зависимости.

Для решения проблемы обработки долгосрочных зависимостей была разработана модификация RNN, получившая название LSTM (Long Short-Term Memory). Ключевым элементом LSTM являются блоки памяти, которые позволяют сети эффективно обрабатывать и запоминать информацию на длительный период времени [35].

LSTM является мощным инструментом для обработки последовательностей данных. Он широко применяется в различных областях, включая финансы и лингвистику.

■ Двухпоточные сети (Two-Stream Networks)

Двухпоточные сети представляют собой архитектуру нейронных сетей, предназначенную для эффективного распознавания действий в видео. Эта архитектура объединяет два отдельных потока обработки данных: RGB-поток, обрабатывающий входные цветные изображения и служащий для распознавания визуальных признаков (объектов или сцен), и градиентный поток, обрабатывающий изображения с градиентом, отражающим движение и изменения в видео. На более поздних этапах сети оба потока объединяются, и полученная информация используется для принятия решений.

Two-Stream Networks представляют собой подход к распознаванию действий в видео, который сочетает в себе визуальную и ориентированную на движение информацию, значительно повышая точность распознавания [36].

■ Генеративно-сопоставительные сети (GAN, Generative Adversarial Networks)

Генеративно-сопоставительные сети (GAN) представляют собой класс нейронных сетей, состоящий из двух моделей: генератора и дискриминатора. Генератор отвечает за создание новых данных, которые должны быть максимально похожи на реальные. Он получает на вход случайные данные и преобразует их в данные, которые имитируют распределение реальных данных.

Дискриминатор – нейронная сеть, которая оценивает вероятность того, что данный образец получен из реального набора данных или создан генератором. Дискриминатор обучается различать реальные и сгенерированные данные. Генератор и дискриминатор обучаются одновременно в сопоставительном режиме. Генератор стремится создавать данные, которые будут приняты дискриминатором за реальные, а дискриминатор – улучшать свою способность различать реальные и сгенерированные данные.

Генеративно-сопоставительные сети представляют собой мощный инструмент, который позволяет создавать новые данные. Эти сети могут быть использованы в различных областях, включая обработку изображений и медицину [37].

■ Локальные бинарные шаблоны (LBP, Local Binary Patterns)

LBP (Local Binary Patterns) – это метод, который анализирует текстуру изображения в локальной области вокруг каждого пикселя. LBP – эффективный инструмент для анализа текстур и распознавания образов, который находит применение в различных областях, связанных с компьютерным зрением и обработкой изображений.

Работа LBP базируется на сравнении значения яркости центрального пикселя с соседними. По результатам формируется двоичный код, который описывает текстуру в локаль-

ной области вокруг центра. Этот код называется LBP-кодом. Сильной стороной такого подхода в том числе является устойчивость к монотонным изменениям яркости [38].

■ Каскады Хаара (*Haar Cascades*)

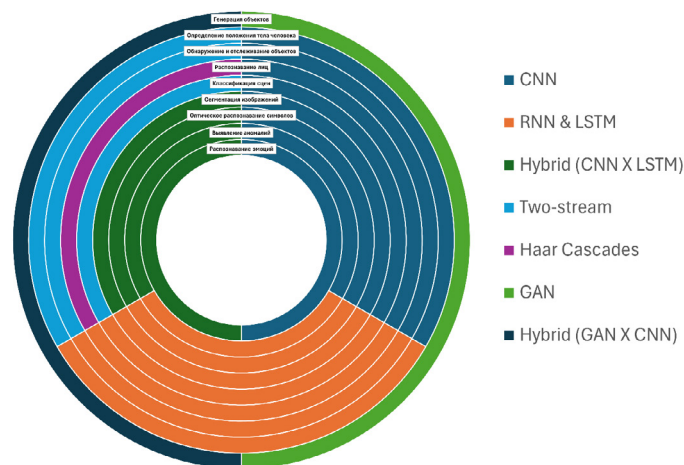
Haar Cascades – один из подходов в машинном обучении, представляющий собой набор классификаторов, обученных для обнаружения объектов посредством каскадной фильтрации, что позволяет быстро отсеивать области изображения, не содержащие целевые объекты [39]. Работа каскадов основана на применении «Признаков Хаара», простых функций, которые вычисляют разницу между суммами пикселей в двух прямоугольных областях на изображении. Они служат для выделения характерных черт объектов, таких как края и линии. Несколько простых классификаторов объединяются в последовательную каскадную структуру. Каждый классификатор в каскаде проверяет наличие объекта в определенном окне изображения и отбрасывает области, которые с высокой вероятностью не содержат искомого объекта. При обучении каскадов используются как положительные, так и отрицательные образцы, которые представляют собой изображения с объектом и без него соответственно.

Haar Cascades представляет собой эффективный и быстрый метод обнаружения объектов на изображениях, который используется в области компьютерного зрения и обработки изображений.

■ Гибридные нейронные сети (*Hybrid Neural Networks*)

Hybrid Neural Networks представляют собой сочетание различных типов нейронных сетей, которые можно адаптировать для решения сложных задач. Эти сети объединяют преимущества различных подходов, таких как рекуррентные нейронные сети (RNN), сверточные нейронные сети (CNN) и глубокие нейронные сети. За счёт сочетания полносвязных, сверточных и рекуррентных слоев гибридные сети могут эффективно обрабатывать разнородные данные, а модульная структура позволяет легко добавлять или удалять компоненты [40], создавая эффективный инструмент для решения сложных задач, объединяющий в себе достоинства разных видов нейронных сетей и методов обработки информации.

Набор задач и инструментов, предназначенных для их решения, можно представить в виде схемы «Применимость инструментов машинного обучения для решения задач анализа графической информации». Заметим, что для решения каждого вида задач доступны программные продукты с открытым исходным кодом, позволяющие с достаточной эффективностью решать эти задачи. Они представлены в последнем столбце таблицы.



Применимость инструментов машинного обучения для решения задач анализа графической информации

Как можно видеть из схемы, сверточные нейронные сети являются наиболее часто используемой технологией анализа графической информации. Примечательно, что нейронные сети этого типа применимы для решения 8 из 9 выделенных классов задач. Исключение составляет задача генерации объектов, что связано с ограничениями, присущими этому классу задач.

RNN и LSTM сети применимы для решения 7 из 9 классов задач, что подчеркивает их важность для обработки последовательных данных и временных зависимостей в графической информации.

Дополнительный анализ литературы для выявления причин такого явного ощутимого доминирования указанных методов выявил, что это обусловлено несколькими факторами, а именно способностью данных методов эффективно фиксировать пространственные и временные характеристики, надежностью обработки сложных наборов данных и более высокой производительностью по сравнению с другими технологиями машинного обучения. Рассмотрим преимущества сверточных и рекуррентных нейронных сетей подробнее.

Среди преимуществ CNN можно выделить:

- **Извлечение признаков:** CNN отлично справляются с извлечением признаков из изображений. Сверточные слои сканируют входные данные и захватывают локальные шаблоны, включая края и текстуры, которые необходимы для задач классификации изображений. Несколько научных работ уже обнаружили, что сверточные нейронные сети превосходят традиционные методы извлечения признаков, поскольку они обучаются самостоятельно, без значительного объема ручной разработки признаков [41, 42]

- **Инвариативность трансляции:** слои объединения в сверточных нейронных сетях обеспечивают инвариативность трансляции, позволяя модели распознавать объекты независимо от их положения на изображении. Это свойство особенно полезно в задачах обнаружения объектов и распознавания лиц, где точное положение объекта может существенно отличаться [41, 43]
- **Масштабируемость и эффективность:** архитектуры CNN способны масштабироваться, поскольку более глубокие сети, такие как ResNet и DenseNet, обеспечивают большую эффективность обучения на больших объемах данных ImageNet [42, 44].

К преимуществам RNN & LSTM можно отнести:

- **Захват временных зависимостей:** LSTM-сети специально разработаны для захвата долгосрочных зависимостей в последовательных данных с помощью механизмов стробирования. За счет этого они способны запоминать информацию в течение длительного времени, что имеет решающее значение для задач с анализом видео, где требуется контекст из предыдущих кадров [43, 45]
- **Переменный размер входных данных:** в отличие от традиционных нейронных сетей RNN могут обрабатывать последовательности любой длины. Поэтому они особенно хорошо подходят для задач, размеры входных данных могут меняться [41, 45]
- **Возможность интеграции с CNN:** было доказано, что интеграция CNN с LSTM (гибридные архитектуры) повышает эффективность итоговой нейронной сети за счет

использования преимуществ обеих моделей. Этот гибридный подход доказал свою эффективность во многих задачах, включая распознавание действий на основе видео и предиктивное моделирование данных временных рядов [43, 46].

ЗАКЛЮЧЕНИЕ

К числу основных областей применения графического информационного анализа, выявленных в ходе этого исследования, относятся обнаружение объектов, распознавание лиц, анализ эмоций и положения тела человека.

Результаты исследования показывают, что CNN являются наиболее используемой технологией в задачах обработки графических данных из-за их эффективности в извлечении пространственных признаков из изображений и видео. Это возможно благодаря особенностям устройства сверточных нейронных сетей.

RNN и LSTM также играют важную роль, особенно в тех задачах, где крайне важны временные зависимости.

Каждая задача, выявленная в ходе этого исследования, иллюстрирует, как технологии искусственного интеллекта могут быть использованы на практике для различных приложений. В этом отношении текущая исследовательская работа подчеркивает потенциал, который несут методы машинного обучения в области анализа графической информации.

Список использованных источников и литературы

1. Манюкова Н.В. Компьютерное зрение как средство извлечения информации из видеоряда // Математические структуры и моделирование. 2015. №4 (36). С. 123–128.
2. Guo Z., Li X., Huang H., Guo N., Li Q. Deep Learning-based Image Segmentation on Multimodal Medical Imaging// IEEE transactions on radiation and plasma medical sciences. 2019 T. 3(2). С. 162–169. DOI: <https://doi.org/10.1109/trpms.2018.2890359>
3. Zhang Y., Gorris J. M., Dong Z. Deep Learning in Medical Image Analysis // Journal of Imaging. 2021. T. 7(4). С. 74. DOI: <https://doi.org/10.3390/jimaging7040074>
4. Van der Velden B.H.M., Kuijff H.J., Gilhuijs K.G.A., & Viergever M.A. Explainable artificial intelligence (XAI) in deep learning-based medical image analysis // Medical image analysis. 2022. T. 79. 102470. DOI: <https://doi.org/10.1016/j.media.2022.102470>
5. Meenakshi S., Gomathy N. A survey on Real time Object Detection and Tracking Algorithms // International Journal of Applied Engineering Research. 2015. T. 10. С. 8290–8297.
6. Santos M. M. D., Hoffmann J. E., Tosso H. G., Malik A. W., Rahman A. U., Justo J. F. Real-time adaptive object localization and tracking for autonomous vehicles// IEEE Transactions on Intelligent Vehicles. 2020. T. 6 (3). – С. 450–459. DOI: <https://doi.org/10.1109/TIV.2020.3037928>.
7. Haris M., Glowacz A. Road Object Detection: A Comparative Study of Deep Learning-Based Algorithms // Electronics. 2021. T. 10. С. 1932. doi: <https://doi.org/10.3390/electronics10161932>
8. Cheng Y.C., Hung Y.C., Huang G.H., Chen T.B., Lu N.H., Liu K.Y., Lin K.H. Deep Learning-Based Object Detection Strategies for Disease Detection and Localization in Chest X-Ray Images // Diagnostics (Basel). 2024. T. 14(23). 2636. doi: <https://doi.org/10.3390/diagnostics14232636>.
9. Song L., Yu G., Yuan J., Liu Z. Human pose estimation and its application to action recognition: A survey // Journal of Visual Communication and Image Representation. 2021. T. 76. С. 103055. DOI: <https://doi.org/10.1016/j.jvcir.2021.103055>.
10. Tharatipyakul A., Thanawat S., Suporn P. Deep learning-based human body pose estimation in providing feedback for physical movement: A review // Heliyon. 2024. T. 10(17). e36589. ISSN 2405-8440, DOI: <https://doi.org/10.1016/j.heliyon.2024.e36589>.

11. Kulkarni S., Shetty A., Ashwath S., Kolkar R., Salian P. Exploring Challenges and Advances in Human Pose Estimation: An Investigation into Deep Learning Research and Artificial Intelligence // *International Journal of intelligent systems and applications in engineering*. 2024. С. 347–354.
12. Фисенко А.М., Трубоченко И.А., Русаков А.Д. Применение технологии распознавания лиц в сфере безопасности // *Научный взгляд в будущее*. 2018. Т. 1, № 21. С. 42–49. DOI: <https://doi.org/10.30888/2415-7538.2021-21-01-021>
13. Kortli Y., Jridi M., Falou A.A., Atri M. Face Recognition Systems: A Survey. *Sensors (Basel)*. 2020. 20(2): 342. DOI: <https://doi.org/10.3390/s20020342>.
14. Hupont I., Tolan S., Gunes H., et al. The landscape of facial processing applications in the context of the European AI Act and the development of trustworthy systems. *Sci Rep*. 2022. № 12. С. 10688. DOI: <https://doi.org/10.1038/s41598-022-14981-6>
15. Gupta T., Bansal S. AI-Driven Emotional Recognition in Digital Ads: A Novel Approach to Consumer Engagement // *Journal of Marketing & Supply Chain Management*. 2023. Т. 2. С. 1–7. DOI: [https://doi.org/10.47363/JMSCM/2023\(2\)131](https://doi.org/10.47363/JMSCM/2023(2)131).
16. Dhuheir M., Albaseer A., Baccour E., Erbad A., Abdallah M., Hamdi M. Emotion Recognition for Healthcare Surveillance Systems Using Neural Networks: A Survey // *International Wireless Communications and Mobile Computing (IWCMC)*, Harbin. 2021. С. 681–687. DOI: <https://doi.org/10.1109/IWCMC51323.2021.9498861>.
17. Mohiuddin A., Abdun M., Md Rafiqul I. A survey of anomaly detection techniques in financial domain // *Future Generation Computer Systems*. 2016. Т. 55. С. 278–288. DOI: <https://doi.org/10.1016/j.future.2015.01.001>.
18. Rodríguez M., Tobón D.P., Múnera D. Anomaly classification in industrial Internet of things: A review // *Intelligent Systems with Applications*. 2023. Т. 18. С. 200232. DOI: <https://doi.org/10.1016/j.iswa.2023.200232>.
19. Singh A., Bacchuwar K., Bhasin A. A Survey of OCR Applications // *International Journal of Machine Learning and Computing (IJMLC)*. 2012. Т. 2. № 3. С. 314–318. DOI: <https://doi.org/10.7763/IJMLC.2012.V2.137>.
20. Kanagarathinam K., Chitra S. Review of Optical Character Recognition and its applications // *Journal of Engineering and Applied Sciences*. 2016. № 11. С. 3441–3444.
21. Голышева Е.Н., Медведев А.А., Масалитин Н.С. Основные подходы к генерации изображений с помощью нейронных сетей // *Инновационная наука*. 2023. № 11-2. С. 25–28.
22. Singh N., Raza K. Medical Image Generation using Generative Adversarial Networks // *ArXiv e-prints*. 2020. DOI: 10.48550/arXiv.2005.10687.
23. Ma J., He Y., Li F. et al. Segment anything in medical images // *Nature Communications*. 2024. № 15. С. 654. DOI: <https://doi.org/10.1038/s41467-024-44824-z>.
24. Seo H., Badiel Khuzani M., Vasudevan V., Huang C., Ren H., Xiao R., Jia X., Xing L. Machine learning techniques for biomedical image segmentation: An overview of technical aspects and introduction to state-of-art applications // *Medical physics*. 2020. Т. 47(5). e148–e167. DOI: <https://doi.org/10.1002/mp.13649>.
25. Mahmud M.N., Osman M.K., Ismail A.P., Ahmad F., Ahmad K.A., Ibrahim A. Road Image Segmentation using Unmanned Aerial Vehicle Images and DeepLab V3+ Semantic Segmentation Model // 2021 11th IEEE International Conference on Control System, Computing and Engineering (ICCSCE), Penang, Malaysia. 2021. С. 176–181. DOI: <https://doi.org/10.1109/ICCSCE52189.2021.9530950>.
26. Haloi P., Bhuyan M. Video Searching and Retrieval using Scene Classification in Multimedia Databases // 2nd International Conference for Emerging Technology (INCET). 2021. С. 1–7. DOI: <https://doi.org/10.1109/INCET51464.2021.9456317>.
27. Xu C., Shu J., Zhu G. Scene Classification Based on Heterogeneous Features of Multi-Source Data // *Remote Sensing*. 2023. Т. 15 (2). С. 325. DOI: <https://doi.org/10.3390/rs15020325>.
28. Apostolidis K., Abesser J., Cuccovillo L., Mezaris V. Visual and audio scene classification for detecting discrepancies in video: a baseline method and experimental protocol // In *Proceedings of the 3rd ACM International Workshop on Multimedia AI against Disinformation (MAD '24)*. Association for Computing Machinery, New York, USA. 2024. С. 30–36. DOI: <https://doi.org/10.1145/3643491.3660287>.
29. Khoury K., Zanella M., Gérin B., Godelaine T., Macq B., Saïd M., Vleeschouwer C., Ben Ayed I. Enhancing Remote Sensing Vision-Language Models for Zero-Shot Scene Classification // *arXiv e-prints*. 2024. DOI: <https://doi.org/10.48550/arXiv.2409.00698>.
30. Neural network flavors [Электронный ресурс] <https://caisplusplus.usc.edu/curriculum/neural-network-flavors/convolutional-neural-networks> (дата обращения 12.01.2025).
31. Taye M.M. Theoretical Understanding of Convolutional Neural Network: Concepts, Architectures, Applications, Future Directions // *Computation*. 2023. Т. 11(3). С. 52. DOI: <https://doi.org/10.3390/computation11030052>
32. O'Shea K., Nash R. An Introduction to Convolutional Neural Networks // *ArXiv e-prints*. 2015. DOI: <https://doi.org/10.48550/arXiv.1511.08458>.
33. Глазкова А.В. Сравнение нейросетевых моделей для классификации текстовых фрагментов, содержащих биографическую информацию // *Программные продукты и системы*. 2019. № 2. С. 263–267.
34. Бендерская Е.Н., Никитин К.В. Рекуррентная нейронная сеть как динамическая система и подходы к ее обучению // *Информатика, телекоммуникации и управление*. 2013. №4 (176). С. 29–40.

35. Лабусов М.В. Нейронные сети долгой краткосрочной памяти и их использование для моделирования финансовых временных рядов // Инновации и инвестиции. 2020. №3. С. 167–171.
36. Simonyan K., Zisserman A. Two-Stream Convolutional Networks for Action Recognition in Videos // University of Oxford: Oxford, UK. 2014. С. 568–576.
37. Лаптев В.В., Данилов В.В., Гергет О.М. Исследование генеративно-состязательных сетей для синтеза новых медицинских данных // Автоматизация и моделирование в проектировании и управлении. 2020. № 2 (8). С. 17–23.
38. Карцов С.К., Куприянов Д.Ю., Зыков А.Н. Анализ методов фрактального кодирования и применение метода LBP (local binary patterns) для обработки изображения // Известия ТулГУ. Технические науки. 2013. №11. С. 235–240.
39. Амеличев Г.Э., Панина В.С., Белов Ю.С. Распознавание лиц с использованием каскадов хаара // E-Scio. 2020. № 8(47). С. 221–228.
40. Аверкин А. Н. Гибридные модулярные нейронные сети // Статистика и экономика. 2016. №4. С. 8–11.
41. Xiping Q. A Comparison Study of Convolutional Neural Network and Recurrent Neural Network on Image Classification // In Proceedings of the 2022 10th International Conference on Information Technology: IoT and Smart City (ICIT '22). Association for Computing Machinery, New York. 2023. С. 112–117. DOI: <https://doi.org/10.1145/3582197.3582215>.
42. Salehin I., Islam M.S., Amin N., Baten M.A., Noman S.M., Saifuzzaman M., Yazmyradov S. Real-Time Medical Image Classification with ML Framework and Dedicated CNN-LSTM Architecture // Journal of Sensors. 2023. T. 19. DOI: <https://doi.org/10.1155/2023/3717035>.
43. Mienye I.D., Swart T.G., Obaido G. Recurrent Neural Networks: A Comprehensive Review of Architectures, Variants, and Applications // Information. 2024. T. 15(9). С. 517. DOI: <https://doi.org/10.3390/info15090517>.
44. Karaman Y., Akdeniz F., Kir Savaş B., Becerkli Y. A Comparative Analysis of SVM, LSTM and CNN-RNN Models for the BBC News Classification // Innovations in Smart Cities Applications. 2023. T. 6. С. 473–483. DOI: https://doi.org/10.1007/978-3-031-26852-6_44.
45. Takale D., Wattamwar A., Saipatwar S., Saindane H., Patil T. Comparative Analysis of LSTM, RNN, CNN and MLP Machine Learning Algorithms for Stock Value Prediction // Journal of Firewall Software and Networking. 2024. T. 2. С. 1–10. DOI: <https://doi.org/10.48001/JoFSN.2024.211-10>.
46. Zhen L., Bărbulescu A.A. Comparative Analysis of Convolutional Neural Network-Long Short-Term Memory, Sparrow Search Algorithm-Backpropagation Neural Network, and Particle Swarm Optimization-Extreme Learning Machine Models for the Water Discharge of the Buzău River, Romania // Water. 2024. T. 16(2) С. 289. DOI: <https://doi.org/10.3390/w16020289>.

ANALYSIS OF GRAPHIC INFORMATION USING MACHINE LEARNING METHODS

Nemkovskiy G.B., Head of R&D, WESTTRADE LTD, Assistant professor of Department of Informational and Internet Technologies I.M. Sechenov First Moscow State Medical University Ministry of Healthcare of the Russian Federation

Ashapkina M.S., Assistant professor of Department of Informational and Internet Technologies I.M. Sechenov First Moscow State Medical University Ministry of Healthcare of the Russian Federation

Efimov S.A., I.M. Sechenov First Moscow State Medical University Ministry of Healthcare of the Russian Federation

Derevianko E.A., I.M. Sechenov First Moscow State Medical University Ministry of Healthcare of the Russian Federation

Modern computer vision and artificial intelligence technologies are widely used in various fields, such as medicine, transportation, video surveillance systems and security. This paper analyzes the methods of visual information processing used in machine learning. Particular attention is paid to the tasks of object detection, face recognition, emotion analysis, determination of human body position, detection of anomalies and optical character recognition. Key applications of these technologies in healthcare, security, transportation and other industries are discussed. The purpose of the work is to study and classify the methods of visual data processing, as well as to identify the main problems arising in this process. This article attempts to classify existing open-source tools.

Keywords: computer vision, machine learning, graphical information analysis, object detection, AI-technologies, open-source software.

For citation: Nemkovskiy G.B., Ashapkina M.S., Efimov S.A., Derevianko E.A. Analysis of graphic information using machine learning methods. *Information and Economic Aspects of Standardization and Technical Regulation*. 2025; 2 (83): 66–74. (In Russ.).

References

1. Manyukova N.V. Komp'yuternoe zrenie kak sredstvo izvlecheniya informacii iz videoryada. *Matematicheskie struktury i modelirovanie*, 2015, no 4 (36), pp. 123–128.
2. Guo Z., Li X., Huang H., Guo N., Li Q. Deep Learning-based Image Segmentation on Multimodal Medical Imaging. *IEEE transactions on radiation and plasma medical sciences*, 2019 v. 3(2). pp. 162–169. DOI: <https://doi.org/10.1109/trpms.2018.2890359>
3. Zhang Y., Gorris J. M., Dong Z. Deep Learning in Medical Image Analysis. *Journal of Imaging*, 2021, vol. 7(4). C. 74. DOI: <https://doi.org/10.3390/jimaging7040074>.
4. Van der Velden B. H. M., Kuijf H. J., Gilhuijs K. G. A., & Viergever M. A. Explainable artificial intelligence (XAI) in deep learning-based medical image analysis. *Medical image analysis*, 2022, vol. 79, 102470. DOI: <https://doi.org/10.1016/j.media.2022.102470>
5. Meenakshi S., Gomathy N. A survey on Real time Object Detection and Tracking Algorithms. *International Journal of Applied Engineering Research*, 2015, vol. 10, pp. 8290–8297.
6. Santos M. M. D., Hoffmann J. E., Tosso H. G., Malik A. W., Rahman A. U., Justo J. F. Real-time adaptive object localization and tracking for autonomous vehicles. *IEEE Transactions on Intelligent Vehicles*, 2020, vol. 6 (3), pp. 450–459. DOI: <https://doi.org/10.1109/TIV.2020.3037928>.
7. Haris M.; Glowacz A. Road Object Detection: A Comparative Study of Deep Learning-Based Algorithms. *Electronics*, 2021, vol. 10, p. 1932. doi: <https://doi.org/10.3390/electronics10161932>.
8. Cheng Y.C., Hung Y.C., Huang G.H., Chen T.B., Lu N.H., Liu K.Y., Lin K.H. Deep Learning-Based Object Detection Strategies for Disease Detection and Localization in Chest X-Ray Images. *Diagnostics (Basel)*, 2024, vol. 14(23), 2636. doi: <https://doi.org/10.3390/diagnostics14232636>.
9. Song L., Yu G., Yuan J., Liu Z. Human pose estimation and its application to action recognition: A survey. *Journal of Visual Communication and Image Representation*, 2021, vol. 76, p. 103055. DOI: <https://doi.org/10.1016/j.jvcir.2021.103055>.
10. Tharatipyakul A., Thanawat S., Suporn P. Deep learning-based human body pose estimation in providing feedback for physical movement: A review. *Heliyon*, 2024, vol. 10(17), e36589. ISSN 2405-8440, DOI: <https://doi.org/10.1016/j.heliyon.2024.e36589>.
11. Kulkarni S., Shetty A., Ashwath S., Kolkar R., Salian P. Exploring Challenges and Advances in Human Pose Estimation: An Investigation into Deep Learning Research and Artificial Intelligence. *International Journal of intelligent systems and applications in engineering*, 2024, pp. 347–354.
12. Fisenko A.M., Trubchenko I.A., Rusakov A.D. Primenenie tekhnologii raspoznavaniya lic v sfere bezopasnosti. *Nauchnyj vzglyad v budushchee*. 2018. no 21, vol. 1, pp. 42–49, DOI: <https://doi.org/10.30888/2415-7538.2021-21-01-021>.
13. Kortli Y., Jridi M., Falou AA, Atri M. Face Recognition Systems: A Survey. *Sensors (Basel)*, 2020 Jan 7, vol. 20(2). P. 342. DOI: <https://doi.org/10.3390/s20020342>.
14. Hupont I., Tolan S., Gunes H. et al. The landscape of facial processing applications in the context of the European AI Act and the development of trustworthy systems. *Sci Rep* 12, 2022, 10688, DOI: <https://doi.org/10.1038/s41598-022-14981-6>
15. Gupta T., Bansal S. AI-Driven Emotional Recognition in Digital Ads: A Novel Approach to Consumer Engagement. *Journal of Marketing & Supply Chain Management*, 2023, vol. 2, pp. 1–7. DOI: [https://doi.org/10.47363/JMSCM/2023\(2\)131](https://doi.org/10.47363/JMSCM/2023(2)131).
16. Dhuheir M., Albaseer A., Baccour E., Erbad A., Abdallah M., Hamdi M. Emotion Recognition for Healthcare Surveillance Systems Using Neural Networks: A Survey. *International Wireless Communications and Mobile Computing (IWCMC)*, Harbin, 2021. pp. 681–687. – DOI: <https://doi.org/10.1109/IWCMC51323.2021.9498861>.
17. Mohiuddin A., Abdun .M, Md Rafiqul I. A survey of anomaly detection techniques in financial domain. *Future Generation Computer Systems*, 2016, vol. 55, pp. 278–288. DOI: <https://doi.org/10.1016/j.future.2015.01.001>.
18. Rodríguez M., Tobón D.P., Múnera D. Anomaly classification in industrial Internet of things: A review. *Intelligent Systems with Application*, 2023, vol. 18. P. 200232. DOI: <https://doi.org/10.1016/j.iswa.2023.200232>.
19. Singh A., Bacchuwar K., Bhasin A. A Survey of OCR Applications. *International Journal of Machine Learning and Computing (IJMLC)*, 2012, vol. 2(3), pp. 314–318, DOI: <https://doi.org/10.7763/IJMLC.2012.V2.137>.
20. Kanagarathinam K., Chitra S. Review of Optical Character Recognition and its applications. *Journal of Engineering and Applied Sciences*, 2016, vol. 11, pp. 3441–3444.
21. Golysheva E.N., Medvedev A.A., Masalitin N.S. Osnovnye podhody k generacii izobrazhenij s pomoshch'yu nejronnyh setej. *Innovacionnaya nauka*, 2023, Vol. 11(2), pp. 25–28.
22. Singh N., Raza K. Medical Image Generation using Generative Adversarial Networks. *ArXiv e-prints*, 2020, DOI: 10.48550/arXiv.2005.10687.

23. Ma J., He Y., Li F., et al. Segment anything in medical images. *Nature Communications*, 2024, vol. 15. P. 654. DOI: <https://doi.org/10.1038/s41467-024-44824-z>.
24. Seo H., Badieli Khuzani M., Vasudevan V., Huang C., Ren H., Xiao R., Jia X., Xing L. Machine learning techniques for biomedical image segmentation: An overview of technical aspects and introduction to state-of-art applications. *Medical physics*, 2020, vol. 47(5), e148–e167. DOI: <https://doi.org/10.1002/mp.13649>.
25. Mahmud M.N., Osman M.K., Ismail A.P., Ahmad F., Ahmad K.A., Ibrahim A. Road Image Segmentation using Unmanned Aerial Vehicle Images and DeepLab V3+ Semantic Segmentation Model. 2021 11th IEEE International Conference on Control System, Computing and Engineering (ICCSCE), Penang, Malaysia, 2021, pp. 176–181. DOI: <https://doi.org/10.1109/ICCSCE52189.2021.9530950>.
26. Haloi P., Bhuyan M. Video Searching and Retrieval using Scene Classification in Multimedia Databases //2nd International Conference for Emerging Technology (INCET), 2021, pp. 1–7. DOI: <https://doi.org/10.1109/INCET51464.2021.9456317>.
27. Xu C., Shu J., Zhu G. Scene Classification Based on Heterogeneous Features of Multi-Source Data. *Remote Sensing*, 2023, vol. 15(2). P. 325. DOI: <https://doi.org/10.3390/rs15020325>.
28. Apostolidis K., Abesser J., Cuccovillo L., Mezaris V. Visual and audio scene classification for detecting discrepancies in video: a baseline method and experimental protocol. In *Proceedings of the 3rd ACM International Workshop on Multimedia AI against Disinformation (MAD '24)*. Association for Computing Machinery, New York, NY, USA, 2024, pp. 30–36. DOI: <https://doi.org/10.1145/3643491.3660287>.
29. Khoury K., Zanella M., Gérin B., Godelaine T., Macq B., Saïd M., Vleeschouwer C., Ben Ayed I. Enhancing Remote Sensing Vision-Language Models for Zero-Shot Scene Classification, *arXiv e-prints*. 2024. DOI: <https://doi.org/10.48550/arXiv.2409.00698>.
30. Neural network flavors [Электронный ресурс] <https://caisplusplus.usc.edu/curriculum/neural-network-flavors/convolutional-neural-networks> (accessed January 12, 2025).
31. Taye M.M. Theoretical Understanding of Convolutional Neural Network: Concepts, Architectures, Applications, Future Directions. *Computation*, 2023, vol. 11(3). P. 52. DOI: <https://doi.org/10.3390/computation11030052>.
32. O'Shea K., Nash R. An Introduction to Convolutional Neural Networks. *ArXiv e-prints*, 2015, DOI: <https://doi.org/10.48550/arXiv.1511.08458>.
33. Glazkova A.V. Sravnenie nejrosetevykh modelej dlya klassifikacii tekstovykh fragmentov, sodержashchih biograficheskuyu informaciyu. *Programmnye produkty i sistemy*, 2019, no. 2, pp. 263–267.
34. Benderskaya E.N., Nikitin K.V. Rekurrentnaya nejronnaya set' kak dinamicheskaya sistema i podhody k ee obucheniyu. *Informatika, telekommunikacii i upravlenie*, 2013, no. 4 (176), pp. 29–40.
35. Labusov M.V. Nejronnye seti dolgoj kratkosrochnoj pamyati i ih ispol'zovanie dlya modelirovaniya finansovykh vremennykh ryadov. *Innovacii i investicii*, 2020, no 3, pp. 167–171.
36. Simonyan K., Zisserman A. Two-Stream Convolutional Networks for Action Recognition in Videos. *University of Oxford: Oxford, UK*, 2014, NIPS'14, pp. 568–576.
37. Laptev V.V., Danilov V.V., Gerget O.M. Issledovanie generativno-sostyazatel'nykh setej dlya sinteza novykh medicinskih dannykh. *Avtomatizaciya i modelirovanie v proektirovanii i upravlenii*, 2020, no 2 (8), pp. 17–23.
38. Karcov S.K., Kupriyanov D.YU., Zykov A.N. Analiz metodov fraktal'nogo kodirovaniya i primenenie metoda LBP (local binary patterns) dlya obrabotki izobrazheniya. *Izvestiya TulGU. Tekhnicheskie nauki*, 2013, no11, pp. 235–240.
39. Amelichev G.E., Panina V.S., Belov Y.S. Raspoznavanie lic s ispol'zovaniem kaskadov haara. *E-Scio*, 2020, no. 8(47), pp. 221–228.
40. Averkin A.N. Gibridnye modulyarnye nejronnye seti. *Statistika i ekonomika*, 2016, no. 4, pp. 8–11.
41. Xiping Q.A. Comparison Study of Convolutional Neural Network and Recurrent Neural Network on Image Classification. In *Proceedings of the 2022 10th International Conference on Information Technology: IoT and Smart City (ICIT '22)*. Association for Computing Machinery, New York, NY, USA, 2023, pp. 112–117. DOI: <https://doi.org/10.1145/3582197.3582215>.
42. Salehin I., Islam M.S., Amin N., Baten M.A., Noman S.M., Saifuzzaman M., Yazmyradov S. Real-Time Medical Image Classification with ML Framework and Dedicated CNN-LSTM Architecture. *Journal of Sensors*, 2023, vol. 19. DOI: <https://doi.org/10.1155/2023/3717035>.
43. Mienye, I.D., Swart, T.G., Obaido, G. Recurrent Neural Networks: A Comprehensive Review of Architectures, Variants, and Applications. *Information*, 2024, vol. 15(9). P. 517. DOI: <https://doi.org/10.3390/info15090517>.
44. Karaman Y., Akdeniz F., Kir Savaş B., Becerkli Y. A Comparative Analysis of SVM, LSTM and CNN-RNN Models for the BBC News Classification. *Innovations in Smart Cities Applications*, 2023, vol. 6, pp. 473–483. DOI: https://doi.org/10.1007/978-3-031-26852-6_44.
45. Takale D., Wattamwar A., Saipatwar S., Saindane H., Patil T. Comparative Analysis of LSTM, RNN, CNN and MLP Machine Learning Algorithms for Stock Value Prediction. *Journal of Firewall Software and Networking*, 2024, vol. 2, pp. 1–10. DOI: <https://doi.org/10.48001/JoFSN.2024.211-10>.
46. Zhen L., Bărbulescu A.A. Comparative Analysis of Convolutional Neural Network-Long Short-Term Memory, Sparrow Search Algorithm-Backpropagation Neural Network, and Particle Swarm Optimization-Extreme Learning Machine Models for the Water Discharge of the Buzău River, Romania. *Water*, 2024, vol. 16(2). P. 289. DOI: <https://doi.org/10.3390/w16020289>.