

УДК 004.932

EDN: LZPVVM

ИССЛЕДОВАНИЕ ПРИМЕНЕНИЯ ТЕХНОЛОГИЙ КОМПЬЮТЕРНОГО ЗРЕНИЯ ДЛЯ АНАЛИЗА БИМЕДИЦИНСКИХ ИЗОБРАЖЕНИЙ

А.Е. Савкин

ORCID: 0000-0003-4826-9087 e-mail: sae.20@bk.ru

Нижегородский государственный технический университет им. Р.Е. Алексеева
*Нижний Новгород, Россия***П.А. Шагалова**

ORCID: 0000-0002-6676-4228 e-mail: polli-shagalova@yandex.ru

Нижегородский государственный технический университет им. Р.Е. Алексеева
*Нижний Новгород, Россия***О.Н. Корелин**

ORCID: 0000-0002-5928-6045 e-mail: o_korelin@mail.ru

Нижегородский государственный технический университет им. Р.Е. Алексеева
Нижний Новгород, Россия

Исследованы возможности применения алгоритмов компьютерного зрения для распознавания клеток крови на биомедицинских изображениях. В качестве основного инструмента используются искусственные нейронные сети. Выполнена разметка и аугментация анализируемых изображений, подготовлены данные для обучения нейронных сетей, представлено сравнение различных нейросетевых архитектур. Также разработан алгоритм распознавания пересекающихся объектов, основанный на разделении контура объекта на сегменты по ключевым точкам. Приведены результаты работы алгоритмов на реальных изображениях микроскопии крови, выполнено сравнение их эффективности, выделены достоинства и недостатки. Наиболее высокие показатели точности распознавания были достигнуты с применением нейронной сети YOLOv8. Проведенные в рамках работы эксперименты показали эффективность применения нейросетевого подхода для анализа биомедицинских изображений.

Ключевые слова: искусственная нейронная сеть, распознавание образов, ключевые точки, микроскопия.

ДЛЯ ЦИТИРОВАНИЯ: Савкин, А.Е. Исследование применения технологий компьютерного зрения для анализа биомедицинских изображений / А.Е. Савкин, П.А. Шагалова, О.Н. Корелин // Труды НГТУ им. Р.Е. Алексеева. 2024. № 1. С. 16-23. EDN: LZPVVM

RESEARCH OF APPLICATION OF COMPUTER VISION TECHNOLOGIES FOR BIOMEDICAL IMAGE ANALYSIS

A.E. Savkin

ORCID: 0000-0003-4826-9087 e-mail: sae.20@bk.ru

Nizhny Novgorod State Technical University n.a. R.E. Alekseev
*Nizhny Novgorod, Russia***P.A. Shagalova**

ORCID: 0000-0002-6676-4228 e-mail: polli-shagalova@yandex.ru

Nizhny Novgorod State Technical University n.a. R.E. Alekseev
Nizhny Novgorod, Russia

O.N. Korelin

ORCID: 0000-0002-5928-6045 e-mail: o_korelin@mail.ru

Nizhny Novgorod State Technical University n.a. R.E. Alekseev

Nizhny Novgorod, Russia

Abstract. The paper presents studies on the possibility of using computer vision algorithms for blood cells recognition in biomedical images. Artificial neural networks are used as the main tool. The analyzed images were labeled and augmented, data for training neural networks were prepared, and a comparison of various neural network architectures was presented. An algorithm for intersecting objects recognition were developed. It is based on dividing the contour of an object into segments using key points. The results of the algorithms operation on real blood microscopy images are presented. Their effectiveness is compared and their advantages and disadvantages are highlighted. The highest recognition accuracy rates were achieved using the YOLOv8 model. Experiments showed the effectiveness of using a neural network approach for biomedical images analysis.

Key words: artificial neural network, image recognition, key points, microscopy.

FOR CITATION: A.E. Savkin, P.A. Shagalova, O.N. Korelin. Research of application of computer vision technologies for biomedical image analysis. Transactions of NNSTU n.a. R.E. Alekseev. 2024. № 1. Pp. 16-23. EDN: LZPVRM

1. Введение

Применение технологий компьютерного зрения для автоматизации анализа изображений, полученных при микроскопии крови, позволяет снизить влияние человеческого фактора, уменьшить нагрузку на специалистов в области лабораторной диагностики и повысить скорость обработки биологических образцов. Для распознавания объектов на биомедицинских снимках могут использоваться как искусственные нейронные сети, так и алгоритмы, основанные на математических методах [1-4]. Каждый из этих подходов имеет свои сильные и слабые стороны. Так, для обработки изображений с использованием традиционных математических алгоритмов в большинстве случаев необходимо меньшее количество вычислений по сравнению с нейронной сетью, при этом не требуются специализированные нейронные или графические ускорители. С другой стороны, модели глубокого обучения более гибки и универсальны. В отличие от математических алгоритмов, они позволяют избежать сложной процедуры ручной настройки параметров, а также зачастую показывают более высокую точность при распознавании объектов. Выбор конкретных алгоритмов или их комбинаций определяется особенностями решаемой задачи и требует проведения исследований.

2. Постановка задачи и входные данные

Данная работа посвящена исследованию применения алгоритмов компьютерного зрения в задаче анализа биомедицинских изображений, полученных при микроскопии образцов крови. На рис. 1. приведены примеры изображений, которые будут рассматриваться далее.

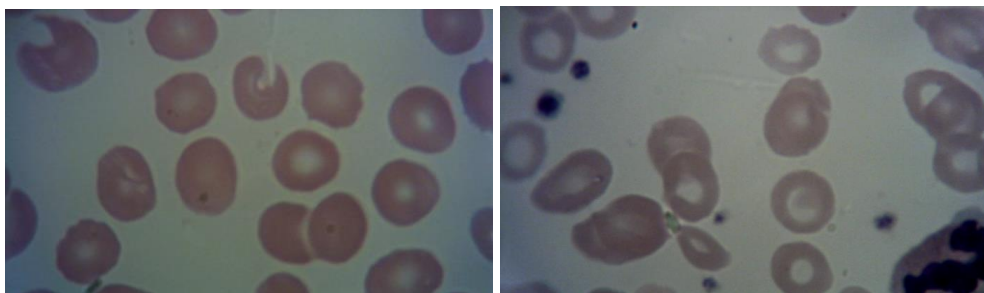


Рис. 1. Примеры изображений, полученных при микроскопии образцов крови

Fig. 1. Examples of microscopy images of blood samples

Исходный набор изображений состоит из 275 снимков, при этом объектами интереса являются клетки крови (эритроциты, лейкоциты, тромбоциты). Основным методом обнаружения и распознавания изучаемых объектов является мощный и гибкий инструмент, успешно применяемый для анализа изображений – искусственные нейронные сети. Чтобы подготовить изображения для обучения, необходима их разметка – данный этап был выполнен с использованием пакета *labelimg*; в результате для каждого изображения был сформирован файл разметки, содержащий координаты вершин, ограничивающих прямоугольников каждой клетки крови. Для увеличения объема обучающей выборки была выполнена аугментация данных с использованием методов библиотеки *imgaug*. При этом использовались: отражение по горизонтали и вертикали, применение шума, изменение резкости и контрастности изображений. После аугментации обучающий датасет из 2376 снимков был сформирован и разделен на обучающую и валидационную выборки в соотношении 80:20.

Решение рассматриваемой задачи детекции объектов предполагает определение наличия искомого объекта на изображении и поиск координат вершин ограничивающих прямоугольников каждого объекта. Существуют два основных подхода к решению данной задачи с использованием нейронных сетей: методы, основанные на регионах интереса (двухэтапные), и методы, не использующие отдельный алгоритм для генерации регионов (одноэтапные). В двухэтапных методах сначала происходит выделение регионов с высокой вероятностью содержания объектов, а затем выбранные области рассматриваются классификатором с целью определить, действительно ли область содержит искомым объект, и регрессором, который указывает расположение ограничивающей рамки. В одноэтапных методах определяются координаты ограничивающих рамок с различными характеристиками (например, вероятность принадлежности классу), и далее корректируется положение рамок.

3. Экспериментальная часть

В процессе выполнения работы были апробированы четыре нейросетевые архитектуры: *Faster R-CNN* (реализует двухэтапный метод детектирования объектов) и три архитектуры, принадлежащие к популярному семейству *YOLO* – *YOLOv3*, *YOLOv4*, *YOLOv8* (реализуют одноэтапный метод детектирования объектов).

Первой среди рассмотренных архитектур для распознавания объектов на основе глубокого обучения выступила нейронная сеть *Faster R-CNN* [5]. Она представляет собой полностью сверточную сеть с отдельным модулем генератора регионов интереса по признакам исходного изображения (RPN). Сгенерированные регионы передаются в два полносвязных слоя: *box-regression-layer*, прогнозирующий значения смещения для ограничивающих рамок, и *box-classification-layer*, классифицирующий изображения в пределах предлагаемой области. После прохождения слоя *RPN* следует слой *RoIPooling* для преобразования регионов к одному размеру и дальнейшей классификации и смещения границ ограничивающих рамок. Обучение сети производилось с помощью открытой нейросетевой библиотеки *Tensorflow*. В качестве основы для извлечения регионов интереса была использована нейронная сеть *VGG-16*. Обучение производилось в 14 эпох (10 эпох со скоростью обучения 0,001 и 4 эпохи со скоростью обучения 0.0001), в качестве оптимизатора алгоритма градиентного спуска был использован стохастический градиентный спуск [6]. Пример работы сети представлен на рис. 2.

Поскольку двухстадийные детекторы подразумевают две стадии, а соответственно, и две нейронные сети, по своей природе они проигрывают одностадийным детекторам, состоящим только из одной нейронной сети в части скорости обработки фотографий, показывая при этом схожий или даже меньший уровень точности. Скорость и точность обработки данных напрямую влияет на качество постановки диагноза, поэтому в данном исследовании нами были рассмотрены некоторые модели одностадийных детекторов семейства *YOLO*. Первой из них является *YOLOv3* – архитектура модели глубокого обучения, используемая для распознавания объектов на изображениях [7].

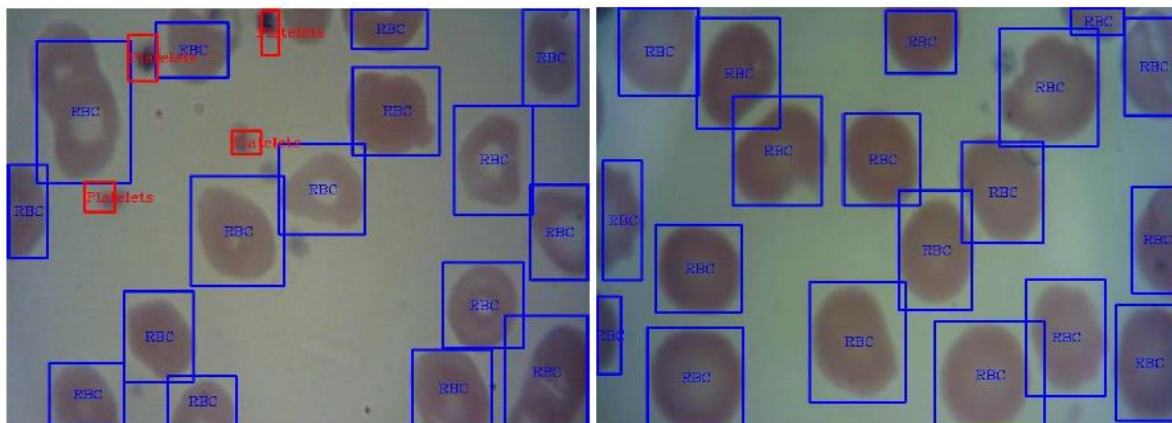


Рис. 2. Пример работы сети на базе архитектуры Faster R-CNN для распознавания клеток крови

Fig. 2. An example of a blood cells recognition network based on the Faster R-CNN architecture

Данная нейронная сеть подразделяется на три основные части: первая состоит из сверточных слоев для извлечения объектов из входного изображения, вторая объединяет карты признаков объектов разных масштабов для улучшения точности распознавания, третья состоит из полносвязных слоев, которые предсказывают местоположение и класс каждого объекта на изображении. В *YOLOv3* также используются якорные коробки, которые представляют собой предварительно определенные ограничивающие рамки разных размеров и соотношений сторон. В сочетании с сеткой, которая делит изображения на якорные ячейки, они используются для прогнозирования местоположения и размера объектов на изображении. Обучение проводилось с помощью фреймворка *darknet*, в качестве реализации модели *YOLO* на *Tensorflow* использован *darkflow*. Обучение проводилось в 250 эпох с размером батча 48, со скоростью обучения, равной $5e-5$. Пример работы сети представлен на рис. 3.

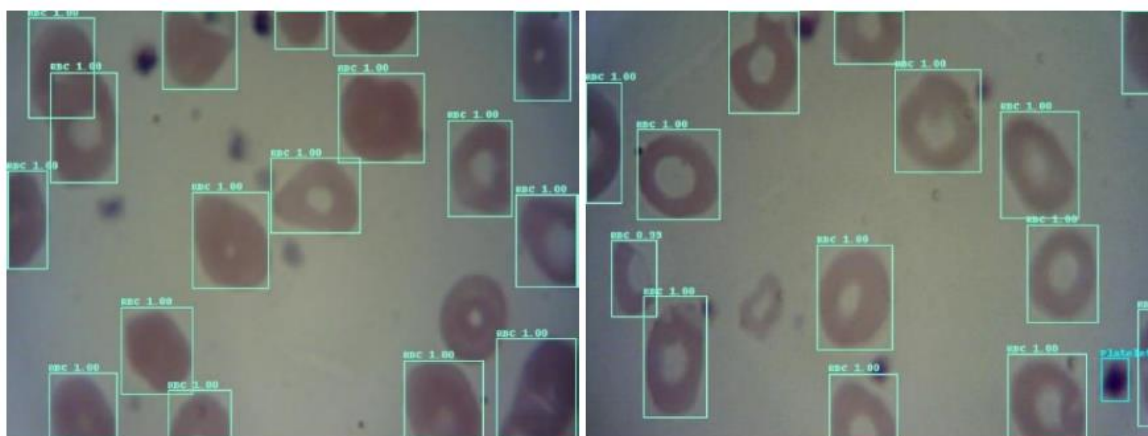


Рис. 3. Пример работы сети на базе архитектуры YOLOv3 для распознавания клеток крови

Fig. 3. An example of a blood cells recognition network based on the YOLOv3 architecture

Проведенные эксперименты показали, что в большом количестве случаев *YOLOv3* неудовлетворительно справляется с распознаванием клеток крови (средняя точность *mAP* составляет 46,1 %). В связи с этим было принято решение апробировать следующее поколение детекторов – *YOLOv4* [8]. Данная архитектура содержит новые методы, повышающие точность и скорость работы сверточной нейронной сети, среди которых можно выделить как наиболее важные и универсальные пакетную нормализацию и остаточные соединения. Для обучения также использовалась открытая нейросетевая библиотека *darknet*.

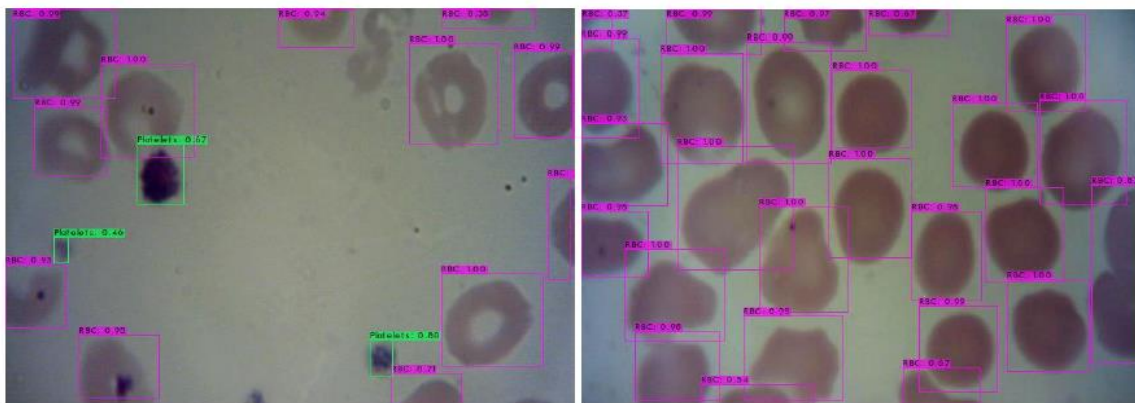


Рис. 4. Пример работы сети на базе архитектуры YOLOv4 для распознавания клеток крови

Fig. 4. An example of a blood cells recognition network based on the YOLOv4 architecture

Количество итераций – 6000, размер батча 32, скорость обучения = 0,001. Пример работы сети представлен на рис. 4. Кроме архитектуры *YOLOv4*, особый интерес для исследования представляет новейшее семейство моделей обнаружения объектов, представленное в начале 2023 г. – *YOLOv8* [9]. По сравнению с предыдущими версиями, модель *YOLOv8* работает быстрее и точнее, обеспечивая при этом единую структуру для обучения моделей для выполнения обнаружение объектов, сегментации экземпляров и классификации изображений. В отличие от рассмотренных *YOLOv3* и *YOLOv4*, в новой версии *YOLOv8* предсказывается непосредственно центр объекта, а не смещение от элементов сетки, а также улучшена аугментация изображений в процессе обучения. Для обучения использовался фреймворк *PyTorch*. Гиперпараметры были подобраны эмперически, использовалась аугментация в процессе обучения. Количество эпох – 1000, размер батча динамически вычисляется при запуске обучения в зависимости от доступного количества памяти, скорость обучения = 0,01. Пример работы сети представлен на рис. 5.

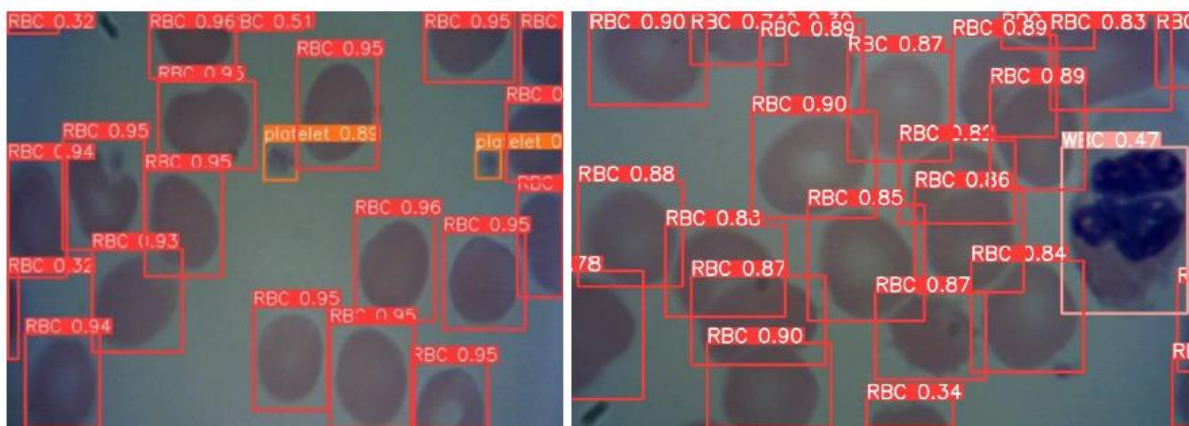


Рис. 5. Пример работы сети на базе архитектуры YOLOv8 для распознавания клеток крови

Fig. 5. An example of a blood cells recognition network based on the YOLOv8 architecture

Результаты проведения экспериментов показали, что сложными для распознавания являются случаи визуального пересечения объектов друг другом. При обработке таких участков изображений показатели точности падают. Для решения данной проблемы в разрабатываемое решение был добавлен обработчик пересекающихся объектов. Пересекающимися объектами в основном являются эритроциты, как наиболее многочисленный класс объектов на изображениях. Множество визуально перекрывающихся друг друга эритроцитов в общем случае имеют произвольную форму, при этом главная визуальная особенность, позволяющая дифференцировать эритроциты – это точки, в которых одни эритроциты пересекают другие. В норме эритроциты круглые, а окружность можно построить по любому ее фраг-

менту (дуге). На основании этих положений разработан и реализован алгоритм, основанный на разделении границ объекта (представляющего собой множество пересекающихся эритроцитов), на фрагменты по точкам пересечения с последующим достраиванием полученных дуг до полной окружности (каждая такая окружность представляет собой эритроцит).

Реализация разработанного алгоритма анализа пересекающихся объектов включает следующие шаги: бинаризация части изображения с пересекающимися объектами, выделение контуров пересекающихся объектов, поиск ключевых (угловых) точек, деление контуров на фрагменты по угловым точкам, аппроксимация контура каждого фрагмента окружностью, удаление слишком мелких объектов, которые точно не являются искомыми объектами, а также удаление повторных срабатываний. Наиболее важными из перечисленных выше шагов являются бинаризация и поиск ключевых точек, поскольку от качества их выполнения зависит качество работы всего алгоритма в целом. Рассмотрим данные шаги более подробно.

Для реализации этапа бинаризации выбран алгоритм Брэдли [10]. В нем пиксель изображения заменяется на черный в случае, если его яркость на t % меньше среднего значения яркости окружающих его пикселей (в противном случае цвет пикселя белый). Выбранный алгоритм бинаризации имеет два параметра, которыми можно варьировать, подбирая значения наилучшим образом подходящие для конкретной задачи – размер окна бинаризации (окрестность пикселя для определения среднего значения яркости) и порог бинаризации t . В результате настройки значений параметров, размер окна бинаризации равен 0,25 ширины изображения, порог бинаризации $t = 5$ %. В качестве алгоритма поиска угловых точек отобран детектор, основанный на локальных и глобальных свойствах кривизны [9]. Детектор анализирует границы изображения, причем каждая граница рассматривается как функция и для каждой функции строится соответствующая ей функция кривизны с последующим анализом точек экстремума. Параметром детектора поиска угловых точек является коэффициент R , необходимый для расчета адаптивного порога по которому будут определяться «сильные» углы. В статье с описанием детектора, основанного на локальных и глобальных свойствах кривизны [11], для поиска угловых точек на изображениях с большим количеством углов было использовано значение $R = 1,5$. Для нашей задачи эмпирическим путем было подобрано значение параметра $R = 1$. Наглядно работа алгоритма представлена на рис. 6.

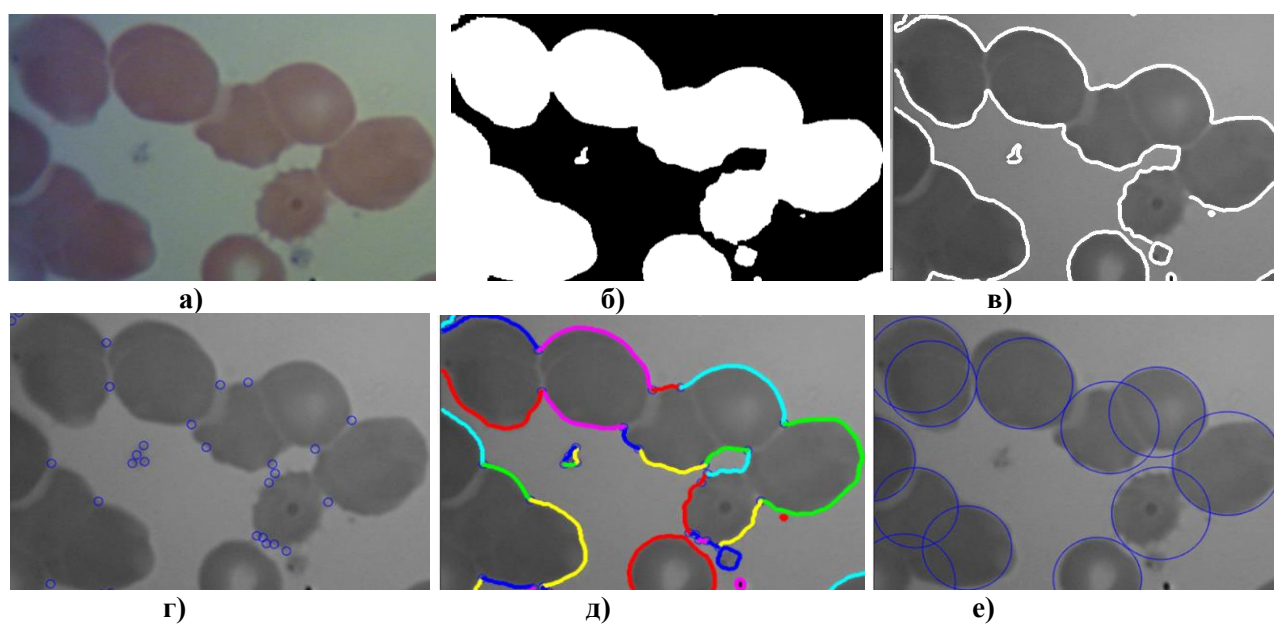


Рис. 6. Пример обработки фрагмента изображения микроскопии крови с использованием алгоритма разделения визуально пересекающихся эритроцитов

Fig. 6. An example of processing a blood microscopy image fragment using an algorithm for visually intersecting red blood cells separation

Результаты

Для оценки эффективности и сравнения нейросетевых архитектур с последующим применением алгоритма распознавания пересекающихся объектов использованы следующие показатели:

- количество эпох обучения – под эпохой подразумевается полный проход датасета через нейронную сеть;
- размер батча – показывает, на партии какого размера были разделены тренировочные данные при обучении;
- скорость обучения – размер шага, который алгоритм стохастического градиентного спуска использует для сходимости;
- *mAP (Mean Average Precision)* – средняя точность, позволяющая оценить, насколько хорошо модель распознает представленные объекты.

Результаты сравнения перечисленных архитектур в проведенных экспериментах представлены в табл. 1.

Таблица 1.
**Параметры и показатели эффективности нейросетевых архитектур
в задаче распознавания клеток крови**

Table 1.
**Metrics used to measure the performance of neural network architectures
in the task of blood cells recognition**

Нейросетевые архитектуры	Количество эпох обучения	Размер батча	Скорость обучения	mAP, %
FasterR-CNN	10 + 4	1	Первые 10: 0.001 и 4: 0.0001	87,3
YOLOv3tiny	250	48	0,00001	46,1
YOLOv4	250	32	0,001	90,7
YOLOv8n	300	динамический	0,01	93,7
YOLOv8s	600	динамический	0,01	95,7

По результатам работы наиболее высокие показатели были достигнуты с применением нейронной сети *YOLOv8*, которая является перспективной для проведения дальнейших исследований и реализации в системах лабораторной диагностики.

Заключение

Для решения задачи распознавания клеток крови на изображениях микроскопии выполнено исследование с применением и сравнением четырех нейросетевых архитектур: *FasterR-CNN*, *YOLOv3tiny*, *YOLOv4*, *YOLOv8*. Разработан алгоритм распознавания пересекающихся объектов, основанный на разделении контура объекта на сегменты по ключевым точкам. По результатам работы наиболее высокие показатели точности распознавания были достигнуты с применением нейронной сети *YOLOv8*, при этом средняя точность распознавания объектов составила 95,7 %.

Проведенные в рамках работы эксперименты показали эффективность применения нейросетевого подхода для анализа биомедицинских изображений.

Библиографический список

1. **Louka, M.** A microfluidic system, utilising image processing methods, for the detection of blood coagulation and erythrocyte aggregation / M. Louka, A. Passos, A. Inglezakis, C. Loizou, E. Kaliviotis // 2022 IEEE 5th International Conference on Image Processing Applications and Systems (IPAS). 2022. Pp. 1-6. DOI: 10.1109/IPAS55744.2022.10053010.
2. **Lee, S.J.** Complete Blood Cell Detection and Counting Based on Deep Neural Networks / S.J. Lee, P.Y. Chen, J.W. Lin // Applied Sciences. 2022. 12(16). Pp.1-16 DOI: <https://doi.org/10.3390/app12168140>
3. **Tavakoli, S.** New segmentation and feature extraction algorithm for classification of white blood cells in peripheral smear images / S. Tavakoli, A. Ghaffari, Z.M. Kouzehkanan, // Nature. Sci. Rep. 2021. № 11. Pp. 1-13 DOI: <https://doi.org/10.1038/s41598-021-98599-0>
4. **Kang, M.** CST-YOLO: A Novel Method for Blood Cell Detection Based on Improved YOLOv7 and CNN-Swin Transformer / M. Kang, C.M. Ting, F.F. Ting, R. Phan // Computer Vision and Pattern Recognition. 2023. Pp. 1-5. DOI: <https://doi.org/10.48550/arXiv.2306.14590>
5. **Ren, S.** Towards Real-Time Object Detection with Region Proposal Networks / S. Ren, K. He, R. Girshick, J. Sun // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2017. Vol. 39. №. 6. Pp. 1137-1149
6. **Аггарвал, Ч.** Нейронные сети и глубокое обучение / Ч. Аггарвал. – СПб.: Диалектика, 2020. – 752 с.
7. **Redmon J.** YOLOv3: An Incremental Improvement / J. Redmon, A. Farhadi // Computer Vision and Pattern Recognition. – 2018. – Pp.1-6. DOI: <https://doi.org/10.48550/arXiv.1804.02767>
8. **Bochkovskiy, A.** YOLOv4: Optimal Speed and Accuracy of Object Detection / A. Bochkovskiy, C.Y. Wang, H.M. Liao // Computer Vision and Pattern Recognition. 2020. Pp. 1-17.
9. **Liu, S.** ADA-YOLO: Dynamic Fusion of YOLOv8 and Adaptive Heads for Precise Image Detection and Diagnosis / S. Liu, J. Zhang, R. Song, T.T. Teoh // Computer Vision and Pattern Recognition. – 2023. – Pp. 1-16. DOI: <https://doi.org/10.48550/arXiv.2312.10099>
10. **Bradley, D.** Adapting Thresholding Using the Integral Image / D. Bradley, G. Roth // Journal of Graphics Tools. 2007. Vol. 12. №. 2. Pp. 13-21.
11. **Yung, N.H.C.** Corner detector based on global and local curvature properties // Optical Engineering. 2008. 47(5). Pp. 1-12.

*Дата поступления
в редакцию: 06.12.2023*

*Дата принятия
к публикации: 22.01.2024*