

ИНФОРМАТИКА И ИНФОРМАЦИОННЫЕ ПРОЦЕССЫ/INFORMATICS AND INFORMATION PROCESSES

DOI: <https://doi.org/10.60797/IRJ.2026.170.38> EDN: JDPDEB

СРАВНИТЕЛЬНЫЙ АНАЛИЗ РЕАЛИЗАЦИЙ МЕТОДОВ КОМПЬЮТЕРНОГО ЗРЕНИЯ ДЛЯ АВТОМАТИЗИРОВАННОГО ПОДСЧЕТА ОБЪЕКТОВ НА СИНТЕТИЧЕСКИХ ИЗОБРАЖЕНИЯХ С ВАРИАТИВНОЙ ПЛОТНОСТЬЮ И НАЛОЖЕНИЯМИ

Научная статья

Проценко И.Г.^{1,*}, Колесник Р.С.²^{1,2} Камчатский государственный технический университет, Петропавловск-Камчатский, Российская Федерация

* Корреспондирующий автор (ip1954[at]list.ru)

Предложена: 06.05.2026; Принята: 23.06.2026; Опубликовано: 17.08.2026

Аннотация

В работе представлен сравнительный анализ реализаций методов компьютерного зрения для автоматизированного подсчета объектов на синтетических изображениях с вариативной плотностью, размерами и частичным наложением объектов. Исследование ориентировано на прикладные задачи визуального мониторинга биоресурсов, где требуется оценивать количество объектов в условиях плотных сцен и неполной видимости отдельных экземпляров.

Рассмотрены четыре метода: сегментация связанных компонент, эвристический подсчет на основе угловых признаков Харриса, регрессионная сверточная нейронная сеть RegressCNN и модель обнаружения объектов YOLOv8n. Для обучения нейросетевых моделей использовался единый синтетический набор данных объемом 95 090 изображений. Итоговая проверка выполнялась на отдельной тестовой выборке объемом 6030 изображений, не входившей в обучающий набор и включавшей сцены с различной плотностью, пропорциями, разбросом размеров и наложениями объектов.

Основной метрикой сравнения выбрана средняя абсолютная ошибка подсчета (MAE), так как конечной задачей является оценка количества объектов на изображении. По результатам экспериментов наименьшая средняя ошибка получена при использовании YOLOv8n. Метод Харриса показал меньшую ошибку по сравнению с методом связанных компонент и RegressCNN, однако его результат основан на эвристическом пересчете угловых признаков в количество прямоугольных объектов. Метод связанных компонент и простая регрессионная модель RegressCNN продемонстрировали более выраженную зависимость ошибки от плотности, геометрии и наложений объектов.

Ключевые слова: компьютерное зрение, автоматический подсчет, синтетические данные, YOLOv8n, связанные компоненты, метод Харриса, средняя абсолютная ошибка (MAE).

A COMPARATIVE ANALYSIS OF COMPUTER VISION METHODS FOR THE AUTOMATED COUNTING OF OBJECTS ON SYNTHETIC IMAGES WITH VARIABLE DENSITY AND OVERLAYS

Research article

Procenko I.G.^{1,*}, Kolesnik R.S.²^{1,2} Kamchatka State Technical University, Petropavlovsk-Kamchatsky, Russian Federation

* Corresponding author (ip1954[at]list.ru)

Suggested: 06.05.2026; Accepted: 23.06.2026; Published: 17.08.2026

Abstract

The work presents a comparative analysis of implementations of computer vision methods for the automated counting of objects on synthetic images characterised by varying densities, sizes and partial occlusion of objects. The research focuses on applied tasks in the visual monitoring of biological resources, where it is necessary to evaluate the number of objects in dense scenes and where individual instances are not fully visible.

Four methods were examined: connected component segmentation, a heuristic approach based on Harris corner detector, the RegressCNN regression convolutional neural network, and the YOLOv8n object detection model. A single synthetic dataset comprising 95,090 images was used to train the neural network models. Final validation was carried out on a separate test set comprising 6,030 images, which was not included in the training set and contained scenes with varying densities, aspect ratios, size distributions and object overlaps.

The mean absolute error (MAE) was chosen as the primary comparison metric, as the ultimate objective is to assess the number of objects in an image. The experimental results showed that the lowest mean error was achieved using YOLOv8n. The Harris method showed a lower error compared with the connected components method and RegressCNN; however, its result is based on a heuristic conversion of angular features into the number of rectangular objects. The connected components method and the simple RegressCNN regression model demonstrated a more significant dependence of the error on the density, geometry and overlap of objects.

Keywords: computer vision, automatic counting, synthetic data, YOLOv8n, connected components, Harris method, mean absolute error (MAE).

Введение

Автоматизированный визуальный учет объектов является одной из прикладных задач компьютерного зрения [1], [2], [3]. Подобные методы могут применяться в системах мониторинга биоресурсов, аквакультуры, экологического наблюдения и других областях, где требуется получать количественную оценку объектов по изображениям или видеопотоку. В таких условиях ручной подсчет оказывается трудоемким, плохо масштабируется и зависит от субъективности оператора.

Для задач визуального мониторинга особую сложность представляют плотные сцены, в которых объекты расположены близко друг к другу, частично перекрываются, имеют различные размеры и пропорции. В реальных условиях эти трудности дополнительно усиливаются изменением освещения, фона, качества изображения и положения объектов относительно камеры. Поэтому перед практическим применением алгоритмов важно оценить, насколько разные методы подсчета устойчивы к изменению структуры сцены.

В данной работе рассматривается модельная задача подсчета объектов на синтетических изображениях. Такой подход позволяет контролировать параметры сцены и сравнивать методы в одинаковых условиях: при заданном количестве объектов, изменяемой плотности, масштабе, пропорциях и наличии наложений.

Данная статья является экспериментальным продолжением предварительного анализа подходов компьютерного зрения к задаче подсчета объектов. В рамках настоящей работы из рассмотренных подходов выбраны четыре реализации, различающиеся принципом получения оценки количества объектов: связные компоненты, метод Харриса, RegressCNN и YOLOv8n.

Одной из основных проблем автоматического подсчета является взаимное наложение объектов. В таких случаях несколько отдельных экземпляров могут восприниматься алгоритмом как одна связная область. Это особенно критично для классических методов обработки изображений, основанных на бинаризации, выделении контуров или подсчете связных компонент.

Другой важный фактор — изменение геометрических характеристик объектов. Если объекты отличаются по размеру, вытянутости или степени перекрытия, методы подсчета могут давать различные ошибки. Например, регрессионная нейросеть может учитывать не только количество объектов, но и суммарную площадь изображения, а методы локальных признаков могут зависеть от выраженности углов и границ.

В работе используется модельная постановка задачи подсчета объектов на синтетических изображениях. Синтетическая генерация данных позволяет заранее задавать и контролировать параметры сцены: количество объектов, плотность размещения, размеры, пропорции и наличие наложений. За счет этого становится возможным сравнить разные методы подсчета в воспроизводимых условиях и проанализировать, как изменение геометрии сцены влияет на ошибку.

Целью данной работы является сравнительная оценка реализаций методов компьютерного зрения для автоматизированного подсчета объектов на синтетических изображениях в условиях вариативной плотности, изменения геометрических характеристик и частичного наложения объектов.

Для достижения поставленной цели решаются следующие задачи:

1. Сформировать синтетический набор данных с заданными параметрами сцены: количеством объектов, размерами, пропорциями, разбросом размеров и степенью наложения.
2. Обучить нейросетевые модели на едином синтетическом наборе данных и выполнить итоговую проверку на отдельной тестовой выборке, не входившей в обучение.
3. Реализовать и сравнить четыре метода подсчета: метод связных компонент, метод угловых признаков Харриса, регрессионную сверточную нейронную сеть RegressCNN и модель обнаружения объектов YOLOv8n.
4. Оценить точность подсчета с использованием средней абсолютной ошибки MAE и проанализировать зависимость ошибки от параметров сцены.
5. Сформулировать ограничения рассмотренных методов и определить условия, при которых каждый подход дает наименьшую ошибку подсчета.

Методы и принципы исследования

Для проведения эксперимента использовалась модельная постановка задачи подсчета объектов на синтетических изображениях. В качестве объектов применялись прямоугольные геометрические примитивы. Такой выбор позволил заранее задавать параметры сцены и контролировать количество объектов, их размеры, пропорции, плотность размещения и наличие наложений.

Для обучения нейросетевых моделей был сформирован единый синтетический набор данных объемом 95 090 изображений размером 608×608 пикселей. Для итоговой проверки использовалась отдельная тестовая выборка объемом 6030 изображений, не входившая в обучающий набор. Тестовая выборка включала пустые сцены, изображения с одним или двумя крупными объектами, сцены без наложения, а также сцены с частичным наложением объектов. В линейных сериях количество объектов изменялось от 1 до 100, для каждого варианта использовалось два повтора.

Разметка данных формировалась автоматически в ходе процедурной генерации изображений. Для RegressCNN использовалось табличное значение количества объектов на изображении. Для YOLOv8n применялась разметка в виде координат ограничивающих рамок. Классические методы не требовали обучения и применялись непосредственно к изображениям тестовой выборки.

В работе сравнивались четыре реализации методов подсчета. Метод связных компонент выполнял бинаризацию изображения и считал количество изолированных областей. Метод Харриса использовал угловые признаки [7]; итоговая оценка количества объектов рассчитывалась как число найденных угловых точек, деленное на четыре. RegressCNN представляла собой сверточную регрессионную модель, предсказывающую количество объектов

напрямую. YOLOv8n [8], [9] после предсказания ограничивающих рамок количество объектов определялось по числу найденных объектов.

Регрессионная модель RegressCNN [10] включала три сверточных слоя с числом каналов 32, 64 и 128, после которых использовался полносвязный регрессионный блок. Выходной слой модели формировал одно числовое значение — прогнозируемое количество объектов на изображении. Для подхода на основе обнаружения объектов использовалась модификация YOLOv8n, обученная на том же синтетическом наборе данных, что и RegressCNN

Порядок эксперимента был одинаковым для всех методов. Каждое изображение тестовой выборки обрабатывалось четырьмя алгоритмами. Полученное значение количества объектов сравнивалось с эталонным значением, известным из процедуры генерации данных. Для каждого изображения вычислялась абсолютная ошибка подсчета, после чего рассчитывалось среднее значение ошибки по всей тестовой выборке.

Вычислительный эксперимент проводился на рабочей станции с графическим ускорителем AMD Radeon RX 7800 XT. Для выполнения тензорных вычислений использовалась библиотека DirectML, обеспечивающая работу с GPU через API DirectX 12 [13].

На рисунке 1 приведены примеры синтетических сцен, использованных при формировании тестовой выборки.

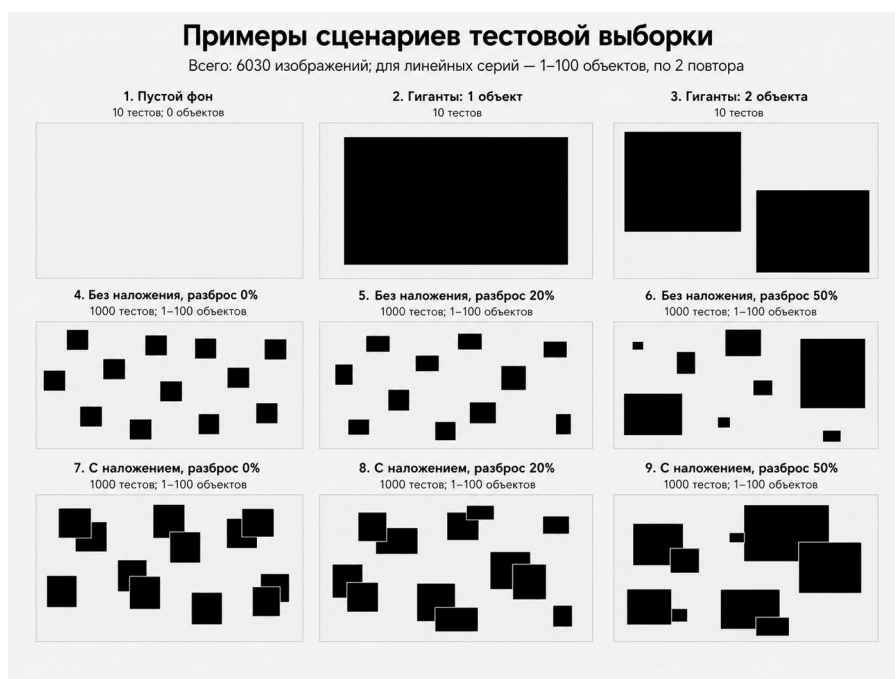


Рисунок 1 - Примеры синтетических сцен, использованных при формировании тестовой выборки
DOI: <https://doi.org/10.60797/IRJ.2026.170.38.1>

Результаты и обсуждение

В данном разделе приводится детальный анализ результатов работы четырех алгоритмов. Оценка производилась на тестовой выборке с использованием метрики средней абсолютной ошибки (Mean Absolute Error, MAE) [11], [12]:

$$MAE = \frac{1}{N} \sum_{i=1}^N |\hat{y}_i - y_i|$$

где $\hat{y}_i$ — предсказанное количество объектов на изображении, y_i — истинное количество объектов, N — количество изображений в тестовой выборке.

Для анализа результатов была рассмотрена зависимость ошибки подсчета от истинного количества объектов на изображении. Для каждого значения количества объектов рассчитывалось среднее значение абсолютной ошибки по соответствующим изображениям тестовой выборки.

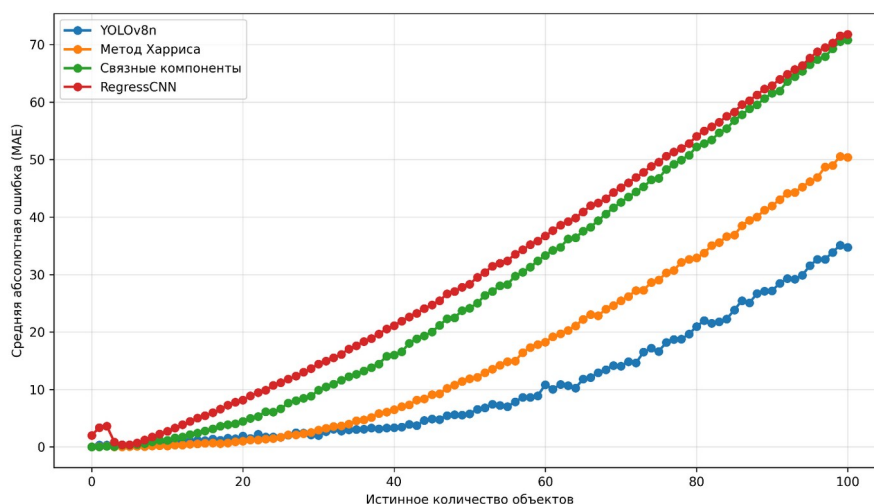


Рисунок 2 - Зависимость средней абсолютной ошибки подсчета от истинного количества объектов на изображении
DOI: <https://doi.org/10.60797/IRJ.2026.170.38.2>

Анализ графиков на рисунке 2 демонстрирует общую тенденцию к росту ошибки подсчета при увеличении числа объектов. При этом алгоритмы обладают различной устойчивостью к уплотнению сцены. Модель YOLOv8n сохраняет минимальную ошибку на всем исследуемом диапазоне. Метод Харриса показывает промежуточный результат, тогда как точность связных компонент и модели RegressCNN резко снижается на плотных сценах. Сводные значения средней абсолютной ошибки (MAE) представлены в таблице 1.

Таблица 1 - Сравнение методов по средней абсолютной ошибке подсчета

DOI: <https://doi.org/10.60797/IRJ.2026.170.38.3>

Метод	Средняя ошибка (MAE)	Интерпретация
YOLOv8n	10,47	Наименьшая ошибка среди рассмотренных методов
Метод Харриса	16,78	Приближенный подсчет через угловые признаки
Связные компоненты	28,20	Ошибка возрастает при плотном расположении и наложении объектов
RegressCNN	31,22	Прямая оценка количества без локализации объектов

Наименьшая средняя абсолютная ошибка получена при использовании YOLOv8n. Это связано с тем, что модель обнаружения объектов формирует отдельные предсказания для объектов, а итоговый подсчет выполняется по числу найденных ограничивающих рамок.

Метод Харриса показал меньшую ошибку по сравнению со связными компонентами и RegressCNN. В данной постановке это объясняется тем, что угловые признаки сохраняют часть информации о количестве прямоугольных объектов. Вместе с тем результат метода остается приближенным, поскольку алгоритм выделяет угловые точки, а не объекты целиком.

Метод связных компонент оказался чувствителен к наложению объектов. При соприкосновении или перекрытии нескольких прямоугольников они могут восприниматься как одна связная область, что приводит к занижению результата подсчета. RegressCNN также показала высокую ошибку, поскольку модель предсказывает итоговое количество напрямую и не выполняет локализацию отдельных экземпляров.

На рисунке 3 обобщены основные факторы структуры сцены, влияющие на ошибку подсчета.

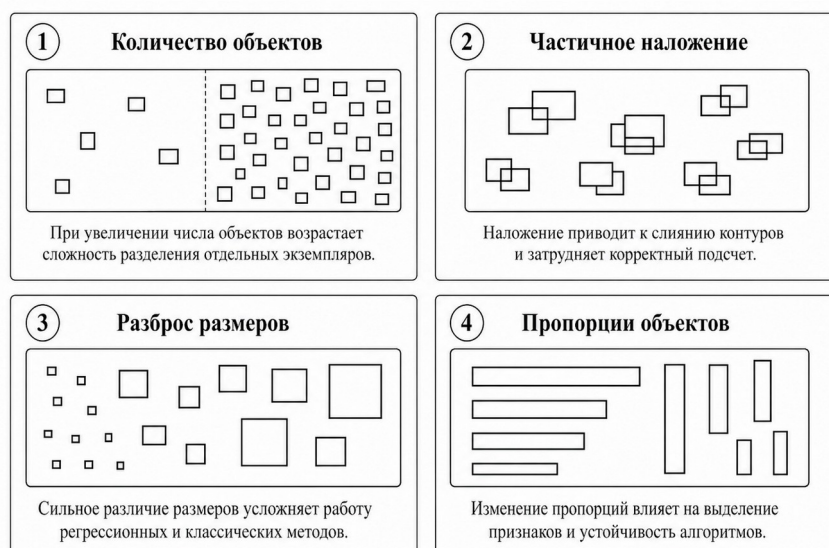
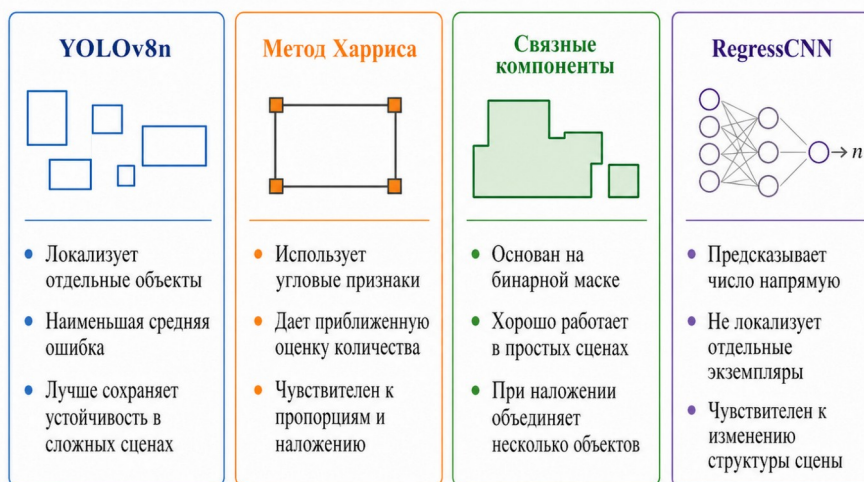


Рисунок 3 - Факторы структуры сцены, влияющие на ошибку подсчета

DOI: <https://doi.org/10.60797/IRJ.2026.170.38.4>

Представленные факторы по-разному влияют на рассматриваемые методы. Увеличение количества объектов и плотности сцены повышает вероятность соприкосновения и перекрытия прямоугольников. Разброс размеров и изменение пропорций дополнительно усложняют выделение устойчивых признаков. Поэтому средняя ошибка зависит не только от выбранного алгоритма, но и от структуры входного изображения.



Сравнение качественных свойств методов по результатам экспериментального исследования.

Рисунок 4 - Характерные особенности и ограничения рассматриваемых методов

DOI: <https://doi.org/10.60797/IRJ.2026.170.38.5>

Таким образом, в условиях проведенного эксперимента наиболее устойчивым оказался подход на основе обнаружения объектов YOLOv8n. Метод Харриса занял промежуточное положение, поскольку позволяет получить приближенную оценку количества объектов через угловые признаки. Связные компоненты и RegressCNN сильнее зависят от структуры сцены: первый метод — от слияния объектов, второй — от отсутствия локализации отдельных экземпляров.

Эксперимент проводился на синтетических изображениях, поэтому полученные результаты следует рассматривать как предварительную оценку поведения методов в контролируемых условиях.

Заключение

В работе выполнено сравнение четырех реализаций методов автоматизированного подсчета объектов на синтетической тестовой выборке. По результатам эксперимента наименьшую среднюю абсолютную ошибку показал YOLOv8n — MAE = 10,47. Метод Харриса показал MAE = 16,78, метод связанных компонент — MAE = 28,20, RegressCNN — MAE = 31,22.

Полученные результаты показывают, что в условиях вариативной плотности, изменения размеров и частичного наложения объектов подход на основе обнаружения объектов оказался наиболее устойчивым среди рассмотренных методов. Метод Харриса может использоваться как приближенный классический способ подсчета, но его точность зависит от качества выделения угловых признаков. Связные компоненты и простая регрессионная модель сильнее зависят от структуры сцены.

Практическая значимость исследования заключается в разработке методики предварительного отбора алгоритмов компьютерного зрения для задач учета биоресурсов. Оценка алгоритмов на синтетических данных позволяет выявить их пределы устойчивости к наложению объектов и изменению плотности сцены, что дает возможность отсеять неэффективные подходы до перехода к работе с реальными изображениями. Такой подход существенно снижает затраты времени и вычислительных ресурсов при проектировании систем автономного визуального мониторинга.

Конфликт интересов

Не указан.

Conflict of Interest

None declared.

Рецензия

Баленко Е.Г., Донской государственный аграрный университет, Персиановский Российская Федерация
DOI: <https://doi.org/10.60797/IRJ.2026.170.38.6>

Review

Balenko E.G., Don State Agrarian University, Persianovsky Russian Federation
DOI: <https://doi.org/10.60797/IRJ.2026.170.38.6>

Список литературы / References

1. Андриянов Н.А. Обнаружение объектов на изображении: от критериев Байеса и Неймана–Пирсона к детекторам на базе нейронных сетей EfficientDet / Н.А. Андриянов, В.Е. Деметьев, А.Г. Ташлинский // Компьютерная оптика. — 2022. — Т. 46. — № 1. — С. 139–159.
2. Калиновский И.А. Обзор и тестирование детекторов фронтальных лиц / И.А. Калиновский, В.Г. Спицын // Компьютерная оптика. — 2016. — Т. 40. — № 1. — С. 99–111.
3. Калинина А.П. Детектирование книг на книжных полках при помощи глубоких нейронных сетей / А.П. Калинина, П.П. Николаев // Компьютерная оптика. — 2020. — Т. 44. — № 6. — С. 968–977.
4. Васильев М.Е. Обзор версий YOLO: одноэтапная модель сверточной нейронной сети / М.Е. Васильев, А.С. Шалимов, О.А. Савина // Universum: технические науки. — 2025. — № 6 (135).
5. Гафаров Ф.М. Сравнение эффективности алгоритмов YOLOv5 и YOLOv8 / Ф.М. Гафаров, А.Ф. Галимянов, Р.Р. Зиннатуллин // интеллектуальные системы и микросистемы. — 2023. — С. 124–131.
6. Морозов А.А. Анализ видеоизображений в реальном времени средствами языка Акторный Пролог / А.А. Морозов, О.С. Сушкова // Компьютерная оптика. — 2016. — Т. 40. — № 6. — С. 947–957. — DOI: 10.18287/2412-6179-2016-40-6-947-957.
7. Harris C. A Combined Corner and Edge Detector / C. Harris, M. Stephens. — 1988.
8. Redmon J. You Only Look Once: Unified, Real-Time Object Detection / J. Redmon, S. Divvala, R. Girshick [et al.] // CVPR — 2016.
9. YOLOv8 Documentation/Repository / Ultralytics. — 2023. — URL: <http://github.com/ultralytics> (accessed: 01.03.2026).
10. Paszke A. PyTorch: An Imperative Style, High-Performance Deep Learning Library / A. Paszke [et al.] // NeurIPS. — 2019.
11. Willmott C.J. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance / C.J. Willmott, K. Matsuura // Climate Research. — 2005. — Vol. 30. — P. 79–82.
12. Chai T. Root mean square error (RMSE) or mean absolute error (MAE)? — Arguments against avoiding RMSE in the literature / T. Chai, R.R. Draxler // Geoscientific Model Development. — 2014. — № 7. — P. 1247–1250. — DOI: 10.5194/gmd-7-1247-2014.
13. DirectML documentation / Microsoft. — URL: <https://learn.microsoft.com/en-us/windows/ai/directml/dml> (accessed: 01.02.2026).

Список литературы на английском языке / References in English

1. Andriyanov N.A. Obnaruzhenie obektov na izobrazhenii: ot kriteriev Baiesa i Neimana–Pirsona k detektoram na baze neironnikh setei EfficientDet [Object Detection in Images: From the Bayes and Neyman–Pearson Criteria to EfficientDet Neural Network-Based Detectors] / N.A. Andriyanov, V.E. Dementev, A.G. Tashlinskii // Kompyuternaya optika [Computer Optics]. — 2022. — Vol. 46. — № 1. — P. 139–159. [in Russian]
2. Kalinovskii I.A. Obzor i testirovanie detektorov frontalnikh lits [Review and testing of frontal face detectors] / I.A. Kalinovskii, V.G. Spitsin // Kompyuternaya optika [Computer Optics]. — 2016. — Vol. 40. — № 1. — P. 99–111. [in Russian]
3. Kalinina A.P. Detektirovanie knig na knizhnikh polkakh pri pomoshchi glubokikh neironnikh setei [Detection of Books on Bookshelves Using Deep Neural Networks] / A.P. Kalinina, P.P. Nikolaev // Kompyuternaya optika [Computer Optics]. — 2020. — Vol. 44. — № 6. — P. 968–977. [in Russian]



4. Vasilev M.E. Obzor versii YOLO: odnoetapnaya model svertochnoi neironnoi seti [An Overview of YOLO Versions: A Single-Stage Convolutional Neural Network Model] / M.E. Vasilev, A.S. Shalimov, O.A. Savina // *Universum: tekhnicheskie nauki* [Universum: Technical Sciences]. — 2025. — № 6 (135). [in Russian]
5. Gafarov F.M. Sravnenie effektivnosti algoritmov YOLOv5 i YOLOv8 [A Comparison of the Performance of the YOLOv5 and YOLOv8 Algorithms] / F.M. Gafarov, A.F. Galimyanov, R.R. Zinnatullin // *intellektualnie sistemi i mikrosistemi* [Intelligent Systems and Microsystems]. — 2023. — P. 124–131. [in Russian]
6. Morozov A.A. Analiz videoizobrazhenii v realnom vremeni sredstvami yazika Aktornii Prolog [Real-time analysis of video images using the Actor Prolog language] / A.A. Morozov, O.S. Sushkova // *Kompyuternaya optika* [Computer Optics]. — 2016. — Vol. 40. — № 6. — P. 947–957. — DOI: 10.18287/2412-6179-2016-40-6-947-957. [in Russian]
7. Harris C. A Combined Corner and Edge Detector / C. Harris, M. Stephens. — 1988.
8. Redmon J. You Only Look Once: Unified, Real-Time Object Detection / J. Redmon, S. Divvala, R. Girshick [et al.] // *CVPR* — 2016.
9. YOLOv8 Documentation/Repository / Ultralytics. — 2023. — URL: <http://github.com/ultralytics> (accessed: 01.03.2026).
10. Paszke A. PyTorch: An Imperative Style, High-Performance Deep Learning Library / A. Paszke [et al.] // *NeurIPS*. — 2019.
11. Willmott C.J. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance / C.J. Willmott, K. Matsuura // *Climate Research*. — 2005. — Vol. 30. — P. 79–82.
12. Chai T. Root mean square error (RMSE) or mean absolute error (MAE)? — Arguments against avoiding RMSE in the literature / T. Chai, R.R. Draxler // *Geoscientific Model Development*. — 2014. — № 7. — P. 1247–1250. — DOI: 10.5194/gmd-7-1247-2014.
13. DirectML documentation / Microsoft. — URL: <https://learn.microsoft.com/en-us/windows/ai/directml/dml> (accessed: 01.02.2026).