КОМПЬЮТЕРНЫЕ ИССЛЕДОВАНИЯ
И МОДЕЛИРОВАНИЕ 2026 Т. 18 № 5 С. 1249–1268
DOI: 10.20537/2076-7633-2026-18-5-1249-1268

Ки&М

МОДЕЛИ В ФИЗИКЕ И ТЕХНОЛОГИИ

УДК: 621.396.96

Детектирование малоразмерных объектов на радиолокационных изображениях

В. И. Вебер[1,a], Д. Д. Иванов[1], В. Ю. Куприц[1], Ю. А. Светличный[2]

[1]Томский государственный университет систем управления и радиоэлектроники,
Россия, 634050, г. Томск, пр. Ленина, д. 40
[2]ООО «ЛЭМЗ-Т»,
Россия, 634055, г. Томск, пр. Развития, д. 8

E-mail: [a] vladweber00@gmail.com

Получено 29.04.2026, после доработки — 05.10.2026.
Принято к публикации 06.10.2026.

В работе исследовано совместное влияние формата аннотации объектов и разрешения входных изображений на качество бинарной детекции малоразмерных наземных объектов на радиолокационных изображениях. Выполнен сравнительный анализ моделей семейства YOLO (v5, v8, v11, v26) и специализированных архитектур ориентированного детектирования, включая Rotated RetinaNet, R3Det, ROI Transformer, Gliding Vertex и другие модели. Для исследования сформирован набор данных binMSTAR на основе открытой базы mixMSTAR, содержащий аннотации в форматах горизонтальных (HBB) и ориентированных (OBB) ограничивающих рамок для 9 вариантов разрешения входных кадров от 640×640 до 1408×1408 пикселей.

Статистическая значимость различий оценивалась с использованием непараметрического U-критерия Манна – Уитни. Установлено, что применение OBB обеспечивает статистически значимое преимущество по метрике $mAP_{50–95}$, в среднем на 6,5 % выше по сравнению с HBB, за счет уменьшения влияния избыточного фонового шума в области разметки. Определен оптимальный диапазон размера входного кадра 896–1024 пикселя. Дальнейшее увеличение разрешения снижает точность локализации из-за перенасыщения признакового описания фоновыми областями. Наилучшие метрики показали модели семейства YOLO, а именно модель YOLOv11(m), для которой при разрешении 960×960 пикселей получено значение $mAP_{50–95} = 0{,}87$. Специализированные OBB-детекторы при разрешении 1024×1024 пикселя продемонстрировали сопоставимую полноту обнаружения с Recall до 0,997. Полученные результаты обосновывают целесообразность использования ориентированных рамок и выбора входного разрешения в диапазоне 896–1024 пикселя для задач бинарной детекции малоразмерных объектов на РЛИ.

Ключевые слова: радиолокационное изображение, радиолокатор с синтезированием апертуры, детекция объектов, модели YOLO, малоразмерный объект, binMSTAR, mixMSTAR, горизонтальные ограничивающие рамки, ориентированные ограничивающие рамки, входное разрешение, глубокое обучение

© 2026 Владислав Игоревич Вебер, Данил Дмитриевич Иванов, Владимир Юрьевич Куприц, Юрий Алексеевич Светличный
Статья доступна по лицензии Creative Commons Attribution-NoDerivs 3.0 Unported License.
Чтобы получить текст лицензии, посетите веб-сайт http://creativecommons.org/licenses/by-nd/3.0/
или отправьте письмо в Creative Commons, PO Box 1866, Mountain View, CA 94042, USA.

COMPUTER RESEARCH AND MODELING
2026 VOL. 18 NO. 5 P. 1249–1268
DOI: 10.20537/2076-7633-2026-18-5-1249-1268

КиМ

MODELS IN PHYSICS AND TECHNOLOGY

UDC: 621.396.96

Small object detection in radar images

V. I. Weber[1,a], D. D. Ivanov[1], V. Yu. Kuprits[1], Yu. A. Svetlichny[2]

[1]Tomsk State University of Control System and Radioelectronics,
40 Lenina ave., Tomsk, 634050, Russia
[2]LEMZ-T Ltd.,
8 Razvitiya ave., Tomsk, 634055, Russia

E-mail: [a] vladweber00@gmail.com

Received 29.04.2026, after completion — 05.10.2026.
Accepted for publication 06.10.2026.

This study examines how object annotation format and input image resolution jointly affect the quality of binary detection of small objects in radar images. The study aims to compare YOLO family models, including YOLOv5, YOLOv8, YOLOv11, and YOLOv26, with specialized oriented object detection architectures such as Rotated RetinaNet, R3Det, ROI Transformer, Gliding Vertex, and related models in the task of binary detection of small ground objects in radar imagery. The study also aims to identify data preparation parameters that improve object localization accuracy.

For the experimental analysis, we formed a new binMSTAR dataset from the open mixMSTAR database. The dataset contains annotations in horizontal bounding box (HBB) and oriented bounding box (OBB) formats for nine input frame resolutions ranging from 640×640 to 1408×1408 pixels. We assessed the statistical significance of the observed differences using the nonparametric Mann–Whitney U test. The results show that OBB annotation provides a statistically significant advantage in terms of $mAP_{50–95}$, with an average improvement of 6.5 %, because it reduces the influence of redundant background noise. The experiments also reveal an optimal input frame size range of 896–1024 pixels. Further resolution growth decreases localization accuracy, since excessive background content weakens the object-related feature representation. YOLO models demonstrate the best overall balance of detection quality, with YOLOv11(m) reaching a peak $mAP_{50–95}$ of 0.87 at a resolution of 960 pixels. At the same time, specialized OBB detectors at a resolution of 1024 pixels achieve comparable recall values, reaching up to 0.997. These results support the use of oriented bounding boxes and input resolutions within the 896–1024 pixel range for binary detection tasks in radar imagery.

Keywords: radar image, synthetic aperture radar, object detection, YOLO models, small object, binMSTAR, mixMSTAR, horizontal bounding boxes, oriented bounding boxes, input resolution, deep learning

Citation: *Computer Research and Modeling*, 2026, vol. 18, no. 5, pp. 1249–1268 (Russian).

© 2026 Vladislav I. Weber, Danil D. Ivanov, Vladimir Yu. Kuprits, Yuri A. Svetlichny
This work is licensed under the Creative Commons Attribution-NoDerivs 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nd/3.0/
or send a letter to Creative Commons, PO Box 1866, Mountain View, CA 94042, USA.

1. Введение

На радиолокационных изображениях (РЛИ) целевые объекты могут занимать лишь несколько десятков пикселей при общем размере кадра в десятки тысяч пикселей. Кроме того, полезный сигнал может интенсивно отражаться от подстилающей поверхности, что создает значительные помехи и приводит к высокому уровню ложных срабатываний. В данной работе решается задача бинарной детекции наземных транспортных средств на РЛИ. Под бинарной детекцией понимается автоматическое обнаружение объектов, которое рассматривается в постановке бинарной классификации («объект» или «фон») с определением координат объектов на изображении. Вопросы последующей категориальной классификации вынесены за рамки текущего исследования и рассматриваются как отдельный этап обработки. Отказ от категориальной классификации на текущем этапе обусловлен как спецификой радиолокационных изображений (малоразмерность объектов, размер менее 1 % относительно всего изображения), так и методологической задачей исследования: провести сравнительный анализ эффективности различных архитектур нейронных сетей.

Системы компьютерного зрения (СКЗ) на основе анализа данных радиолокаторов с синтезированной апертурой (РСА) являются критически важным инструментом для всепогодного мониторинга земной поверхности. Прикладная актуальность таких систем подтверждается их широким применением в поисково-спасательных операциях [Gao et al., 2025; Yin et al., 2026; Красноперова и др., 2025], мониторинге инфраструктуры [Li et al., 2024] и обеспечении безопасности морского и наземного пространств [Han et al., 2024; Купряшкин, 2025; Купряшкин, 2022]. В этих работах авторы не представляют обоснование. Однако, несмотря на высокую практическую ценность существующих решений, в научно-технической литературе наблюдается ряд методологических упущений. В частности, в большинстве упомянутых работ авторы не акцентируют внимание на обосновании выбора типа аннотации объектов, будь то стандартные горизонтальные рамки (Horizontal Bounding Box, HBB) или ориентированные рамки (Oriented Bounding Box, OBB). Вместо сравнительного анализа эффективности того или иного подхода исследователи, как правило, используют один тип разметки, предопределенный структурой выбранного набора данных. Это ограничивает понимание того, как геометрическая специфика аннотаций влияет на точность локализации малоразмерных или наклонных объектов на РЛИ. Аналогичная тенденция прослеживается и в отношении пространственных параметров обрабатываемых данных. Авторы используют фиксированные габариты изображений. Такой унифицированный подход игнорирует тот факт, что изменение масштаба и разрешения входных данных может критически влиять на качество обнаружения малоразмерных объектов.

В этой статье рассматривается как семейство моделей YOLO (v5, v8, v11, v26) [Jocher et al., 2026], зарекомендовавшее себя в задачах реального времени, так и специализированные алгоритмы для ориентированных объектов, такие как R3Det и Rotated RetinaNet и др. архитектуры, описанные в [Liu et al., 2023]. В работе приводится сравнение подходов с использованием горизонтальных и ориентированных ограничивающих рамок, что позволяет оценить целесообразность усложнения архитектуры модели для более точного описания геометрии малоразмерных объектов на РЛИ. Кроме этого, особое внимание в работе уделено влиянию входного разрешения на качество формирования карт признаков. Целью исследования является сравнительный анализ различных архитектур нейронных сетей в задаче обнаружения малоразмерных объектов на РЛИ для выявления наиболее перспективных решений.

2. Обнаружение наземных малоразмерных объектов на РЛИ с РСА

В данном исследовании используется сформированная авторами база данных binMSTAR. Эта база данных является модификацией mixMSTAR для решения задачи бинарной детекции.

Научная новизна работы заключается в апробации современных архитектур нейронных сетей на базе данных binMSTAR, что позволяет оценить их потенциал для выявления целей в условиях сложной помеховой обстановки без внесения погрешности, связанной с категориальной классификацией. Подробно об используемой базе данных описано в § 4.

Одним из ключевым аспектом исследования выступает сопоставление точности двух подходов к локализации: с использованием горизонтальных и ориентированных ограничивающих рамок. В первом случае объект описывается вектором $B = \{x, y, w, h\}$, где (x, y) — координаты центра, а w, h — размеры прямоугольника, стороны которого параллельны осям координат. При использовании формата OBB в вектор параметров вводится угол ориентации θ относительно горизонтальной оси изображения: $B = \{x, y, w, h, \theta\}$.

Формирование итогового РЛИ представляет собой обработку массива данных большой размерности. Важным этапом является фрагментация исходного массива на рабочие кадры фиксированного размера $S \times S$. Таким образом, существует возможность формирования РЛИ разных размеров при сохранении размеров объектов.

Пусть d — линейный размер объекта интереса в пикселях, $A_c = A_i - A_o$ — площадь контекстной области, $A_i = S^2$ — площадь анализируемого РЛИ, $A_o \approx d^2$ — площадь занимаемая объектом. При фиксированном разрешении на местности значение d остается инвариантным относительно S. Однако выбор размера S критически определяет геометрический контекст и соотношение «сигнал/фон». Увеличение S влияет на следующие факторы.

1. Увеличение контекста. С ростом S расширяется площадь A_c, содержащая окружение объекта. Контекст несет дополнительную информацию о структуре сцены и взаимосвязях объектов, что повышает различимость и способствует росту качества детекции.

2. Избыточность фоновой информации. Одновременно с ростом S увеличивается доля областей, не содержащих полезной информации. При фиксированном размере объекта это приводит к снижению относительной доли полезного сигнала объекта и росту отрицательных примеров, что ухудшает обучение модели и снижает точность детектирования.

Количественная оценка эффективности исследуемых моделей проводится на основе стандартного аппарата метрик детекции объектов: точности (Precision), полноты (Recall) и усредненной точности (mean average precision, mAP). В качестве агрегированных критериев качества рассматриваются показатели mAP_{50} и $\text{mAP}_{50–95}$, рассчитываемые для фиксированного порога метрики пересечения областей (Intersection over Union, IoU), равного 0,5, а также путем усреднения значений в диапазоне от 0,5 до 0,95 с шагом 0,05 соответственно.

В системах мониторинга РЛИ при решении задачи бинарной детекции стоимость пропуска объекта (False Negative, FN) критически превышает стоимость ложной тревоги, поскольку ложная тревога может быть отфильтрована на последующих стадиях анализа данных, в то время как пропущенный объект теряется безвозвратно. В связи с этим при сравнении моделей приоритетным критерием считается максимизация показателя Recall как основного индикатора надежности обнаружения малоразмерных объектов. При этом максимизация Recall не должна достигаться за счет неконтролируемого снижения Precision. Приоритет показателя Recall дополняется анализом $\text{mAP}_{50–95}$, который выступает индикатором строгой локализации. Для малоразмерных объектов на РЛИ это позволяет оценить, насколько эффективно модель подавляет влияние фона при разных размерах S.

Таким образом, в основу данного исследования положены следующие гипотезы.

1. Применение горизонтальных ограничивающих рамок приводит к снижению метрики $\text{mAP}_{50–95}$ в сравнении с ориентированными рамками. Данный эффект объясняется включением в область аннотирования избыточной фоновой составляющей, которая вносит дополнительный шум в вектор признаков объекта и затрудняет процесс обучения модели.

2. Эффективность распознавания определяется балансом между семантически значимым контекстом и избыточным фоном. Это обуславливает существование оптимального пиксельного разрешения РЛИ, при котором максимизируется метрика $mAP_{50–95}$.

3. Модели для обнаружения наземных объектов на РЛИ с РСА

Для проведения сравнительного анализа в настоящем исследовании используются архитектуры нейронных сетей, реализованные в рамках открытых программных библиотек. Оценка эффективности специализированных алгоритмов ориентированного детектирования проводится на базе библиотеки MMRotate [Zhou et al., 2022], в то время как исследование эволюционной линейки универсальных детекторов семейства YOLO (v5, v8, v11, v26) реализовано с использованием библиотеки Ultralytics.

Все исследуемые модели разделены на две крупные методологические группы:

1) специализированные алгоритмы для ориентированных ограничивающих рамок;

2) алгоритмы семейства YOLO.

3.1. Специализированные алгоритмы для ориентированных ограничивающих рамок

В этот раздел входят модели, математический аппарат которых изначально спроектирован для учета угловой ориентации объектов. В соответствии с принятой в литературе классификацией их можно разделить на четыре типа [Liu et al., 2023].

1. Одноэтапные алгоритмы (One-Stage): напрямую предсказывают класс и координаты ограничивающих рамок по картам признаков, что делает их вычислительно эффективными.

2. Двухэтапные алгоритмы (Two-Stage): разделяют процессы локализации и классификации. Сначала сеть выделения предложений (например, Region Proposal Network, RPN) генерирует потенциальные области интереса, которые затем классифицируются на втором этапе. Они обычно точнее, но работают медленнее.

3. Алгоритмы с этапом уточнения (S2A-Net, R3Det): используют вспомогательные регрессоры для итеративной корректировки положения и габаритов рамок.

4. Безъякорные методы (Rotated FCOS, Oriented RepPoints): предсказывают координаты напрямую, избавляясь от ручного подбора параметров эталонных рамок.

В работе Retinanet [Lin et al., 2017] утверждается, что основная причина, по которой одноэтапные детекторы уступают двухэтапным моделям, заключается в резком дисбалансе между объектами переднего плана и фоном во время обучения. Для решения этой проблемы была предложена функция Focal Loss, которая добавляет два коэффициента к бинарной кросс-энтропийной функции потерь, чтобы сбалансировать важность положительных и отрицательных выборок. Кроме того, она снижает значимость простых выборок (изображения фона), чтобы обучение было сосредоточено на сложных малочисленных примерах (изображения объекта). Retinanet — первая одноэтапная модель, по точности превосходящая двухэтапные модели. На ее основе была создана модель Rotated Retinanet, которая предсказывает дополнительный угол в регрессионной ветви (x, y, w, h).

S2A-Net [Han et al., 2021] — это модель с этапом уточнения, в которой используется модуль выравнивания признаков (Feature Alignment Module, FAM) на основе усовершенствованной деформируемой свертки (Deformable Convolution, DCN). На этапе уточнения горизонтальная якорная рамка преобразуется в повернутую якорную рамку с помощью сети уточнения якорей (Anchor Refinement Network, ARN). Это обучаемый модуль уточнения, который напрямую контролируется с помощью аннотаций. Затем карта признаков внутри якоря выравнивается, после чего на нее накладывается свертка с выравниванием (Alignment Convolution, ACL).

R3Det [Dai et al., 2017] — это модель с этапом уточнения, в которой используется модуль уточнения признаков (Feature Refinement Module, FRM) для восстановления карты признаков

в соответствии с уточненной ограничительной рамкой. Каждая точка на восстановленной карте признаков получается путем добавления пяти векторов признаков, состоящих из пяти точек (четыре угловые точки и центральная точка в уточненной ограничительной рамке), после интерполяции. Модуль FRM позволяет решить проблему несовпадения признаков, характерную для уточненных одноэтапных детекторов, и может быть добавлен несколько раз для повышения эффективности. Кроме того, предложена приблизительная функция потерь SkewIoU, которая лучше отражает реальную функцию потерь SkewIoU, сохраняя при этом дифференцируемость.

ROI Transformer [Ding et al., 2019] — это двухэтапная модель, которая добавляет обучаемый модуль для преобразования горизонтальной области интереса (Horizontal Region of interest, HRoI) в повернутую область интереса (Rotated RoI, RRoI). Модель генерирует HRoI на основе небольшого количества горизонтальных якорных рамок и формирует RRoI, смещая повернутую эталонную область относительно HRoI. Эта операция избавляет от необходимости заранее задавать большое количество повернутых якорных рамок с разными углами для непосредственного создания RRoI. На следующем этапе предложенная модель извлекает инвариантные к повороту признаки из карты признаков и RRoI для улучшения последующей классификации и регрессии.

Oriented RCNN [Xie et al., 2021] основан на Faster RCNN [Ren et al., 2015] и предлагает эффективную сеть с ориентированной регрессией положения вершин (RPN). Ориентированная RPN использует новое представление смещения средней точки с шестью параметрами для отображения смещений повернутых эталонных объектов относительно горизонтальной якорной рамки и генерации четырехугольных предложений. По сравнению с RRPN [Ma et al., 2018] эта сеть позволяет избежать огромного количества вычислений, связанных с предварительным созданием большого количества повернутых якорных рамок. По сравнению с ROI Transformer она преобразует горизонтальные якорные рамки в ориентированные предложения за один шаг, значительно сокращая количество параметров сети RPN.

В модели Gliding Vertex [Xu et al., 2021] используется надежное представление OBB, которое позволяет устранить ограничения, связанные с предсказанием вершин и углов. В частности, в регрессионной ветви используются четыре дополнительных параметра соотношения длин для сдвига соответствующей вершины с каждой стороны горизонтальной ограничительной рамки. Такой подход позволяет избежать неоднозначность с порядком при непосредственном предсказании положения четырех вершин. Кроме того, он снижает чувствительность к ошибкам при предсказании угла. Помимо этого, авторы ввели параметр соотношения площадей r для предсказания наклона рамки.

В таблице 1 приведены краткое описание и сравнение специализированных алгоритмов. Здесь 1 GFLOPs — это число операций с плавающей точкой, выраженное в миллиардах операций. Скорость обработки FPS (кадров в секунду) приведена авторами [Liu et al., 2023] для моделей, реализованных в MMRotate, при тестировании на вычислительном стенде с GPU NVIDIA GeForce RTX 3090 с 24 ГБ видеопамяти.

Для всех девяти моделей в статье [Liu et al., 2023] использовались единые условия обучения, чтобы результаты можно было сопоставлять напрямую. Эксперименты выполнены в MMRotate на Ubuntu 20.04.4, Python 3.8.10, PyTorch 1.11.0 и CUDA 11.3. В качестве основной сети для всех детекторов применялась ResNet50 с предобучением на ImageNet, а в качестве шеи сети использовалась FPN. Входные изображения обучающей выборки разрезались на фрагменты 1024×1024 пикселя. Тестовые крупные сцены также делились на фрагменты 1024×1024 пикселя, но с шагом 824 пикселя, то есть с перекрытием соседних окон. Для аугментации применялось случайное отражение по горизонтали, вертикали или диагонали с вероятностью 0,25. Обучение проводилось в течение 180 эпох при размере пакета 2 изображения. В качестве оптимизатора использовался SGD с начальной скоростью обучения 0,005, и коэффициент весового

Таблица 1. Описание и сравнение специализированных алгоритмов

Модель	Тип модели	Ключевые особенности	Количество параметров, млн	GFLOPs	FPS
Rotated RetinaNet	Одноэтапная	Использует Focal Loss; к ветви регрессии добавлен параметр угла, поэтому модель предсказывает x, y, w, h и ориентацию объекта	36,74	218,18	29,2
S2A-Net	Модель с уточняющим этапом	Использует Feature Alignment Module; горизонтальный якорь уточняется до ориентированного якоря через Anchor Refinement Network; применяется Alignment Convolution	38,85	198,12	26,3
R3Det	Модель с уточняющим этапом	Использует Feature Refinement Module; уточнение признаков выполняется по пяти точкам ориентированной рамки, включая четыре угла и центр; применяется приближенная SkewIoU loss	37,52	235,19	26,1
ROI Transformer	Двухэтапная	Преобразует горизонтальную область интереса HRoI в ориентированную RRoI; используется Rotated Position Sensitive RoI Align	55,39	225,32	25,3
Oriented RCNN	Двухэтапная	Построена на Faster RCNN; использует oriented RPN и шестипараметрическое представление смещения средних точек для формирования ориентированных предложений	41,37	211,44	26,5
ReDet	Двухэтапная	Использует ReResNet с вращательно-эквивариантными признаками и Rotation-Invariant RoI Align для получения признаков, инвариантных к повороту	31,70	54,48	18,4
Gliding Vertex	Двухэтапная	Ориентированная рамка задается через четыре коэффициента смещения вершин и параметр отношения площадей r, что снижает чувствительность к углу	41,37	211,31	28,5
Rotated FCOS	Безъякорная	Одноэтапная полносверточная схема без заранее заданных якорей; предсказывает расстояния l, r, t, b, centerness и дополнительный угол	32,16	207,16	29,7
Oriented RepPoints	Безъякорная	Представляет объект набором точек с последующим преобразованием в OBB; использует DCN, keypoint detection, spatially constrained loss и Adaptive Points Assessment and Assignment	36,83	194,35	26,8

затухания (weight de-cay) составлял $1 \cdot 10^{-4}$. Для стабилизации обучения применялось ограничение градиента по L2-норме с максимальным значением 35.

3.2. Алгоритмы семейства YOLO

В качестве альтернативного подхода рассматривается линейка моделей семейства YOLO (v5, v8, v11, v26). Выделение данных алгоритмов в отдельный подраздел обусловлено целью исследования: проследить, как внутренние архитектурные изменения семейства влияют на способность модели обнаруживать объекты на РЛИ. Большинство современных моделей детекции, включая YOLO, включает три основных компонента [Jocher et al., 2026].

1. Модуль извлечения признаков (Backbone), отвечающий за извлечение иерархических признаков из входного изображения. В процессе обработки пространственное разрешение карт признаков постепенно уменьшается, увеличивая при этом глубину каналов, что позволяет модели переходить от локальных низкоуровневых признаков к более сложным структурным представлениям объекта.

2. Модуль объединения разноуровневых признаков (Neck), предназначенный для объединения карт признаков разных масштабов. Для этого используются схемы пирамидального объединения признаков, включая FPN и PAN, что повышает устойчивость детекции объектов различных размеров.

3. Финальный предсказательный блок (Head), который на основе собранных признаков осуществляет классификацию объекта и регрессию координат его границ.

Далее приведены особенности различных поколений нейронных сетей семейства YOLO [Hidayatullah, Tubagus, 2026]. Используемые значения задержки инференса на CPU заимствованы из официального бенчмарка Ultralytics. В указанном бенчмарке оценка выполнялась для ONNX-представления моделей при размере входного изображения 640×640 пикселей на процессоре Intel Xeon CPU 2,00 ГГц. В связи с этим приведенные значения задержки используются только для ориентировочного сопоставления различных поколений YOLO между собой и не рассматриваются как абсолютная характеристика быстродействия моделей в условиях настоящего исследования.

YOLOv5 (2020) стала одной из базовых реализаций экосистемы Ultralytics, предложив гибкую модульную реализацию на базе PyTorch с масштабированием конфигураций (от n до x). Архитектурно модель использует модуль извлечения признаков с активацией SiLU и модуль объединения разноуровневых признаков PANet. Однако производительность данной версии ограничивается необходимостью постобработки через алгоритм Non-Maximum Suppression (NMS), а также общей вычислительной сложностью используемой архитектуры, что создает дополнительные накладные расходы при инференсе. Согласно используемому бенчмарку, базовая версия YOLOv5(n) демонстрирует точность mAP $\approx$ 34,3 % при задержке около 73,6 мс на CPU.

В YOLOv8 (2023) авторы переходят к архитектуре безъякорных методов, внедрив разделенный финальный предсказательный блок (decoupled head) для более точного предсказания классов и координат. Обновленный модуль извлечения признаков на основе блоков C2f и модуля SPPF позволил оптимизировать количество параметров при сохранении информативности извлекаемых признаков. С точки зрения метрик YOLOv8n достигает 37,3 % mAP, время обработки на CPU осталось сопоставимым с предшественником ($\sim$ 80,4 мс).

YOLOv11 (2024) ориентирована на повышение эффективности и быстродействии модели при работе в условиях ограниченных ресурсов. Внедрение компактных блоков C3k2 и модуля внимания C2PSA направлено на улучшение агрегации признаков. Это указывает на снижение времени инференса по сравнению с YOLOv5: YOLOv11(n) достигает точности 39,5 % mAP при задержке 56,1 мс на CPU.

YOLOv26 (2026) представляет собой итерацию, ориентированную на прямое внедрение в конечные устройства. К числу изменений относятся отказ от NMS, а также исключение Distribution Focal Loses (DFL), что снижает нагрузку на процессор. Благодаря интеграции прогрессивной балансировки потерь (ProgLoss) и механизмов Small Target Label Assignment (STAL) модель YOLO26v(n) показывает 40,3 % mAP при задержке 38,9 мс на CPU.

Кроме того, YOLO характеризуется масштабируемостью архитектуры, которая реализуется через изменение коэффициентов глубины (числа последовательных блоков в слоях) и ширины (количества каналов в сверточных слоях) нейронной сети. В рамках одной версии модели представлены конфигурациями разной вычислительной сложности (Nano — n, Small — s, Medium — m, Large — l, Extra Large — x), что определяет объем извлекаемых признаков и общую емкость гипотез модели. Увеличение размера сети напрямую коррелирует с ростом числа обучаемых параметров и вычислительной сложности в операциях с плавающей точкой (FLOPs), что позволяет более крупным вариантам формировать более детализированные карты признаков при локализации объектов, в то время как компактные модели ориентированы на снижение вычислительных затрат и использование в условиях ограниченных ресурсов.

4. Формирование базы данных

Формирование репрезентативного набора данных для обучения нейросетевых моделей на РЛИ является сложной и нетривиальной задачей, обусловленной спецификой радиолокационной съемки. Существует три основных подхода к формированию таких баз данных: получение данных в ходе натурных экспериментов, математическое моделирование радиолокационных отражений от объектов и подстилающей поверхности, а также использование открытых наборов данных. Первый подход требует значительных временных и финансовых затрат, а второй — высокой точности моделей отражения. В данной работе используется третий вариант — использование баз данных из открытых источников, что позволяет обеспечить сопоставимость результатов с мировыми исследованиями. Проведение исследований было реализовано на наборе данных binMSTAR, модификацией набора mixMSTAR [Liu et al., 2023]. Данный набор данных имеет синтетическую природу формирования ввиду нехватки количества изображений, сформированных в реальных условиях. База данных mixMSTAR содержит информацию о 20 классах распознаваемой техники. Исходные изображения mixMSTAR имеют разных размер ($\sim 1470 \times 1784$, $17\,720 \times 1784$, 8858×1784, $10\,346 \times 1784$, $13\,286 \times 1784$). Таким образом, были сформированы 9 баз данных с разным размером на основе исходного набора данных (640×640, 704×704, 768×768, 832×832, 896×896, 960×960, 1024×1024, 1216×1216, 1408×1408). При формировании изображений использовалось окно размером $S \times S$ с коэффициентом перекрытия 0,5, который обеспечивает перекрытие соседних фрагментов на 50 %. В аннотации были изменены метки 20 классов на метку класса «объект». Каждая из сформированных баз данных была разделена на обучающую и валидационную в соотношении 70 % на 30 %. Тестовая выборка для всех баз данных была получена на основе одних и тех же исходных тестовых изображений из mixMSTAR. В таблице 2 представлена информация о сформированных базах данных.

Несмотря на сокращение выборки при высоком разрешении, данные фрагменты характеризуются высокой плотностью малоразмерных объектов. В качестве иллюстрации на рис. 2 представлен пример фрагмента, полученного путем кадрирования исходного радиолокационного изображения (РЛИ), приведенного на рис. 1.

К сформированным изображениям подготовлены 3 вида аннотаций для обучения в форматах ориентированных рамок (OBB) и неориентированных рамок (HBB) для моделей YOLO, для специализированных алгоритмов аннотация представлена в формате DOTA. В формате DOTA

Таблица 2. Детальная информация о сформированных выборках

Название базы данных	Размер РЛИ, пикселей	Количество изображений, test/train/val	Формат аннотации для ограничивающих рамок
binMSTAR-640	640 × 640	580/509/191	YOLO, DOTA
binMSTAR-704	704 × 704	527/519/204	YOLO
binMSTAR-768	768 × 768	371/257/108	YOLO
binMSTAR-832	832 × 832	339/265/103	YOLO
binMSTAR-896	896 × 896	212/177/72	YOLO
binMSTAR-960	960 × 960	196/177/75	YOLO
binMSTAR-1024	1024 × 1024	184/89/35	YOLO, DOTA
binMSTAR-1216	1216 × 1216	77/46/20	YOLO
binMSTAR-1408	1408 × 1408	65/46/20	YOLO

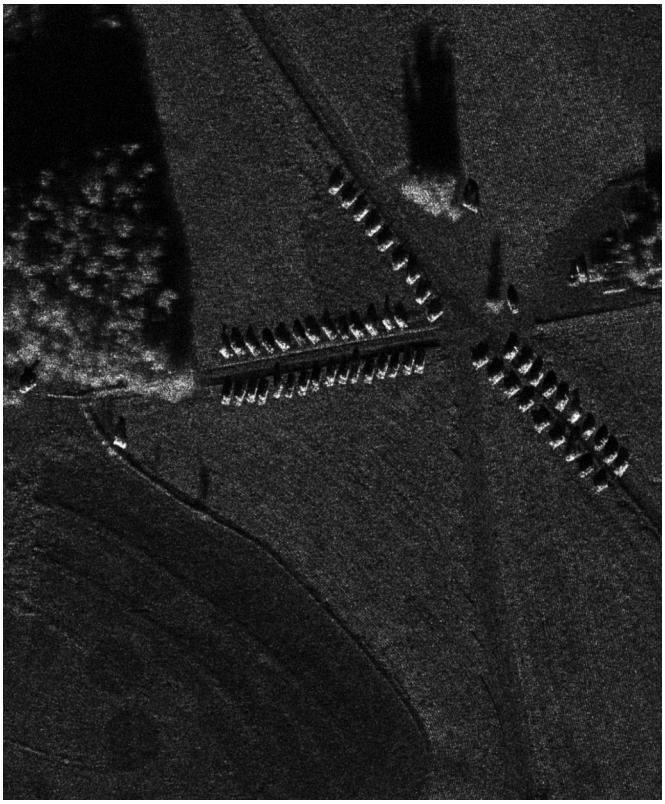


Рис. 1. Исходное изображение из обучающей выборки mixMSTAR

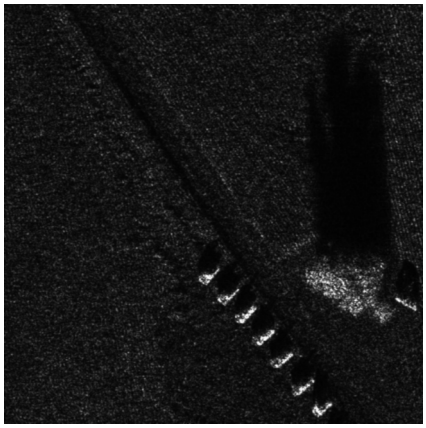


Рис. 2. Сформированное изображение из обучающей выборки binMSTAR-640

(Dataset for Object detection in Aerial Images) каждый объект описан строкой, содержащей 10 значений: 8 координат вершин ограничивающей рамки, строковое обозначение класса и бинарный флаг сложности (difficult).

5. Исследование моделей нейронных сетей

Экспериментальное исследование направлено на верификацию гипотез о влиянии размерности входного кадра S на стабильность детекции и выявление условий, при которых использование OBB становится более эффективным по сравнению с HBB. В данном разделе приводятся

количественные показатели работы нейросетевых архитектур, полученные в ходе тестирования на базе сформированных наборов данных binMSTAR. Для комплексной оценки предложенных подходов и верификации теоретических положений о геометрической неопределенности нами было проведено две серии вычислительных экспериментов.

Первая серия экспериментов посвящена сравнительному анализу эффективности архитектурных решений при фиксированном стандартном разрешении кадра $S = 640$. Данный этап предназначен для проведения оценки внутренней устойчивости различных семейств нейронных сетей в равных условиях. В этом блоке сопоставляются весовые модификации моделей YOLO (от n до l), подходы HBB и OBB в рамках одного семейства алгоритмов, а также специализированные решения, представленные в работе [Liu et al., 2023].

Вторая серия экспериментов представляет собой расширенное исследование масштабируемости, в ходе которого изучается чувствительность различных поколений моделей семейства YOLO в модификациях HBB и OBB к изменению входной размерности кадра S в диапазоне от 640 до 1408. Этот этап позволяет определить динамику изменения точности обнаружения при изменении масштаба рецептивного поля, что критично для выделения малоразмерных объектов на сложном фоне подстилающей поверхности. Проведение комплексного анализа на различных разрешениях позволяет установить границы применимости современных детекторов и выявить оптимальные условия для достижения максимальной точности локализации.

Специализированные алгоритмы обучались с использованием гиперпараметров, соответствующих работе [Liu et al., 2023]. В первой серии экспериментов обучение и тестирование выполнялись при размере входного кадра РЛИ 640×640 пикселей, а во второй серии дополнительно рассматривался размер 1024×1024 пикселя.

Для моделей семейства YOLO в первой серии экспериментов использовался размер входного кадра 640×640 пикселей. Во второй серии экспериментов для оценки влияния пространственного масштаба размер входного кадра варьировался от 640×640 до 1408×1408 пикселей с шагом, кратным 32. При увеличении входного разрешения размер пакета данных пропорционально снижался от 16 до 2 изображений для предотвращения переполнения видеопамяти на предельных разрешениях. Для обеспечения более устойчивой сходимости алгоритмов семейства YOLO при малых размерах пакета в схеме обучения использовался оптимизатор AdamW вместо стохастического градиентного спуска SGD. Выбор AdamW обусловлен тем, что данный оптимизатор сочетает адаптивную коррекцию шага обновления весов с отделенным весовым затуханием, что позволяет стабилизировать обучение и уменьшить риск избыточной подстройки модели под обучающую выборку. Это особенно важно при обработке РЛИ, где неоднородные фоновые области могут иметь признаки, близкие к откликам целевых объектов. В конфигурации эксперимента были заданы начальная скорость обучения 0,003, коэффициент конечной скорости обучения 0,01 и коэффициент регуляризации 0,0005. Обучение всех моделей YOLO проводилось в течение 100 эпох. Мониторинг функции потерь на валидационной выборке показал выход на плато, что указывает на достижение сходимости и снижает вероятность недообучения моделей.

Результаты первой серии экспериментов представлены в таблице 3.

Среди исследованных архитектур наиболее стабильные результаты продемонстрировали модели семейства YOLO. Специализированные модели ориентированного детектирования, обученные и протестированные при размере входного фрагмента 640×640 пикселей, характеризуются высокими значениями Recall, находящимися преимущественно в диапазоне 0,90–0,95, однако для ряда моделей наблюдается заметное снижение Precision. Наиболее низкие значения Precision получены для FCOS, Gliding Vertex, Oriented RepPoints, S2ANet и R3Det. Одним из факторов, влияющих на снижение Precision, может быть низкий порог уверенности предсказаний, который составлял 0,05. При таком значении сохраняется большее число потенциальных

Таблица 3. Метрики бинарной детекции объектов на радиолокационных изображениях

Название модели	Precision	Recall	mAP_{50}	$mAP_{50–95}$
YOLOv5(n) HBB	0,977	0,947	0,979	0,761
YOLOv5(s) HBB	0,967	0,957	0,983	0,752
YOLOv5(m) HBB	0,984	0,965	0,987	0,788
YOLOv5(l) HBB	0,966	0,963	0,985	0,768
YOLOv8(n) HBB	0,984	0,95	0,984	0,762
YOLOv8(s) HBB	0,973	0,962	0,985	0,772
YOLOv8(m) HBB	0,977	0,945	0,977	0,758
YOLOv8(l) HBB	0,929	0,951	0,931	0,716
YOLOv11(n) HBB	0,979	0,96	0,985	0,769
YOLOv11(s) HBB	0,974	0,95	0,981	0,752
YOLOv11(m) HBB	0,981	0,949	0,983	0,777
YOLOv11(l) HBB	0,984	0,96	0,986	0,789
YOLOv26(n) HBB	0,964	0,951	0,983	0,775
YOLOv26(s) HBB	0,98	0,955	0,983	0,781
YOLOv26(m) HBB	0,976	0,947	0,985	0,773
YOLOv26(l) HBB	0,976	0,954	0,984	0,781
YOLOv8(n) OBB	0,982	0,952	0,98	0,847
YOLOv8(s) OBB	0,985	0,946	0,98	0,839
YOLOv8(m) OBB	0,98	0,943	0,976	0,85
YOLOv8(l) OBB	0,978	0,955	0,98	0,851
YOLOv11(n) OBB	0,981	0,943	0,975	0,819
YOLOv11(s) OBB	0,98	0,955	0,98	0,843
YOLOv11(m) OBB	0,968	0,909	0,961	0,778
YOLOv11(l) OBB	0,976	0,937	0,975	0,83
YOLOv26(n) OBB	0,914	0,89	0,952	0,798
YOLOv26(s) OBB	0,976	0,924	0,971	0,823
YOLOv26(m) OBB	0,983	0,943	0,98	0,855
YOLOv26(l) OBB	0,981	0,935	0,977	0,848
Rotated RetinaNet (OBB)	0,545	0,952	0,907	0,611
Oriented Rep Points (OBB)	0,457	0,928	0,906	0,610
Oriented R-CNN (OBB)	0,754	0,930	0,906	0,607
RoI transformer (OBB)	0,916	0,920	0,906	0,605
R3Det (OBB)	0,523	0,930	0,905	0,593
S2ANet (OBB)	0,516	0,922	0,886	0,533
Gliding Vertex (OBB)	0,272	0,928	0,903	0,427
ReDet (OBB)	0,672	0,909	0,895	0,569
FCOS (OBB)	0,233	0,940	0,906	0,597

предсказаний, что снижает вероятность пропуска целевых объектов и способствует повышению Recall, однако одновременно увеличивает вероятность ложноположительных срабатываний. Более вероятной причиной снижения Precision при размере входного фрагмента 640×640 пикселей является сокращение пространственного контекста сцены. При меньшем размере входного фрагмента модель анализирует более локальную область РЛИ, поэтому отдельные яркие отражатели, неоднородности фона и помеховые участки могут восприниматься как признаки целевого объекта. При увеличении размера входного фрагмента модель получает более полное представление об окружении объекта, включая характер фоновой поверхности, взаимное расположение радиолокационных откликов и прилегающие участки радиолокационной тени. Это повышает устойчивость различения объекта и фона и может приводить к росту Precision. На рис. 3 и рис. 4 представлены гистограммы распределения метрик Recall и $mAP_{50–95}$ соответственно.

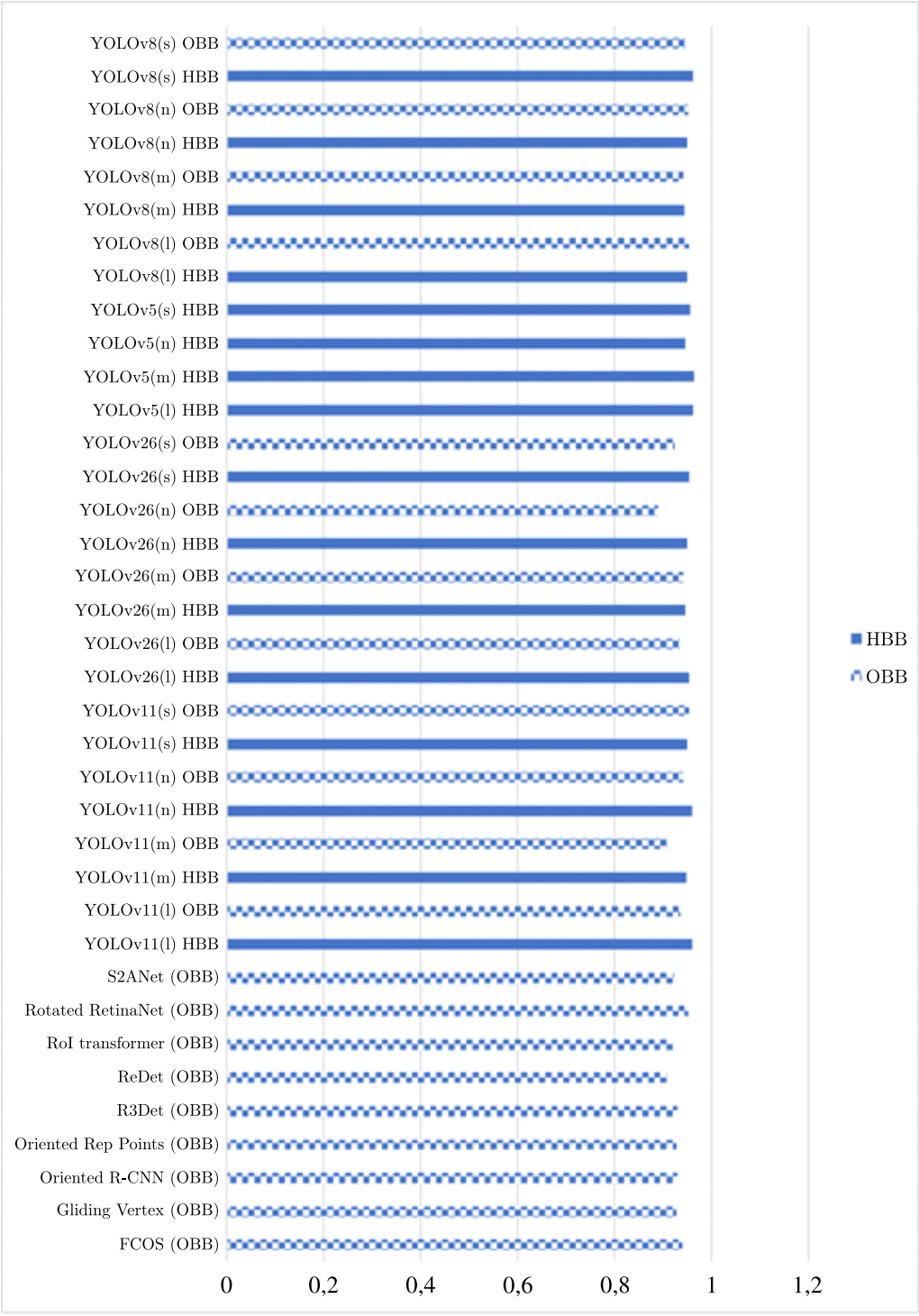


Рис. 3. Гистограмма распределения метрики Recall для разных моделей

На рис. 5 и рис. 6 представлены средние значения метрик для различных размеров и различных поколений моделей семейства YOLO соответственно.

Для оценки статистической значимости превосходства методов с ориентированными рамками (OBB) над методами с горизонтальными рамками (HBB) был применен непараметрический U-критерий Манна – Уитни [Mann, Whitney, 1947]. Помимо этого, выполнен статистический ана-

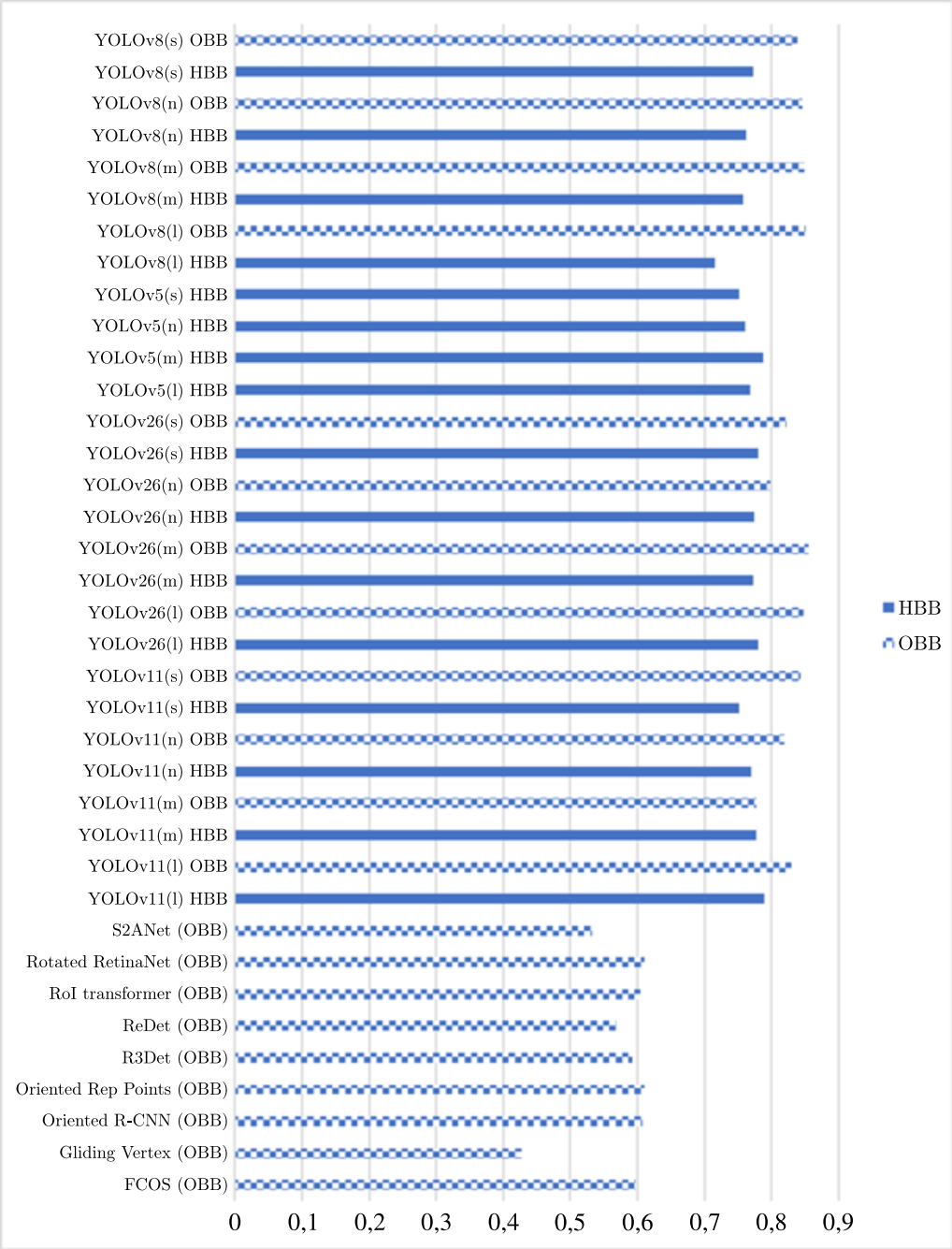


Рис. 4. Гистограмма распределения метрики $mAP_{50–95}$ для разных моделей

лиз влияния различных размерных характеристик и различных поколений нейросетевых моделей семейства YOLO на итоговые метрики качества обнаружения. Значение критерия при уровне значимости p-value < 0,05 подтверждает, что наблюдаемое преимущество предложенного подхода является статистически достоверным и не обусловлено случайным распределением результатов на тестовой выборке.

Сравнение проводилось для моделей семейства YOLO по выборкам значений Precision, Recall, mAP_{50} и $mAP_{50–95}$, полученным на тестовом наборе данных. В таблице 4 приведены результаты расчетов U-критерия. В этой таблице представлены средние значения сравниваемых групп по метрикам. Знак стрелки означает преимущество одной из рассматриваемых групп, на-

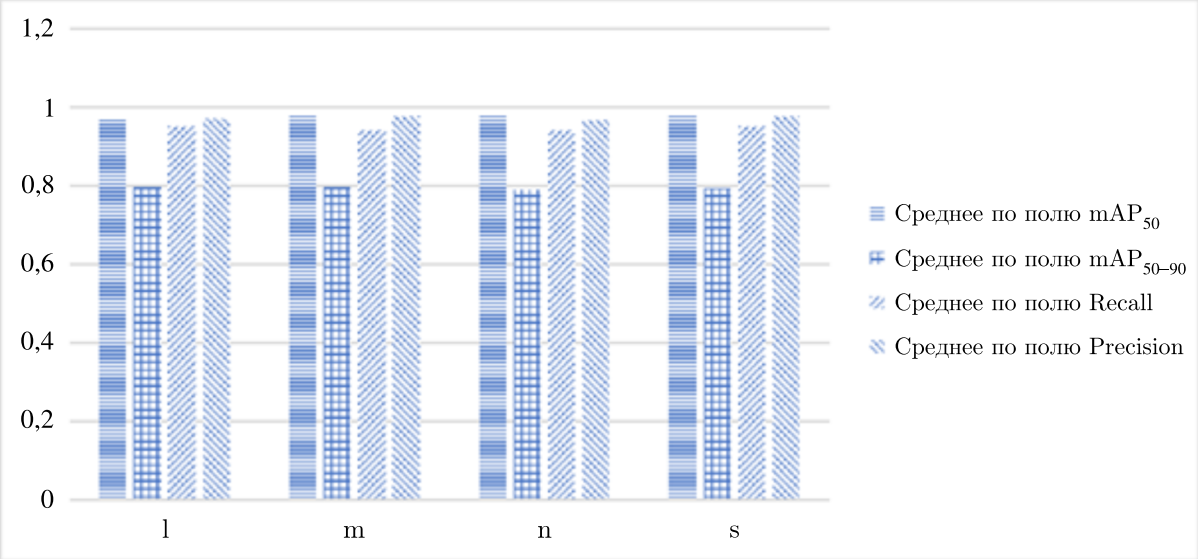


Рис. 5. Средние показатели метрик для моделей YOLO разного размера

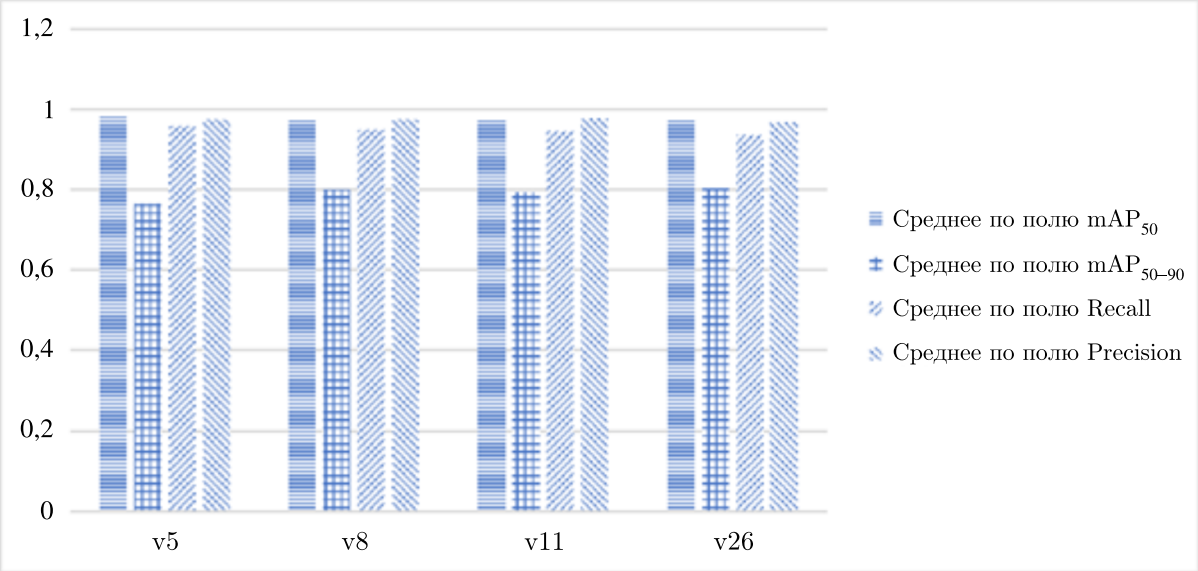


Рис. 6. Средние показатели метрик для моделей YOLO разного поколения

звание которой стоит перед стрелкой. Число означает среднюю разницу метрики. Так, например, при сравнении HBB и OBB в столбце $mAP_{50–95}$ видно, что результат у алгоритмов с OBB лучше в среднем на 6,5 %, знак звездочки означает, что различие статистики значимо (p-value < 0,05). Отсутствие звездочки означает, что различие не значимо. Буква v (version) означает поколение.

Таким образом, сравнительный анализ показал, что формат HBB имеет статистически значимый прирост по показателям первичного обнаружения, но он мал по абсолютной величине (Recall выше на 0,6 %, mAP_{50} — на 1,8 %). Однако при переходе к строгим метрикам локализации наблюдается статистически значимое превосходство формата OBB по метрике $mAP_{50–95}$, что свидетельствует о более высокой точности описания границ объектов при использовании ориентированных рамок. Также на основе статистического U-критерия было выявлено, что ни размер, ни поколение модели не оказывают существенного влияния на итоговые метрики, что связано

Таблица 4. Результаты расчетов U-критерия для алгоритмов YOLO

Сравнение	Precision	Recall	mAP_{50}	$mAP_{50–95}$
v11 пр. v26	v11↑0,009	v11↑0,008	v11↑0,001	v26↑0,010
v5 пр. v11	v11↑0,004	v5↑0,013	v5↑0,005	v11↑0,027
v5 пр. v26	v5↑0,005	v5↑0,021∗	v5↑0,007	v26↑0,037
v5 пр. v8	v8↑0,000	v5↑0,007	v5↑0,009	v8↑0,032
v8 пр. v11	v11↑0,004	v8↑0,005	v11↑0,004	v8↑0,005
v8 пр. v26	v8↑0,005	v8↑0,013	v26↑0,003	v26↑0,005
HBB пр. OBB	OBB↑0,000	HBB↑0,018∗	HBB↑0,006∗	OBB↑0,065∗
Размер l пр. Размер m	Размер m↑0,008	Размер l↑0,008	Размер m↑0,004	Размер l↑0,001
Размер l пр. Размер n	Размер l↑0,001	Размер l↑0,009	Размер n↑0,003	Размер l↑0,007
Размер l пр. Размер s	Размер s↑0,006	Размер l↑0,001	Размер s↑0,006	Размер l↑0,003
Размер m пр. Размер n	Размер m↑0,010	Размер m↑0,001	Размер m↑0,002	Размер m↑0,007
Размер m пр. Размер s	Размер m↑0,002	Размер s↑0,007	Размер s↑0,002	Размер m↑0,002
Размер n пр. Размер s	Размер s↑0,008	Размер s↑0,008	Размер s↑0,004	Размер s↑0,004

со специфичностью радиолокационных изображений и ограниченным набором информативных признаков.

Для второй серии экспериментов было обучено 63 различных моделей YOLO. На рис. 7, 8, 9 представлены зависимости усредненных показателей метрик $mAP_{50–95}$, Recall, Precision соответственно (в зависимости от входного разрешения).

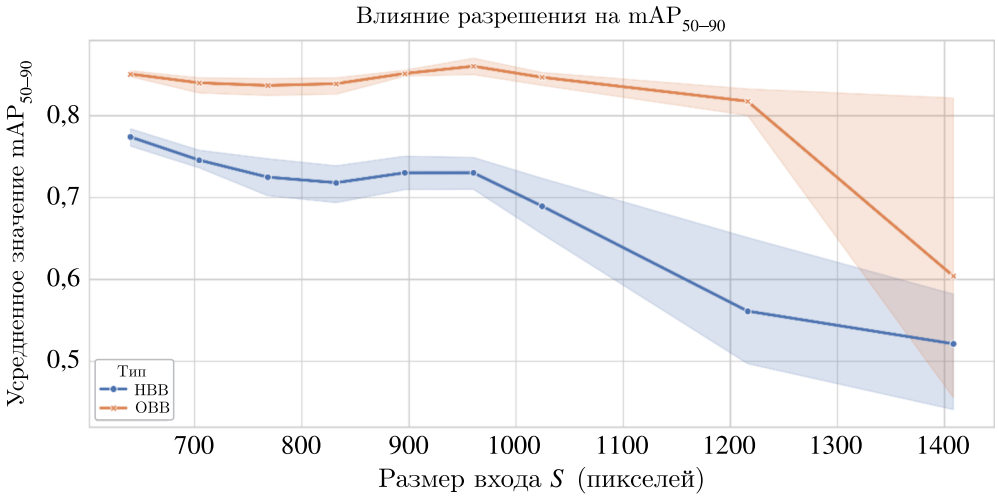


Рис. 7. Зависимость среднего показателя $mAP_{50–95}$ от входного размера S

На рис. 7, 8, 9 области вдоль линий обозначают 95%-й доверительный интервал. Всего моделей для OBB было 27 (YOLOv8(m), YOLOv11(m), YOLOv26(m) для 9 разных входов), а для HBB — 36. Во втором случае больше моделей из-за того, что использовались модели YOLOv5(m), они не имеют модификации с ориентированными ограничивающими рамками.

Экспериментально установлено, что эффективность детектирования нелинейно зависит от разрешения входного изображения. Максимальные значения целевых метрик (mAP_{50} и $mAP_{50–95}$) для ориентированного детектирования достигаются в диапазоне 896–1024 пикселей. Выход за пределы данного диапазона в сторону увеличения разрешения приводит к деградации качества распознавания, что подтверждает наличие оптимального масштаба входных данных для исследуемого набора объектов. Модель YOLOv11(m) при разрешении 960 пикселей обеспечила пиковое значение $mAP_{50–95} = 0{,}87$.

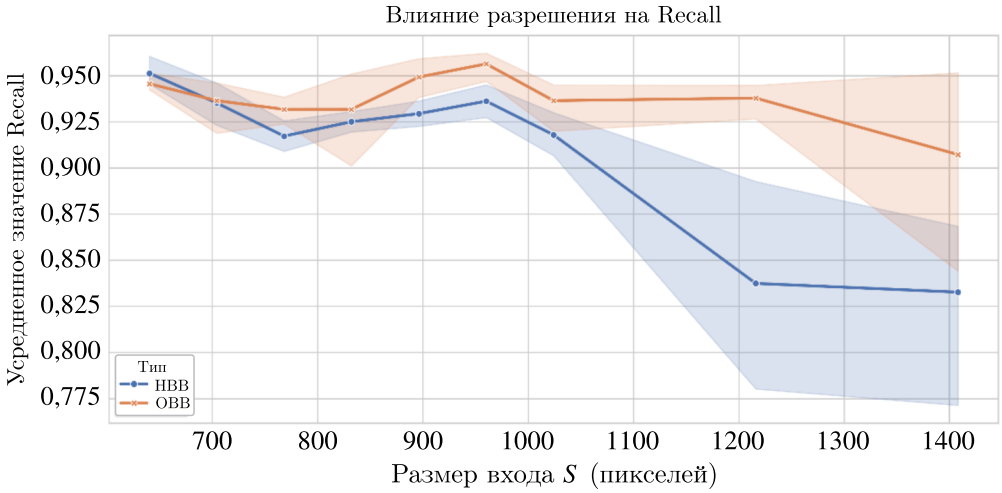


Рис. 8. Зависимость среднего показателя Recall от входного размера S

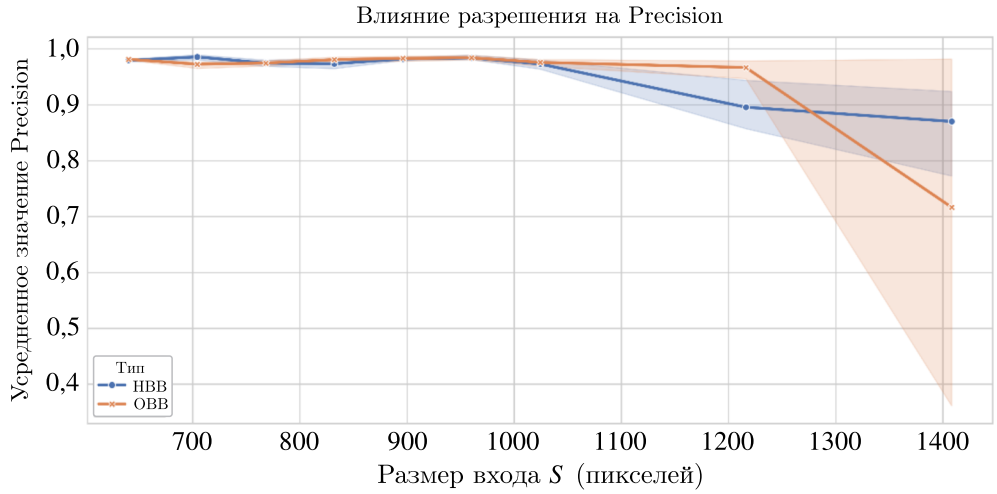


Рис. 9. Зависимость среднего показателя Precision от входного размера S

В таблице 5 представлены метрики некоторых специализированных моделей нейронных сетей, обученных на разрешении $S = 1024$, в сравнении с моделями, обученными на разрешении $S = 640$.

Таблица 5. Метрики качества для специализированных моделей

Модель / размер входного слоя	mAP_{50}	$mAP_{50–95}$	Precision	Recall
Gliding Vertex / 1024	0,906	0,39	0,934	0,980
Gliding Vertex / 640	0,903	0,427	0,272	0,928
Oriented R-CNN / 1024	0,909	0,578	0,924	0,995
Oriented R-CNN / 640	0,906	0,607	0,754	0,930
RoI Transformer / 1024	0,909	0,654	0,934	0,995
RoI Transformer / 640	0,906	0,605	0,916	0,920
Rotated RetinaNet / 1024	0,909	0,649	0,937	0,997
Rotated RetinaNet / 640	0,907	0,611	0,545	0,952

Специализированные модели, обученные при размере входного фрагмента 1024×1024 пикселя, в большинстве случаев показывают более высокие значения Precision и Recall по сравнению с моделями, обученными при размере 640 × 640 пикселей. Наиболее выраженный рост Precision наблюдается для Gliding Vertex, Oriented R-CNN и Rotated RetinaNet. Эти результаты согласуются с предположением о том, что снижение Precision при размере 640 × 640 пикселей связано с недостаточным пространственным контекстом сцены. При размере 1024 × 1024 пикселя модель получает более полное представление об окружении объекта, включая характер фоновой поверхности и взаимное расположение радиолокационных откликов. Это повышает устойчивость различения объекта и фона и может снижать число ложноположительных срабатываний.

6. Заключение

Проведенное исследование позволило верифицировать гипотезы о влиянии методов аннотирования и параметров входных данных на эффективность детектирования малоразмерных объектов на радиолокационных изображениях. На основании полученных результатов можно сформулировать следующие выводы.

1. Методы аннотирования. Установлено, что формат ориентированных ограничивающих рамок обладает статистически значимым преимуществом над горизонтальными по метрике $mAP_{50–95}$. Это обусловлено минимизацией фонового шума в обучающей выборке: исключение избыточных нецелевых пикселей предотвращает эффект «размытия» градиентов, что критично для точного предсказания координат объектов со сложной ориентацией. При этом для первичной локализации (метрики Recall и mAP_{50}) различия между форматами остаются несущественными (в пределах 0,6–1,8 %). Также на основе статистического U-критерия было выявлено, что ни размер, ни поколение модели не оказывают существенного влияния на итоговые метрики, что связано со специфичностью радиолокационных изображений и ограниченным набором информативных признаков.

2. Параметры входных данных. Экспериментально подтверждено наличие оптимального пиксельного разрешения РЛИ. Для задачи бинарной детекции малоразмерных объектов на РЛИ максимизация метрик достигается в диапазоне 896–1024 пикселей. Дальнейшее увеличение разрешения ведет к деградации качества из-за перенасыщения признакового описания фоновой составляющей, тогда как снижение разрешения ниже оптимального порога снижает значение метрик из-за недостаточности контекста.

3. Сравнительная эффективность моделей. Наилучшие метрики продемонстрировали модели семейства YOLO, показав высокую устойчивость к архитектурным изменениям. В частности, модель YOLOv11(m) при разрешении 960 пикселей обеспечила пиковое значение $mAP_{50–95}$ = 0,87. Специализированные алгоритмы ориентированного детектирования (такие как R3Det, S2ANet) показывают сопоставимые или более высокие значения Recall при разрешении 1024 пикселей, по остальным метриками эти алгоритмы проигрывают.

4. Практические рекомендации. Для решения задач бинарной детекции на РЛИ целесообразно использовать модели YOLO с входным разрешением в диапазоне 896–1024 пикселей. Специализированные алгоритмы с высокой полнотой обнаружения могут быть эффективно интегрированы в каскадные системы в качестве первого этапа обработки.

Перспективы дальнейших исследований связаны с разработкой каскадного бинарного детектора, дополненного специализированным классификатором «фон/цель». Такая архитектура позволит объединить высокую чувствительность первичного обнаружения с фильтрацией ложных срабатываний, что может обеспечить повышение метрики Precision.

Список литературы (References)

Красноперова А. С., Твердохлебов А. С., Карташов А. А., Вебер В. И., Куприц В. Ю. Исследование эффективности применения моделей нейронных сетей YOLO для распознавания объектов на радиолокационных изображениях // Российский технологический журнал. — 2025. — Т. 13, № 4. — С. 25–36. — DOI: 10.32362/2500-316X-2025-13-4-25-36

Krasnoperova A. S., Tverdokhlebov A. S., Kartashov A. A., Weber V. I., Kuprits V. Yu. Issledovanie effektivnosti primeneniya modeley neyronnykh setey YOLO dlya raspoznavaniya obektov na radiolokatsionnykh izobrazheniyakh [Study of efficiency of YOLO neural network models for radar image object recognition] // Russian Technological Journal. — 2025. — Vol. 13, No. 4. — P. 25–36 (in Russian). — DOI: 10.32362/2500-316X-2025-13-4-25-36

Купряшкин И. Ф. Детектирование объектов военной техники на зашумленных радиолокационных изображениях с использованием детектора YOLOv4-tiny // Журнал радиоэлектроники. — 2025. — № 2. — DOI: 10.30898/1684-1719.2025.2.7

Kuprjashkin I. F. Detektirovanie obektov voennoy tekhniki na zashumlennykh radiolokatsionnykh izobrazheniyakh s ispolzovaniem detektora YOLOv4-tiny [Detection of military equipment objects in noisy radar images using the YOLOv4-tiny detector] // Zhurnal radioelektroniki. — 2025. — No. 2 (in Russian). — DOI: 10.30898/1684-1719.2025.2.7

Купряшкин И. Ф. Детектирование объектов военной техники с использованием сверточной нейронной сети на радиолокационных изображениях, сформированных в условиях шумовых помех // Журнал радиоэлектроники. — 2022. — № 6. — DOI: 10.30898/1684-1719.2022.6.8

Kuprjashkin I. F. Detektirovanie obektov voennoy tekhniki s ispolzovaniem svertochnoy neyronnoy seti na radiolokatsionnykh izobrazheniyakh, sformirovannykh v usloviyakh shumovykh pomekh [Military objects detection using a convolutional neural network on radar images formed in jamming conditions] // Zhurnal radioelektroniki. — 2022. — No. 6 (in Russian). — DOI: 10.30898/1684-1719.2022.6.8

Dai J., Qi H., Xiong Y., Li Y., Zhang G., Hu H., Wei Y. Deformable convolutional networks // Proceedings of the IEEE International Conference on Computer Vision. — 2017. — P. 764–773.

Ding J., Xue N., Long Y., Xia G.-S., Lu Q. Learning ROI transformer for oriented object detection in aerial images // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. — 2019. — P. 2849–2858.

Gao F., Tian Y., Wu Y., Zhang Y. ST-YOLOv8: Small-target ship detection in SAR images targeting specific marine environments // Applied Sciences. — 2025. — Vol. 15, No. 12. — Art. 6666.

Han J., Ding J., Li J., Xia G.-S. Align deep features for oriented object detection // IEEE Transactions on Geoscience and Remote Sensing. — 2021. — Vol. 60. — P. 1–11.

Han J., Kang L., Tian J., Jiang M., Guo N. Vehicle localization method in complex SAR images based on feature reconstruction and aggregation // Sensors. — 2024. — Vol. 24, No. 20. — Art. 6746.

Hidayatullah P., Tubagus R. YOLO26: A comprehensive architecture overview and key improvements // arXiv preprint. — 2026. — arXiv:2602.14582

Jocher G., Qiu J., Chaurasia A. Ultralytics YOLO. — [Electronic resource]. — https://github.com/ultralytics/ultralytics (accessed: 28.04.2026).

Li Y., Li X., Li W., Hou Q., Liu L., Cheng M.-M., Yang J. SARDet-100K: Towards open-source benchmark and toolkit for large-scale SAR object detection // Advances in Neural Information Processing Systems. — 2024. — Vol. 37. — P. 128430–128461.

Lin T.-Y., Goyal P., Girshick R., He K., Dollár P. Focal loss for dense object detection // Proceedings of the IEEE International Conference on Computer Vision. — 2017. — P. 2980–2988.

Liu Z., Luo S., Wang Y. Mix MSTAR: A synthetic benchmark dataset for multi-class rotation vehicle detection in large-scale SAR images // Remote Sensing. — 2023. — Vol. 15, No. 18. — Art. 4558.

Ma J., Shao W., Ye H., Wang L., Wang H., Zheng Y., Xue X. Arbitrary-oriented scene text detection via rotation proposals // IEEE Transactions on Multimedia. — 2018. — Vol. 20. — P. 3111–3122.

Mann H. B., Whitney D. R. On a test of whether one of two random variables is stochastically larger than the other // Annals of Mathematical Statistics. — 1947. — Vol. 18, No. 1. — P. 50–60.

Ren S., He K., Girshick R., Sun J. Faster R-CNN: Towards real-time object detection with region proposal networks // Advances in Neural Information Processing Systems. — 2015. — Vol. 28. — P. 91–99.

Xie X., Cheng G., Wang J., Yao X., Han J. Oriented R-CNN for object detection // Proceedings of the IEEE/CVF International Conference on Computer Vision. — 2021. — P. 3520–3529. — DOI: 10.1109/ICCV48922.2021.00350

Xu Y., Fu M., Wang Q., Wang Y., Chen K., Xia G.-S., Bai X. Gliding vertex on the horizontal bounding box for multi-oriented object detection // IEEE Transactions on Pattern Analysis and Machine Intelligence. — 2021. — Vol. 43, No. 4. — P. 1452–1459. — DOI: 10.1109/TPAMI.2020.2974745

Yin J., Xu G., Wang N., Wang N., Zhang Z. A survey of small sea-surface target detection for maritime search and rescue // IEEE Transactions on Intelligent Transportation Systems. — 2026. — Vol. 27, No. 1. — P. 211–231. — DOI: 10.1109/TITS.2025.3635199

Zhou Y., Yang X., Zhang G., Wang J., Liu Y., Hou L., Jiang X., Liu X., Yan J., Lyu C., Zhang W., Chen K. MMRotate: A rotated object detection benchmark using PyTorch // Proceedings of the 30th ACM International Conference on Multimedia. — 2022. — P. 7331–7334. — DOI: 10.1145/3503161.3548541