ISSN 2658-5014 (Print) СИИТ ISSN 2686-7044 (Online)
http://siit.ugatu.su 2027. Т. 9, № 1(30)
СИСТЕМНАЯ ИНЖЕНЕРИЯ И ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ
Научная статья

УДК 621.396.969 : 004.8 DOI: 10.54708/SIIT-2027-no1-p31 EDN: LRYXUG

Распознавание наземных объектов по разноракурсным данным многопозиционной системы бортовых РСА: разработка и апробация нейросетевого алгоритма на основе имитационного моделирования

С. Н. Кокутин • И.Р. Мергалиев

Университет Иннополис

Автоматическое распознавание наземных объектов по радиолокационным изображениям, получаемым с борта беспилотных воздушных судов, требует учета сложной фоно-целевой обстановки и выраженной ракурсной изменчивости радиолокационных сигнатур целей, при этом для подходов на основе глубокого обучения характерен острый дефицит обучающих наборов данных. В работе представлены результаты реализации и экспериментальной апробации нейросетевого алгоритма распознавания неподвижных наземных объектов на радиолокационных кадрах, формируемых в многопозиционной системе малогабаритных бортовых радиолокаторов с синтезированной апертурой X-диапазона, в качестве носителей которых используется группа беспилотных воздушных судов самолетного типа. Предложенный подход базируется на аппарате имитационного моделирования, включающем генерацию синтетических радиолокационных кадров посредством программного симулятора на основе модифицированного алгоритма растеризации и слияние в единое трехканальное изображение двух разноракурсных кадров, получаемых в проекции горизонтальной дальности. Сформирован расширенный синтетический датасет для семи целевых классов, относящихся к категориям транспортных средств, зданий и деревьев. Датасет создан путем управляемого варьирования состава, взаимного расположения и ориентации наземных объектов, а также типа подстилающей поверхности. Реализован конвейер обработки на базе облегченной одностадийной архитектуры YOLOv12S и механизма нарезки входного объединенного радиолокационного кадра на перекрывающиеся фрагменты фиксированного размера с последующим пространственным объединением локальных предсказаний в единый набор результатов распознавания. Результаты экспериментов продемонстрировали эффективность предложенного решения при распознавании наземных объектов в различных условиях наблюдения и подтвердили его потенциальную применимость в подсистемах технического зрения беспилотных авиационных систем.

Радиолокатор с синтезированной апертурой; многопозиционная радиолокационная система; беспилотное воздушное судно; имитационное моделирование; синтетический датасет; радиолокационный кадр; нейросетевой алгоритм; распознавание наземных объектов.

Введение

Распознавание наземных объектов по радиолокационным изображениям (РЛИ), получаемым с помощью радиолокаторов с синтезированной апертурой (РСА), на протяжении последних десятилетий остается одной из ключевых и наиболее динамично развивающихся задач в области дистанционного зондирования Земли (ДЗЗ) и интеллектуальной обработки данных [Hen13, Bep10]. В отличие от оптических систем, радиолокационные средства сохраняют работоспособность в условиях ограниченной видимости, не зависят от естественного освещения и позволяют получать информацию о сцене в широком диапазоне метеоусловий. Эти преимущества обусловливают широкое применение РСА в задачах мониторинга чрезвычайных

Кокутин С. Н., Мергалиев И. Р. Распознавание наземных объектов по разноракурсным данным многопозиционной системы бортовых РСА: разработка и апробация нейросетевого алгоритма на основе имитационного моделирования // СИИТ. 2027. Т. 9, № 1(30). С. 31-54. DOI: 10.54708/SIIT-2027-no1-p31. EDN: LRYXUG.

Kokutin S. N., Mergaliev I. R. “Ground object recognition from multi-aspect data of a multistatic onboard SAR system: development and experimental evaluation of a neural network algorithm based on simulation modeling” // SIIT. 2027. Vol. 9, no. 1(30), pp. 31-54. DOI: 10.54708/SIIT-2027-no1-p31. EDN: LRYXUG. (In Russian).

ситуаций, всепогодного наблюдения и картографирования, ледовой разведки и многих других. В последние годы наблюдается экспоненциальный рост числа коммерческих спутников с РСА – такие компании, как Umbra, Capella Space и ICEYE, активно наращивают свои орбитальные группировки, что приводит к лавинообразному увеличению объема доступных радиолокационных данных [Mor26].

Наряду со спутниковым мониторингом все более востребованными становятся бортовые радиолокационные станции (РЛС), устанавливаемые на беспилотные воздушные суда (БВС) [Фед22]. Такие системы обеспечивают оперативное получение данных в локальных районах с возможностью гибкого изменения ракурса и высоты наблюдения, что критически важно для задач поиска и спасения, мониторинга стихийных бедствий и загрязнений, охраны периметра. В отличие от спутников, БВС позволяют реализовать многопозиционные схемы наблюдения, при которых один и тот же объект может быть облучен с нескольких направлений в пределах одного пролета, что создает принципиально новые возможности для повышения информативности распознавания [Kan25]. Именно для таких сценариев с участием авторов настоящего исследования разрабатывается прототип многопозиционной системы малогабаритных бортовых РЛС [Ком26], предназначенной для автоматического обнаружения и распознавания наземных объектов в масштабе времени, близком к реальному.

Однако интерпретация РЛИ, получаемых посредством бортовых авиационных систем (БАС), остается крайне сложной задачей ввиду присущих им характерных особенностей: спекл-шума, искажений сигнала, сложных эффектов многократного рассеяния и сильной зависимости наблюдаемого образа объекта от ракурса облучения и свойств подстилающей поверхности. В прикладных сценариях бортового авиационного мониторинга особую актуальность приобретает задача распознавания ключевых наземных объектов – транспортных средств, зданий различных типов и одиночных деревьев, которые могут служить ориентирами или представлять самостоятельный интерес для системы наблюдения. Для таких объектов характерна высокая внутриклассовая вариативность и выраженная межклассовая близость в радиолокационном представлении, что существенно затрудняет применение универсальных подходов, разработанных для оптических RGB и многоспектральных изображений.

Ключевым сдерживающим фактором развития в области автоматического распознавания целей с применением РСА (Synthetic Aperture Radar Automatic Target Recognition, SAR ATR) остается острый дефицит крупномасштабных размеченных датасетов на основе натурных РЛИ. Сбор и аннотирование реальных РЛИ сопряжены со значительными финансовыми и временными затратами, требуют проведения дорогостоящих натурных экспериментов с использованием авиационных носителей. На протяжении многих лет единственным общепризнанным бенчмарком оставался набор данных MSTAR (Moving and Stationary Target Acquisition and Recognition), собранный еще в 1990-х годах [Che16, Bla20]. Ограниченный объем и идеальные условия регистрации данных MSTAR не позволяют обучать нейросетевые модели для реальных условий эксплуатации, где неизбежно возникают существенные сдвиги распределений, вызванные изменчивостью фоно-целевой обстановки и параметров радиолокационной съемки. Лишь в последние годы появились новые крупные датасеты, такие как ATRNet-STAR, содержащий более 190 000 изображений 40 типов транспортных средств в различных условиях наблюдения [Liu25b], однако и они не в полной мере охватывают все разнообразие реальных сценариев.

Особенно остро проблема дефицита данных проявляется в двух аспектах. **Во-первых**, большинство существующих эталонных датасетов содержит либо изображения единичных объектов, изолированных на сцене [Bla20, Liu25b], либо изображения с несколькими объектами одного класса [Lin23]. Однако такие сцены хорошо подходят только для этапа предварительного обучения (Pre-training). В реальных условиях наблюдения с БВС объекты редко встречаются поодиночке – гораздо чаще на одном РЛИ присутствуют близкорасположенные групповые объекты (скопления транспортных средств, элементы застройки или различные естественные объекты), подверженные взаимным эффектам радиозатенения (Shadowing),

радарного наложения (Layover), и многократных переотражений (Multiple-Bounce) [Gro22]. Отсутствие достаточного количества размеченных датасетов со сценами, содержащими множественные объекты различных классов, существенно ограничивает возможность обучения нейросетевых моделей, реализующих устойчивые алгоритмы автоматического распознавания для работы в практических сценариях. Подобные крупные датасеты, такие как SARDet-100K [Li24], являются редким исключением. **Во-вторых**, существующие бенчмарки характеризуются ограниченным разнообразием подстилающей поверхности (фона). Натурные РЛИ обычно собираются в ограниченных географических районах с определенным и зачастую однородным типом ландшафта, тогда как в реальных миссиях БВС объекты могут располагаться на фоне растительности, почвы, сельскохозяйственных угодий, дорожного покрытия, посреди лесных массивов или городской застройки. Радиолокационный отклик объектов, особенно малоразмерных, существенно зависит от свойств окружающего фона, и модели, обученные в идеальных (однородных) фоновых условиях, демонстрируют снижение точности при переносе на сцены с иным типом подстилающей поверхности. Таким образом, проблема дефицита размеченных натурных данных с групповыми объектами на гетерогенном фоне является одним из главных препятствий на пути создания эффективных и устойчивых систем автоматического распознавания для БВС.

В работах последних лет прослеживаются несколько основных направлений решения задачи распознавания объектов на РЛИ, каждое из которых в той или иной степени пытается преодолеть описанные ограничения.

Первое направление связано с адаптацией и модификацией архитектур глубокого обучения для решения задач обнаружения и распознавания объектов на РЛИ. В рамках этого направления активно применяются нейросетевые детекторы семейства YOLO (You Only Look Once), которые благодаря высокой скорости и точности стали одной из наиболее распространенных базовых архитектур [Qiu23]. Однако специфика радиолокационного сигнала – спекл-шум, сложная фоновая обстановка, малые размеры целей – обусловила необходимость дальнейших архитектурных усовершенствований. В работах последних лет предлагаются гибридные решения, сочетающие детекторы семейства YOLO с многоуровневой интеграцией признаков и блоками динамической маршрутизации внимания [Roc25]. Подобные модификации позволяют повысить устойчивость детекции к сложным фонам и улучшить выделение признаков малоразмерных объектов. Однако такие решения нередко сопровождаются ростом вычислительной сложности и, что более важно, не устраняют фундаментальной зависимости от качества и репрезентативности обучающей выборки, а также сохраняют чувствительность к доменному сдвигу (Domain Shift), вызванному изменениями условий съемки (изменения ракурса съемки, угла падения луча или пространственного разрешения получаемых данных).

Второе направление, получившее особенно интенсивное развитие в последние годы, связано с использованием синтетических данных для обучения моделей SAR ATR. Поскольку сбор и разметка реальных РЛИ требуют значительных ресурсов, синтетическая генерация рассматривается как перспективный путь восполнения дефицита данных. В ряде работ предлагаются различные подходы к синтезу: физически обоснованные симуляторы РЛИ на основе электродинамической модели [Li22], симуляторы на основе трассировки лучей [Aue16], растеризации [Bal06], а также симуляторы с применением генеративно-состязательных сетей (GAN) [Sun24] и диффузионных вероятностных моделей [Zha26]. Синтетические данные позволяют формировать сцены с групповыми объектами и разнообразными типами подстилающей поверхности, что принципиально невозможно обеспечить в натурных экспериментах в необходимом масштабе. В то же время синтетические данные неизбежно содержат упрощающие предположения о физике рассеяния и структуре фона. В результате модели, обученные исключительно на синтетических данных, демонстрируют снижение обобщающей способности при переносе на реальные измерения, что известно как «разрыв между доменами» (Domain Gap).

Третье направление связано с использованием многопозиционных схем наблюдения, построение и оптимизация которых на базе беспилотных носителей рассматриваются в работах [Son25, Xu20]. В таких системах один и тот же объект рассматривается под разными углами обзора. Как известно, РЛИ формируется как измерение коэффициента обратного рассеяния наземных объектов (удельная эффективная площадь рассеяния, УЭПР). При изменении ракурса облучения энергия, отраженная от различных частей объекта, варьируется, что приводит к существенным изменениям наблюдаемого образа [Zha23]. В ряде случаев с некоторого ракурса признаки цели могут оказаться невыраженными, делая объект слабо идентифицируемым. Наблюдение объектов под разными углами дает возможность получать асимметричные характеристики их отражения, тем самым повышая полноту информации и помехоустойчивость, обеспечивая обнаружение небольших наземных целей [Ye25]. Многопозиционное наблюдение позволяет более полно зафиксировать характеристики цели и устранять наличие слепых зон за счет комплементарности ракурсов, а также получать РЛИ более высокого разрешения с меньшим уровнем шума за счет слияния данных с разных ракурсов, тем самым повысить надежность распознавания.

Несмотря на впечатляющий прогресс, современное состояние SAR ATR характеризуется рядом нерешенных проблем. **Во-первых**, сохраняется острая нехватка крупномасштабных размеченных датасетов натурных изображений, особенно для задач распознавания групповых объектов на разнородных фонах. **Во-вторых**, проблема переносимости моделей с синтетических данных на реальные измерения (Sim-to-Real) остается одним из главных препятствий на пути практического внедрения. **В-третьих**, многопозиционные подходы, обладающие значительным потенциалом, пока не получили широкого распространения в силу методологических и вычислительных сложностей их реализации. **В-четвертых**, как отмечается в недавних обзорах [Li23, Lan25, Кок25], глубокое обучение позволило существенно смягчить ряд долговременных проблем SAR ATR, однако одновременно породило и новые вызовы, требующие дальнейшего осмысления: необходимость больших вычислительных мощностей, переобучение моделей глубокого обучения из-за недостатка обучающих данных, ложные срабатывания при наличии сложных сцен и помех.

Настоящая работа направлена на преодоление перечисленных ограничений. Целью исследования является разработка и экспериментальная апробация нейросетевого алгоритма 2D-распознавания неподвижных наземных объектов на РЛК, получаемых с БВС в многопозиционной системе, на основе имитационного моделирования и современной нейросетевой архитектуры.

Научная новизна предлагаемого подхода заключается, **во-первых**, в использовании многопозиционной радиолокационной информации для формирования единого трехканального входа, что обеспечивает сохранение ракурсных радиояркостных контрастов и компенсацию анизотропии отражения; **во-вторых**, в комплексном применении средств имитационного физико-математического моделирования радиолокационных сцен для управляемого обучения нейросетевой модели распознавания, что компенсирует дефицит натурных размеченных данных.

Практическая значимость работы состоит в возможности использования разработанного нейросетевого алгоритма и конвейера обработки в составе подсистемы группового технического зрения БАС для автоматического обнаружения и распознавания наземных объектов на разноракурсных РЛК в близком к реальному масштабе времени.

Для достижения поставленной цели в работе решаются следующие основные задачи:

- обоснование и реализация схемы объединения двух разноракурсных РЛК в единое трехканальное изображение с согласованием пространственной ориентации и аннотаций;
- формирование на базе имитационного моделирования трехпозиционной системы бортовых РСА обучающего и тестового наборов радиолокационных данных, включающих сцены с объектами семи классов на фоне четырех типов подстилающей поверхности;

- разработка алгоритма обработки данных с использованием нейросетевой архитектуры, направленной на распознавание наземных объектов с учетом специфики радиолокационных данных и особенностей синтетического датасета;
- экспериментальная апробация разработанного нейросетевого алгоритма на расширенной тестовой выборке в условиях вариативности геометрических параметров съемки и оценка его вычислительной эффективности.

Материалы и методы

Схема трехпозиционной системы радиолокационного наблюдения

В основе предлагаемого подхода к радиолокационной съемке лежит трехпозиционная схема воздушного базирования на основе БВС самолетного типа. Носители оснащаются малогабаритными РСА сантиметрового диапазона волн (X-диапазон, $\lambda \approx 3$ см).

Система включает в себя две передающие позиции (*Tx*1, *Tx*2) и одну приемную позицию (*Rx*), пространственно разнесенные в горизонтальной плоскости. Данная распределенная конфигурация обеспечивает одновременное облучение земной поверхности с двух ракурсов и формирование двух разноракурсных радиолокационных кадров (РЛК) одной и той же сцены: кадра *GRF*1 от бистатической пары *Tx*1–*Rx* и кадра *GRF*2 от бистатической пары *Tx*2–*Rx* в проекции наклонной дальности на горизонтальную плоскость (*горизонтальная дальность*, англ. Ground Range).

На рис. 1 приведена трехпозиционная схема группового технического зрения, предназначенная для оперативного радиолокационного мониторинга земной поверхности. В предложенной схеме РЛК формируются в передних и переднебоковых зонах обзора бортовых РСА. По сравнению с традиционными однопозиционными РСА бокового обзора такая система позволяет раньше обнаружить цели по курсу движения БВС и сократить время подлета к целевой зоне поиска. Мультистатическая конфигурация обеспечивает повышенные значения пространственного разрешения и отношения «сигнал/шум», а также более полную информацию о наблюдаемой сцене за счет использования разноракурсной информации, что улучшает качество распознавания объектов интереса.

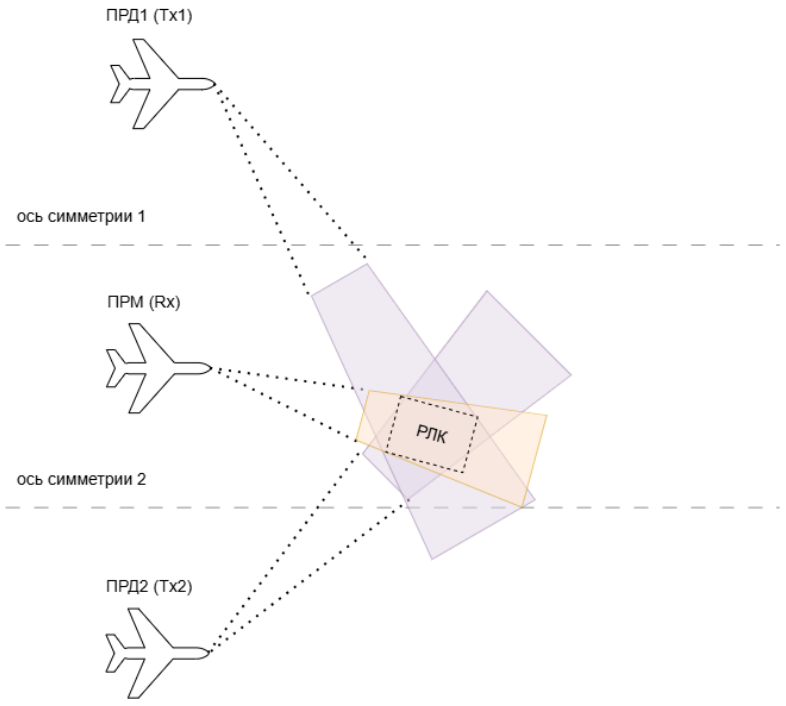


Рис. 1. Трехпозиционная схема группового технического зрения

Моделирование радиолокационных кадров

Формирование наборов входных радиолокационных данных для обучения, валидации и тестирования нейросетевого алгоритма реализовано на основе средств имитационного моделирования многопозиционной системы бортовых РСА переднебокового обзора. Программно-алгоритмическое ядро имитационной модели базируется на модифицированном методе растеризации [Кок26]. Моделирование включает вычисление первичного «сырого» сигнала (Raw Signal) с учетом геометрических параметров сцены и траектории носителей с последующим формированием РЛК алгоритмами цифровой обработки.

Алгоритм формирования РЛК в симуляторе включает следующие последовательные этапы:

- загрузка конфигурационных YAML-файлов, считывание двухмерных масок подстилающей поверхности и трехмерных полигональных моделей наземных объектов с последующим формированием единой 3D-модели радиолокационной сцены;
- задание параметров и режимов РСА, а также параметров траекторий БВС;
- дискретизация видимых элементов сцены в облако точек, имитирующее элементарные радиолокационные отражатели сцены;
- расчет карт видимости, глубин и нормалей элементов сцены с использованием модифицированного алгоритма растеризации;
- расчет комплексных амплитуд и фаз отраженных сигналов для каждого видимого элементарного отражателя с учетом взаимной геометрии объекта и многопозиционной системы наблюдения;
- формирование структурированной радиоголограммы (промежуточного комплексного представления радиолокационного сигнала);
- предобработка радиоголограммы и ее согласованная фильтрация (сжатие) по дальности;
- синтезирование РЛК на основе алгоритма обратной проекции (Back Projection) путем когерентного суммирования отраженных сигналов с учетом их фазы для каждого видимого элементарного отражателя.

Такой подход обеспечивает высокую физическую адекватность синтезируемых данных, корректно воспроизводя когерентную природу радиолокации и спекл-шум, возникающий вследствие интерференции полей от множества элементарных рассеивателей.

РЛК формируются в проекциях наклонной дальности (Slant Range) и горизонтальной дальности (Ground Range), также автоматически создается разметка для каждого наземного объекта на сцене и метаданные.

Применение программного симулятора для пакетной генерации синтетических РЛК позволяет контролируемо варьировать высоту полета носителей, ракурс обзора, тип подстилающей поверхности, насыщенность сцены наземными объектами различных классов.

Структура целевых классов

Постановка задачи ориентирована на 2D-детекцию и распознавание семи классов наземных объектов, объединенных в три укрупненные категории:

1. Транспортные средства:

- *Легковые автомобили* – седаны, хетчбэки, кроссоверы, внедорожники (SUV);
- *Грузовые автомобили и автобусы* – крупногабаритный колесный транспорт регулярной геометрической формы;
- *Специальная техника* – бульдозеры, экскаваторы, тракторы, автопогрузчики и пожарные машины со сложной конфигурацией металлических узлов и навесного оборудования.

2. Здания:

- *Частные здания* – небольшие одно- и двухэтажные жилые дома индивидуальной застройки с наклонными (скатными) крышами;

- *Многоквартирные и общественные здания* – многоэтажные и среднеэтажные здания, имеющие башенную, протяженную (линейную) или сложную пространственную конфигурацию;
- *Промышленные здания* – крупногабаритные ангары, склады и цеха, преимущественно с плоскими металлическими или бетонными перекрытиями.

3. Деревья:

- *Одиночные деревья* – отдельно стоящие лиственные и хвойные древесные насаждения.

Выбор указанного состава классов для задачи распознавания наземных объектов обусловлен типичным содержанием целевых сцен, содержащих антропогенные (транспортные средства, городская и сельская застройка) и природные (древесная растительность) объекты. Разделение на классы объясняется прежде всего принципиальными различиями в механизмах формирования их радиолокационных сигнатур.

Так, легковые автомобили формируют компактную сигнатуру с относительно низкой эффективной площадью рассеяния (ЭПР). Грузовые автомобили и автобусы выделяются более высокой ЭПР и выраженной прямоугольной формой за счет эффекта двойного отражения (*Double-bounce*) от вертикальных бортов и поверхности земли. Специальная техника характеризуется сложной, изрезанной сигнатурой, обусловленной наличием множества металлических навесных агрегатов, стрел и крановых механизмов, создающих эффекты многократного переотражения.

В группе классов зданий критически важным признаком на РЛК является характер взаимодействия волны с геометрией сооружения. Радиолокационный портрет частных зданий с наклонными (двухскатными, четырехскатными и т.д.) крышами определяется характерным пространственным сочетанием контрастных зон высокого отражения и затенения в сочетании с малой высотой стен. Многоквартирные и общественные здания распознаются по высокоинтенсивным линейным аномалиям, возникающим за счет эффекта двойного отскока в зоне стыка вертикальных стен с землей (аналогично работе диэдрального отражателя), а также по протяженным геометрическим теням. Промышленные здания идентифицируются по крупногабаритным геометрически правильным контурам, где обширные участки плоских кровель со стабильной текстурой, имеющие зеркальное отражение (низкая интенсивность обратного рассеяния), сочетаются с периодическими полосами от модульных элементов или шедовых (пилообразных) крыш, а также с точечными яркими отражателями от внешних коммуникаций.

Включение одиночных деревьев в качестве отдельного класса позволяет алгоритмам распознавания разделить механизмы объемного рассеяния (Volume scattering), характерные для крон деревьев, от сигналов, формируемых жесткими антропогенными отражателями (металлическими или бетонными конструкциями).

Формирование синтетического датасета

Сформированный синтетический массив РЛК базируется на подходе, основанном на объединении двух разноракурсных РЛК в проекции Ground Range (Ground Range Frame): *GRF*1 и *GRF*2, полученных симулятором от позиций *Tx*1–*Rx* и *Tx*2–*Rx*.

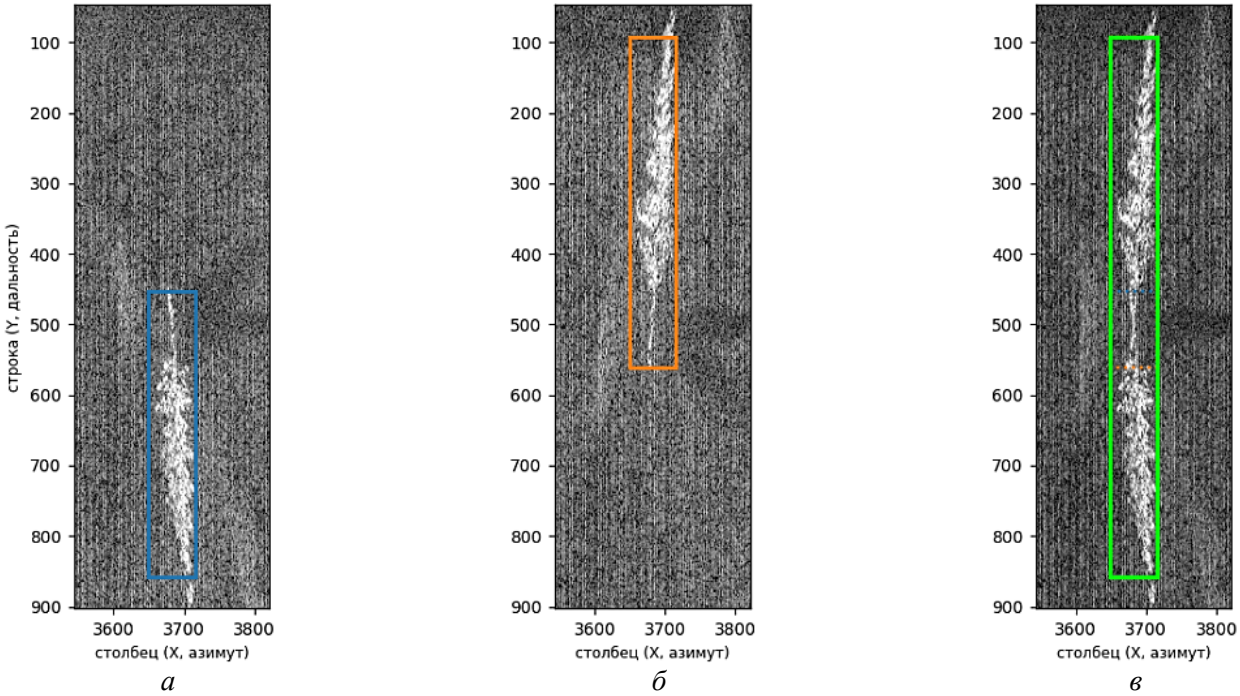
Для адаптации радиолокационных данных к современным нейросетевым детекторам, оптимизированным для обработки трехканальных тензоров, предложена схема прямого бескомпрессионного слияния в формат [*GRF*1, *GRF*2, *Null*], где:

- *GRF*1: монохромный РЛК с первого ракурса (1-й канал);
- *GRF*2: монохромный РЛК со второго ракурса (2-й канал);
- *Null:* канал с нулевыми значениями пикселей (3-й канал).

Обоснование данной компоновки базируется на сохранении независимой радиояркостной информации каждого ракурса. В отличие от эвристических методов компрессии в каналы статистик, например, в виде [max (*GRF*), min (*GRF*), mean (*GRF*)], предложенное представление

[*GRF*1, *GRF*2, *Null*] исключает взаимную деградацию локальных радиоконтрастов и сглаживание точечных рассеивателей, что по результатам предварительных экспериментов обеспечивает прирост валидационной метрики mAP_{50} с 0,916 до 0,942.

Перед объединением *GRF*2 подвергались повороту по горизонтали и вертикали для приведения к согласованной пространственной ориентации с *GRF*1, что обеспечивает корректное сопоставление объектов между ракурсами. В качестве итоговой аннотации было взято объединение аннотаций для *GRF*1 и *GRF*2, которое выполнялось по обеим осям. Пример объединения аннотаций приведен на рис. 2.



а *б* *в*

Рис. 2. Объединение аннотаций кадров *GRF*1 и *GRF*2 для объекта класса «Одиночное дерево»: *а* – аннотация *GRF*1; *б* – аннотация *GRF*2; *в* – аннотация объединенного РЛК

Для задачи распознавания семи выделенных классов наземных объектов с помощью симулятора сформирован датасет из РЛК в проекции *Ground Range*, включающий сцены, синтезированные для четырех основных типов подстилающей поверхности: болотистая местность (болота), луговая растительность (луг), сухой песок и вспаханные сельскохозяйственные угодья (пашня).

С целью повышения обобщающей способности нейросетевых моделей и обеспечения инвариантности детектора к пространственной ориентации целей, при генерации РЛК объекты сцены случайным образом поворачивались на произвольные углы в диапазоне [0–360°]. Кроме того, для приближения синтетических данных к реальным условиям наблюдения случайным образом варьировались пространственное распределение и плотность размещения объектов различных классов на генерируемых сценах, а также углы поворота объектов относительно оси визирования. Сводная статистика распределения объектов по классам в сформированном наборе данных приведена в табл. 1.

В обучающей и валидационной выборках размерность исходных данных составляет 1313×5981×3 пикселей (высота × ширина × число каналов), что соответствует высоте полета носителей 500 м. Для подачи в нейросетевую модель изображения разрезаются на квадратные фрагменты (слайсы) базовой размерности 1313×1313×3 пикселей. Общий исходный объем выборки для обучения и валидации составлял 1200 сцен (960 сцен в обучающей выборке и 240 – в валидационной, соотношение 80:20). После разбития на фрагменты объем выборки увеличился до 1852 сцен.

Таблица 1

Статическое распределение наземных объектов в синтетическом датасете по классам (суммарно для обучающей и валидационной выборок)

№ класса	Класс наземного объекта	Количество объектов	Доля от общего объема объектов (%)	Количество РЛК
1	Легковые автомобили	449	10,8	438
2	Грузовые автомобили и автобусы	424	10,2	422
3	Специальная техника	372	8,9	364
4	Частные здания	569	13,7	569
5	Многоквартирные и общественные здания	420	10,1	420
6	Промышленные здания	211	5,1	211
7	Одиночные деревья	1723	41,3	1200

Распределение типов подстилающих поверхностей как в обучающей, так и в валидационной выборках выдерживалось сбалансированным. Сводная статистика распределения типов подстилающей поверхности в сформированном наборе данных приведена в табл. 2.

Таблица 2

Статическое распределение типов подстилающей поверхности в синтетическом датасете по классам (суммарно для обучающей и валидационной выборок)

№ п/п	Тип подстилающей поверхности	Количество РЛК	Доля от общего объема РЛК (%)
1	Болото	317	26,4
2	Луг	312	26,0
3	Песок	286	23,8
4	Пашня	285	23,8
	Всего:	1200	100

Для независимого тестирования был формирован отдельный набор из 270 сцен (27×10 вариаций) с размерностями данных (526×3352×3), (1313×5981×3), (5251×19124×3) пикселей, в котором варьировались высота съемки (200, 500, 2000 м), ракурс разворота целей относительно оси визирования (0°, 45°, 90°, 135°, 180°) и тип подстилающей поверхности (песок, луг, пашня, болото) (табл. 3 и 4).

Таблица 3

Статическое распределение наземных объектов в тестовой выборке

№ класса	Класс наземного объекта	Количество объектов	Доля от общего объема объектов (%)	Количество РЛК
1	Легковые автомобили	4	14,8	4×10
2	Грузовые автомобили и автобусы	3	11,1	3×10
3	Специальная техника	3	11,1	3×10
4	Частные здания	4	14,8	4×10
5	Многоквартирные и общественные здания	4	14,8	4×10
6	Промышленные здания	2	7,4	2×10
7	Одиночные деревья	7	25.9	7×10
	Объекты всех классов	27	100	27×10

Таблица 4

Параметры формирования тестовой выборки

Номер серии	Высота носителя, м	Тип подстилающей поверхности	Угол ориентации объекта, град	Размерность данных (В × Ш × К)
1	2000	Песок	0	5251×19124×3
2	200	Песок	0	526×3352×3
3	500	Песок	0	1313×5981×3
4	500	Песок	45	1313×5981×3
5	500	Песок	90	1313×5981×3
6	500	Песок	135	1313×5981×3
7	500	Песок	180	1313×5981×3
8	500	Луг	0	1313×5981×3
9	500	Пашня	0	1313×5981×3
10	500	Болото	0	1313×5981×3

Архитектура нейросетевой модели

Традиционно для автоматического обнаружения и распознавания малоразмерных объектов на РЛИ применялись классические алгоритмы, основанные на поддержании постоянного уровня ложных тревог (Constant False Alarm Rate, CFAR) [El13]. Несмотря на высокую скорость работы, математические фильтры класса *CFAR* демонстрируют неудовлетворительную селективность в условиях неоднородного фона подстилающей поверхности и не способны осуществлять многоклассовую классификацию объектов. Переход к двухстадийным детекторам (таким как Faster R-CNN и его модификации) успешно решает задачу распознавания [Li17, Guo21], однако их двухэтапная архитектура (выделение регионов интереса с последующей классификацией) создает избыточную вычислительную нагрузку, практически исключающую обработку данных в реальном времени, что особенно критично для бортовых вычислителей БАС. Гибридные трансформерные архитектуры [Che23, Zha25] обеспечивают высокую точность за счет механизмов глобального внимания. Тем не менее, они обладают квадратичной вычислительной сложностью и неэффективными паттернами доступа к памяти, что приводит к критическому росту задержки (Latency) при анализе как больших РЛИ, так и последовательностей субапертурных РЛК, формируемых с повышенной частотой.

Для преодоления указанных ограничений в качестве базовой модели распознавания наземных объектов на сформированных РЛК была выбрана передовая одностадийная архитектура YOLOv12 [Tia25]. Модель выполняет извлечение признаков, определение пространственного положения целей (координат ограничивающих рамок) и классификацию обнаруженных объектов за один проход нейронной сети. Архитектура YOLOv12 включает механизм ориентированного внимания по областям (Area Attention) и эффективные блоки агрегации признаков, что позволяет учитывать пространственный контекст сцены при сохранении режима быстрого вывода.

По сравнению с классическими алгоритмами обнаружения с постоянным уровнем ложных тревог (CFAR) YOLOv12 не ограничивается статистическим анализом фона и обеспечивает одновременную локализацию и классификацию объектов. В отличие от двухстадийных детекторов, таких как Faster R-CNN, одноэтапная схема YOLO сокращает число последовательных операций и обеспечивает более высокую скорость обработки. По сравнению с трансформерными архитектурами YOLOv12 сохраняет способность моделировать пространственный контекст при меньших вычислительных затратах. Кроме того, за счет интеграции механизма внимания и сверточных нейронных сетей (CNN) архитектура 12-го поколения превосходит предыдущие версии (YOLOv10, YOLOv11), опирающиеся на локальные ядра свертки. Это преимущество достигается за счет способности улавливать глобальные контекстные связи

и сложные пространственные зависимости на изображении без существенного увеличения вычислительной сложности и объема параметров сети, что обеспечивает оптимальный баланс между точностью обнаружения и задержкой вывода [Wan26]. В задачах анализа РЛИ подобный подход помогает эффективно подавлять спекл-шум, отсекать радиолокационные переотражения и устойчиво выделять малоразмерные цели на сложном фоне.

В линейке YOLOv12 предусмотрено несколько вариантов архитектур разного масштаба: от облегченных до крупных и вычислительно тяжелых [Kha25, Wan26]. Для построения бортовой системы технического зрения БАС предварительно была выбрана модификация YOLOv12S. Этот предварительный выбор обусловлен результатами сравнительных экспериментов [Kha25], в ходе которых было установлено, что версия *S* функционирует существенно быстрее старших конфигураций (*M*, *L*, *X*), а также отсутствует статистически значимый прирост точности распознавания при переходе к этим более крупным и ресурсоемким архитектурам.

Конвейер ИИ-обработки (инференс)

Для автоматического распознавания наземных объектов на РЛК высокого разрешения разработан конвейер логического вывода (инференса), сочетающий YOLO-детектор, решающий задачу локализации и классификации объектов, и алгоритм фрагментарной обработки изображений методом скользящего окна. В основу алгоритма положена декомпозиция крупноформатных изображений на локальные фрагменты с последующим пространственным объединением векторных гипотез обнаружения. Такая схема позволяет снизить пропуски малоразмерных слабоконтрастных радиолокационных целей, которые могут возникать при масштабировании исходного изображения под фиксированный входной размер нейросети.

Последовательность выполнения вычислительных операций представлена на функциональной блок-схеме алгоритма обработки РЛК (рис. 3).

Далее опишем содержание и особенности каждого этапа обработки входного РЛК, представляющего собой трехканальное слияние двух разноракурсных РЛК.

Этап 1. Осуществляется загрузка входного РЛК, представленного трехканальной матрицей изображения с целочисленными значениями яркостных отсчетов в диапазоне [0; 255]. Пространственные размеры исходного РЛК могут различаться и учитываются при последующем формировании фрагментов и их подготовке к подаче в нейросетевую модель. Ограничение минимального размера исходного РЛК величиной 960 пикселей не устанавливается. При невозможности чтения изображения РЛК исключается из дальнейшей обработки.

Этап 2. Если пространственные размеры РЛК превышают 1313 пикселей хотя бы по одной из координатных осей, выполняется его разбиение на перекрывающиеся фрагменты с использованием скользящего окна базового размера 1313×1313 пикселей. Базовый шаг перемещения окна составляет 1051 пиксель, что соответствует перекрытию 262 пикселя, или приблизительно 20 % размера окна. Положение крайних окон корректируется по границам изображения для обеспечения полного покрытия РЛК. Если размер РЛК по одной из осей меньше 1313 пикселей, размер фрагмента по этой оси принимается равным соответствующему размеру РЛК. Например, для изображения размером 526×3352 пикселя формируются прямоугольные фрагменты размером 526×1313 пикселей. Если оба пространственных размера РЛК не превышают 1313 пикселей, он обрабатывается как единый фрагмент.

Этап 3. Каждый фрагмент масштабируется с помощью билинейной интерполяции с сохранением исходного соотношения сторон до достижения размера 960 пикселей по большей стороне. При необходимости масштабированное изображение дополняется полями постоянной яркости методом LetterBox[1]. Размеры полей определяются с учетом режима обработки и требований к кратности пространственных размеров входного тензора шагу нейросетевой модели. Поэтому подготовленный фрагмент может иметь как квадратную, так и прямоугольную

[1] Ultralytics YOLO. Data Augment API Reference. docs.ultralytics.com/reference/data/augment#ultralytics.data.augment.LetterBox.

форму. Затем выполняется поэлементное преобразование целочисленных значений яркостных отсчетов в вещественный тип данных с нормировочным делением на 255 и приведением к безразмерному интервалу [0; 1].

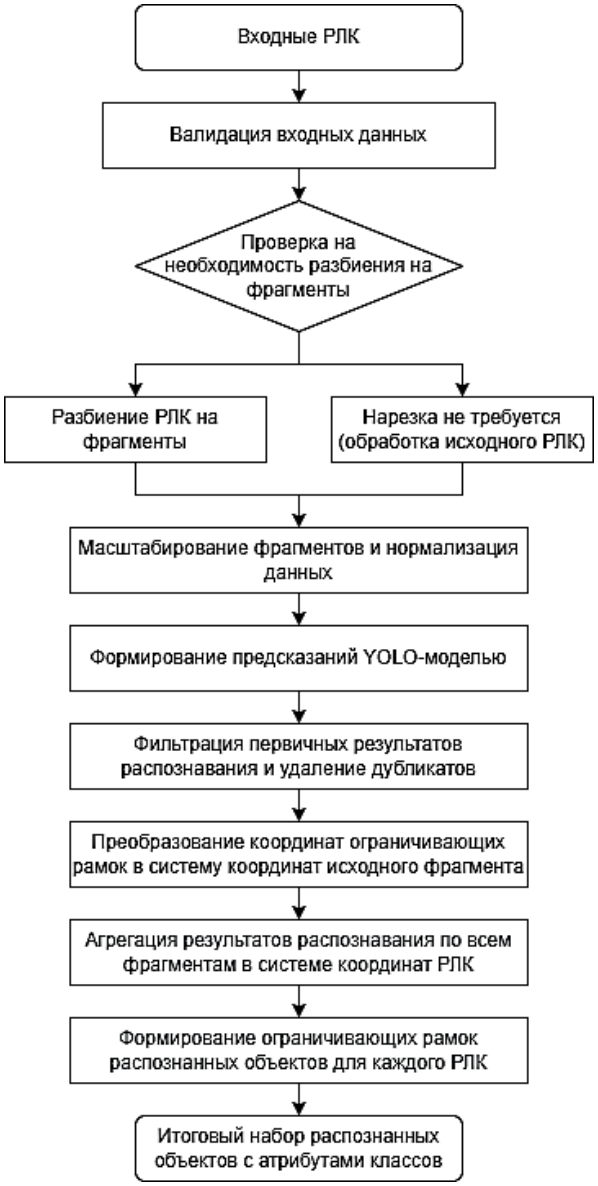


Рис. 3. Алгоритм распознавания наземных объектов на РЛК с применением нейросетевой модели YOLOv12S

Этап 4. Подготовленные и нормализованные фрагменты подаются на вход нейросетевой модели YOLOv12S, выполняющей извлечение признаков, локализацию и распознавание объектов одновременно по 7 целевым классам. Предсказания формируются в локальной системе координат входного изображения после масштабирования и добавления полей. Для каждого обнаруженного объекта определяются координаты ограничивающей рамки, идентификатор распознанного класса и значение уверенности модели в сформированной гипотезе обнаружения и классификации.

Этап 5. Отсечение шумовых откликов выполняется по порогу уверенности 0,1, что обеспечивает чувствительность к слабоконтрастным целям на сложном фоне. Для устранения дублирующих предсказаний внутри отдельных фрагментов применяется алгоритм подавления немаксимумов (*Non-Maximum Suppression, NMS*) [Ren16] с порогом пространственного перекрытия (*Intersection over Union, IoU*), равным 0,7.

Этап 6. Координаты ограничивающих рамок распознанных объектов, оставшихся после внутрифрагментной фильтрации, преобразуются из системы координат подготовленного входного изображения в систему координат исходного фрагмента. Для этого компенсируются смещения, обусловленные добавлением полей, после чего выполняется обратное масштабирование с учетом коэффициента, использованного при подготовке данного фрагмента. Полученные координаты ограничиваются его фактическими пространственными границами. Такое преобразование применяется как к квадратным фрагментам размером 1313×1313 пикселей, так и к прямоугольным фрагментам, например, фрагменты размером 526×1313 пикселей и т. д.

Этап 7. Координаты ограничивающих рамок распознанных объектов из системы координат одиночного фрагмента преобразуются в систему координат исходного РЛК с учетом смещения скользящего окна. Предсказания, полученные для всех фрагментов исходного РЛК, объединяются в общий набор результатов распознавания. Для устранения дублирующихся предсказаний в перекрывающихся областях применяется алгоритм жадного нелинейного слияния GreedyNMM (Greedy Non-Maximum Merging), реализованный в составе фреймворка SAHI (Slicing Aided Hyper Inference) [Aky22]. При использовании конфигурации GREEDYNMM/IOS@0.5 ограничивающие рамки, относящиеся к одному объекту, объединяются. Это позволяет корректно восстановить единую ограничивающую рамку объекта, разделенного границей соседних фрагментов. В результирующем наборе для каждого распознанного объекта сохраняются координаты ограничивающей рамки, установленный класс и значение уверенности модели. В отличие от метода немаксимального подавления (*NMS*), при котором ограничивающие рамки для предсказаний с меньшей уверенностью исключаются, GreedyNMM объединяет согласованные предсказания в итоговую ограничивающую рамку.

Этап 8. В пределах сформированных ограничивающих рамок для каждого распознанного объекта выполняется анализ значений интенсивности входного (объединенного) РЛК. На основе локальной нормализации мощности относительно общего фонового уровня и анализа двумерной пространственной автокорреляционной функции определяется масштаб одного независимого зерна спекл-шума. Далее выполняется адаптивное скользящее усреднение мощности с размером окна фильтрации, пропорциональным вычисленному размеру зерна, после чего отфильтрованные значения мощности переводятся в логарифмический масштаб. Уровень фонового шума подстилающей поверхности оценивается по моде гистограммы распределения, а верхний предел мощности – по высокому перцентилю, что обеспечивает робастность разделения к одиночным выбросам и крупным объектам. При наличии контраста порог бинаризации устанавливается посередине между фоном и пиком, а итоговые пространственные координаты локализованных ограничивающих рамок для распознанных объектов на каждом РЛК в паре формируются по профилям плотности ярких отсчетов вдоль строк и столбцов с отсечением краевых квантилей для устранения изолированных шумовых всплесков.

Таким образом, на выходе конвейера обработки формируется результирующий набор распознанных наземных объектов в системе координат РЛК, который далее передается в смежные программные блоки подсистемы технического зрения БАС для совместного отображения радиолокационных данных и других картографических слоев, а также для принятия дальнейшего решения в ходе мониторинга или поисковых работ.

Методика оценки качества распознавания

Количественная оценка качества распознавания наземных объектов на РЛК выполнялась путем сопоставления сформированных предсказаний с эталонной разметкой тестовых сцен с использованием стандартного критерия относительного геометрического перекрытия – площади пересечения над объединением ограничивающих рамок IoU с порогом не менее 50%. Данное пороговое условие являлось определяющим для расчета всей совокупности локальных показателей качества, формирующих базис для оценки эффективности модели по каждому классу наземных объектов.

Для детальной характеристики качества распознавания каждого из 7 целевых классов независимо применялись три базовые метрики: точность (*Precision*), полнота (*Recall*) и показатель

средней точности (AP_{50}). Точность позволяет оценивать частоту ошибок первого рода (ложноположительных срабатываний) при распознавании объектов конкретного типа на фоне подстилающей поверхности. Полнота характеризует способность модели минимизировать ошибки второго рода (пропуски целей), что критически важно для обнаружения малоразмерных объектов, чья радиолокационная сигнатура представлена небольшим количеством точечных откликов. Метрика средней точности класса (mAP_{50}) комплексно оценивает качество ранжирования предсказаний для конкретного типа объектов на основе площади под кривой «точность–полнота», интегрируя баланс между точностью и полнотой во всем диапазоне порогов уверенности модели.

Для формирования сводной интегральной оценки качества работы системы по всей совокупности целевых категорий выбрана стратегия макро-усреднения (метрика Macro Average). Подход Macro Average рассчитывает итоговые показатели как простое невзвешенное среднее арифметическое результатов локальных метрик по каждому классу, обеспечивая равную значимость всех 7 типов целей.

Результаты и их обсуждение

Разработанный алгоритмический комплекс на базе нейросетевой архитектуры YOLOv12 послужил основой для проведения серии сравнительных экспериментов по оценке эффективности алгоритмов распознавания наземных объектов. Вся серия вычислительных экспериментов проводилась при фиксированном стандартном пороге пространственного перекрытия ограничивающих рамок IoU ≥ 0,50.

Результаты анализа модификаций YOLOv12

Для подтверждения гипотезы о рациональности выбора облегченной модели для системы технического зрения БАС было проведено экспериментальное сопоставление 5 различных модификаций YOLOv12, отличающихся масштабом архитектуры (Extra-Large – X, Large – L, Medium – M, Small – S, Nano – N). Сравнительные измерения выполнялись на едином валидационном массиве, включающем 371 комплексную радиолокационную сцену, полученную после нарезки исходных 240 сцен на квадратные фрагменты, и содержащем в совокупности 1063 экземпляров объектов всех 7 распознаваемых классов. Численные значения интегральных макро-усредненных показателей оценки качества (Precision, Recall, mAP_{50}), полученных по результатам тестирования исследуемых конфигураций нейросети, приведены в табл. 5.

Таблица 5

Показатели качества распознавания наземных объектов для различных модификаций архитектуры YOLOv12

№ п/п	Модификация модели	Precision	Recall	mAP_{50}
1	YOLOv12X	0,931	0,874	0,939
2	YOLOv12L	0,915	0,877	0,932
3	YOLOv12M	0,910	0,891	0,941
4	**YOLOv12S**	**0,927**	**0,881**	**0,942**
5	YOLOv12N	0,915	0,867	0,927

Анализ экспериментальных данных табл. 5 показал, что максимальные значения интегрального показателя средней точности для всех целевым классов mAP_{50} были достигнуты при использовании модификации YOLOv12S (Small). При этом масштабирование архитектуры и переход к более вычислительно тяжелым и крупным конфигурациям (M, L и X) не привели к статистически значимому приросту показателей точности (Precision) и полноты (Recall). Полученные результаты подтвердили корректность предварительного выбора модификации YOLOv12S в качестве базовой архитектуры, поскольку она обеспечивает повышенную скорость инференса и минимальные требования к видеопамяти при сохранении уровня качества распознавания, близкого к максимальному.

Примеры визуализации результатов распознавания наземных объектов

Для иллюстрации результатов, формируемых на основе конвейера ИИ-обработки с использованием выбранной модели YOLOv12S, приведем пример распознавания наземных объектов на синтетической радиолокационной сцене по данным трехпозиционной системы радиолокационного наблюдения.

На рис. 4 представлена моделируемая сцена, содержащая наземные объекты нескольких целевых классов: частное здание, 2 легковых автомобиля (седан и внедорожник), грузовой автомобиль и два дерева. На рис. 5 показан входной трехканальный РЛК, сформированный путем объединения разноракурсных РЛК в проекции Ground Range, вместе с объединенными ограничивающими рамками распознанных объектов, агрегированными в единой системе координат сцены. На рис. 6 и 7 приведены РЛК для первого и второго ракурсов соответственно, на которых отображены локализованные ограничивающие рамки распознанных объектов.

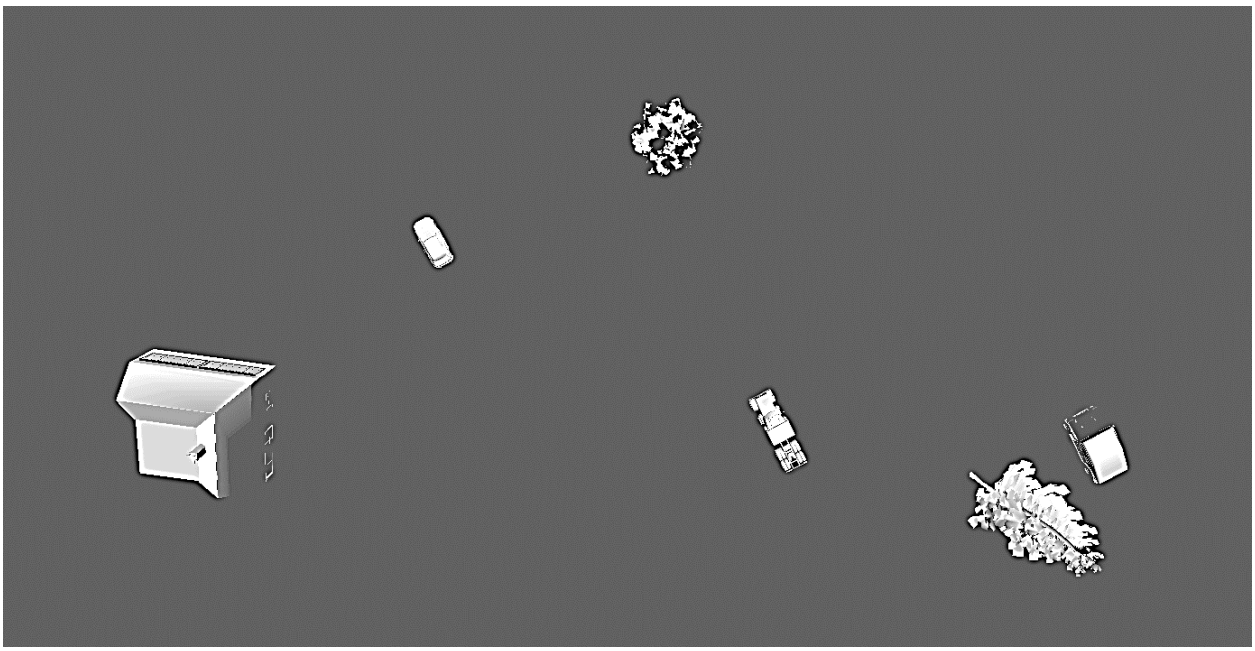


Рис. 4. Пример моделируемой сцены, содержащей наземные объекты 4 классов

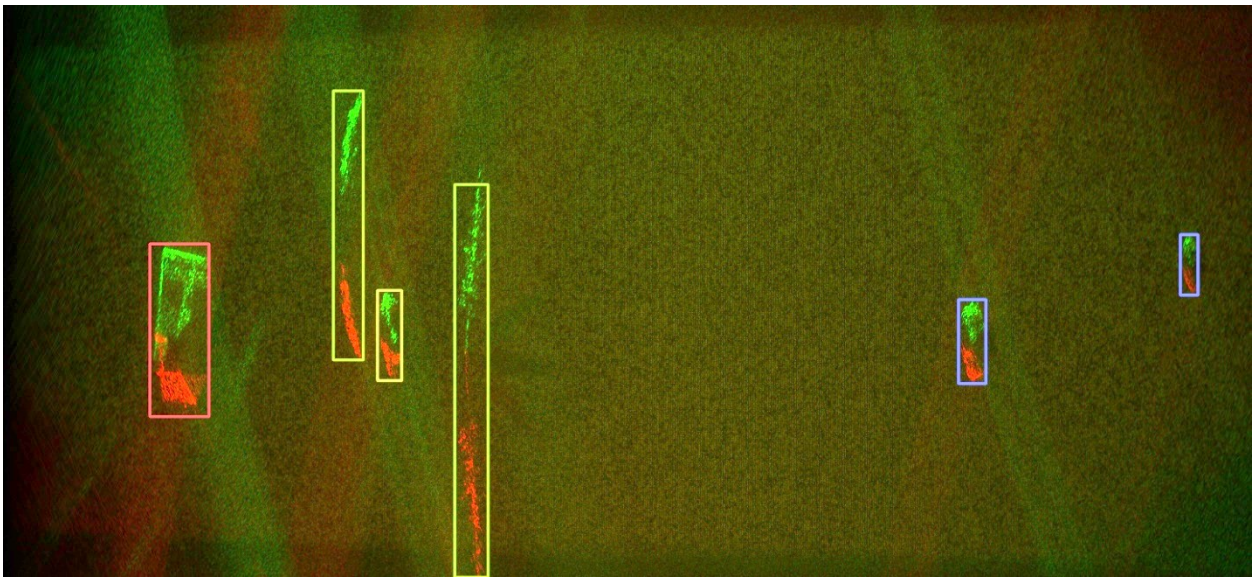


Рис. 5. Пример входного трехканального РЛК в RGB-палитре
с объединенными ограничивающими рамками распознанных объектов



Рис. 6. Пример монохромного РЛК для первого ракурса с локализованными ограничивающими рамками распознанных объектов

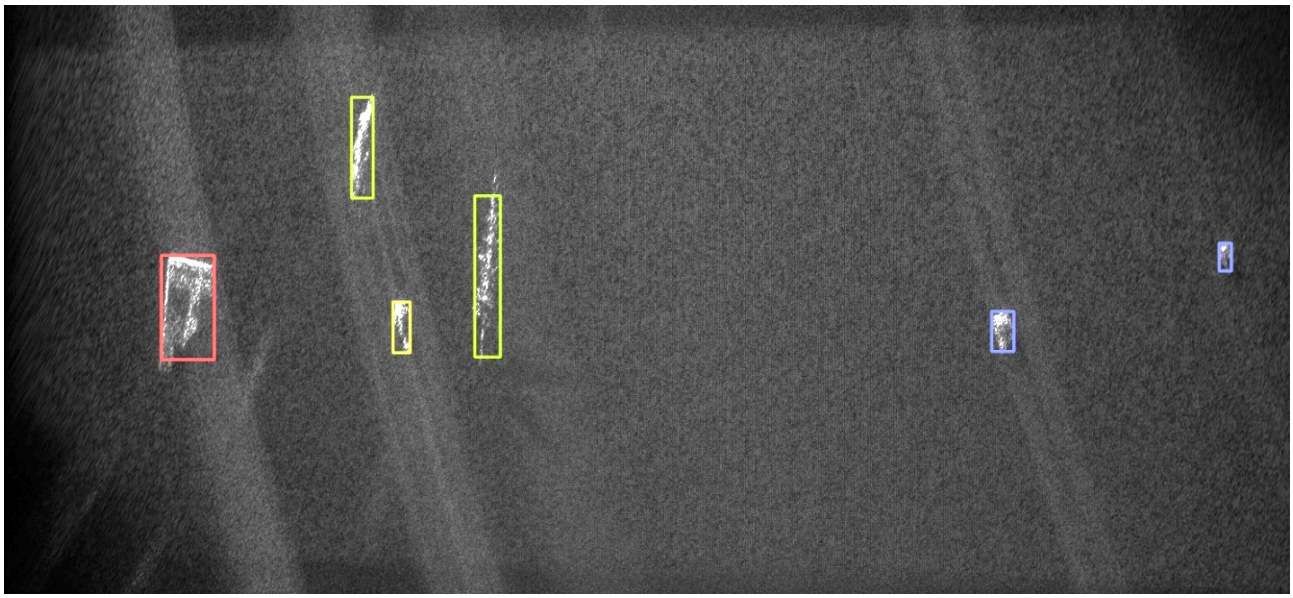


Рис. 7. Пример монохромного РЛК для второго ракурса с локализованными ограничивающими рамками распознанных объектов

Представленный пример демонстрирует корректную локализацию и классификацию объектов различных классов, включая транспортные средства, здания и одиночные деревья, при их совместном размещении на фоне подстилающей поверхности, имеющей тип «болото».

Оценка качества распознавания на валидационной выборке

Численные значения показателей Precision, Recall, AP_{50} для отдельных классов, а также интегральной метрики Macro Average по всем классам, полученные по завершении финальной эпохи оптимизации весов на основе валидационной выборки объемом 371 сцена, приведены в табл. 6.

Анализ данных табл. 6 демонстрирует выраженную дифференциацию качества распознавания модели в зависимости от класса наземного объекта и его отражательных свойств.

Таблица 6

Показатели качества распознавания целевых классов наземных объектов на валидационной выборке (конфигурация YOLOv12S)

№ класса	Класс наземного объекта	Кол-во сцен с классом	Кол-во объек-тов	Precision	Recall	AP_{50}
1	Легковые автомобили	111	112	0,974	0,848	0,951
2	Грузовые автомобили и автобусы	97	97	0,847	0,773	0,872
3	Специальная техника	86	87	0,837	0,827	0,907
4	Частные здания	156	156	0,972	0,901	0,979
5	Многоквартирные и общественные здания	97	97	0,919	0,938	0,950
6	Промышленные здания	64	64	0,969	0,962	0,972
7	Одиночные деревья	324	450	0,972	0,917	0,964
	***Macro Average* по всем классам**			**0,927**	**0,881**	**0,942** (mAP_{50})

Категории зданий и деревьев демонстрируют очень высокую точность распознавания (AP_{50} = 0,950–0,979). Для зданий это объясняется прежде всего четкой геометрически правильной формой крыш и устойчивым радиояркостным контрастом кровли и элементов стен в сочетании со специфической геометрией сопряженных зон радиолокационного затенения. Уверенное распознавание объектов категории «деревья» объясняется принципиальными физическими различиями между процессами объемного диффузного рассеяния, характерного для крон деревьев, и детерминированного рассеяния от жестких металлических или бетонных конструкций и антропогенных объектов. В категории транспортных средств наивысшая точность распознавания зафиксирована у легковых автомобилей (Precision = 0,974, AP_{50} = 0,951), тогда как грузовые автомобили и специальная техника показали умеренные значения (AP_{50} = 0,872 и 0,907 соответственно). Это возможно вызвано эффектом геометрического и радиолокационного сходства этих двух классов, в том время как отражательные характеристики легковых машин в виде компактных групп точечных рассеивателей локализуются нейросетевым алгоритмом достаточно уверенно.

Достигнутый интегральный показатель (mAP_{50} = 0,942) на валидационной выборке сопоставим с результатами известных современных аналогов, полученных на полусинтетических валидационных выборках. Так, в работе [Pan24] облегченная архитектура YOLOv5S, обученная на 2006 сценах размером 1478×1784 пикселей композитного датасета с разрешением 0,3 м, достигла интегрального показателя mAP_{50} = 0,958 по 10 классам военной техники без статистически значимого отставания от более тяжелых модификаций YOLOv5M, YOLOv5L и YOLOv7. Генерация многообъектных сцен выполнялась путем встраивания сегментированных целевых сигнатур, извлеченных из набора MSTAR в реальные РЛИ, содержащие различные типы фона, с помощью алгоритма SARBake. При этом трехмерные CAD-модели объектов использовались исключительно для создания базовой разметки (точные границы контура корпуса машины и падающей тени), а не для прямого моделирования сцен. В работе [Son23] модель CF-YOLO (модификация YOLOv5 с модулями внимания CAM и иерархической фильтрацией фона) на валидационной выборке полусинтетического датасета WSVD показала близкий показатель mAP_{50} = 0,9555 для самой тяжелой модификации. Набор данных WSVD был получен с помощью слияния целевых объектов из датасета MSTAR и сложного фона из датасета MiniSAR[2] с предварительной нормализацией яркостных распределений в цветовом пространстве HSV. Ограничением является то, что на основе реализованной архитектуры CF-YOLO решается только одноклассовая задача обнаружения («транспортное средство – фон»), а здания и деревья рассматриваются исключительно как мешающий фон. В работе

[2] Sandia National Laboratories. Complex SAR Data. www.sandia.gov/radar/pathfinder-radar-isr-and-synthetic-aperture-radar-sar-systems/complex-data-clone.

[Liu23] наилучший результат (mAP_{50} = 0,7537) на наборе данных Mix MSTAR, включающем в себя 20 классов транспортных средств, были получен для вычислительно сложной архитектуры Oriented RepPoints (модели YOLO в бенчмарк из 9 архитектур включены не были). Полусинтетический датасет Mix MSTAR был также получен слиянием целевых объектов и фоновых РЛИ, отличием его от двух предыдущих является расширенное количество объектов, сложные сценарии их размещения, а также ориентированные ограничивающие рамки. Из-за высокой межклассовой схожести модификаций транспортных средств итоговое качество распознавания на Mix MSTAR оказывается существенно ниже, чем на наборах, состоящих из 10 классов.

В отличие от всех перечисленных аналогов, ограниченных распознаванием исключительно наземной техники и рассматривающих здания и деревья как неразмеченный помеховый фон, разработанный конвейер на основе YOLOv12S решает более сложную задачу одновременной многоклассовой дифференциации антропогенных и природных объектов по 7 целевым классам. Незначительное расхождение интегрального показателя (mAP_{50} = 0,942) с результатами моделей YOLOv5s [Pan24] и CF-YOLO [Son23] (mAP_{50} = 0,955–0,958) объясняется переходом от локализованных полунатурных вырезок MSTAR, снятых при фиксированных углах визирования, к полномасштабным синтетическим сценам со свободным вращением объектов (0–360°) на четырех типах подстилающей поверхности. Достигнутый уровень точности подтверждает высокое качество сходимости модели в среде имитационного моделирования и обосновывает возможность перехода к следующему этапу исследований – оценке устойчивости алгоритма при тестировании на независимой выборке с комбинированным варьированием условий наблюдения.

Оценка качества распознавания на тестовой выборке

Экспериментальная оценка разработанного алгоритма распознавания выполнялась на независимой тестовой выборке, составляющей 18,4 % от общего объема синтетического датасета. Формирование тестовых поднаборов осуществлялось путем варьирования трех ключевых групп факторов: типа подстилающей поверхности, высоты съемки и азимутального ракурса объектов. Такая постановка эксперимента направлена на количественную оценку устойчивости нейросетевой модели к изменениям фоновой обстановки, пространственного масштаба и геометрии наблюдения. Количественные показатели качества работы созданной нейросетевой архитектуры для отдельных классов также оценивались по стандартным метрикам: Precision, Recall, AP_{50}. Итоговая интегральная эффективность рассчитывалась с использованием метрики Macro Average, обеспечивающей равную статистическую значимость всех классов независимо от их доли в выборке. Сводные результаты оценки качества по каждому из 7 целевых классов и общая оценка на основе макро-усреднения представлены в табл. 7.

Таблица 7

Показатели качества распознавания целевых классов наземных объектов на тестовой выборке (конфигурация YOLOv12S)

№ класса	Класс наземного объекта	Кол-во сцен с классом	Кол-во вариаций объектов	Precision	Recall	AP_{50}
1	Легковые автомобили	270	40	1,000	0,858	0,959
2	Грузовые автомобили и автобусы	270	30	0,963	0,872	0,958
3	Специальная техника	270	30	0,799	1,000	0,877
4	Частные здания	270	40	0,983	0,800	0,919
5	Многоквартирные и общественные здания	270	40	0,488	0,786	0,477
6	Промышленные здания	270	20	0,693	0,700	0,774
7	Одиночные деревья	270	70	0,984	0,986	0,990
	***Macro Average* по всем классам**			**0,844**	**0,857**	**0,851** (mAP_{50})

Анализ экспериментальных данных табл. 7 позволяет сделать следующие выводы. Интегральный показатель mAP_{50} = 0,851 на тестовой выборке подтверждает общую устойчивость модели YOLOv12S к варьированию ракурсов и других условий наблюдения, однако детальный анализ выявляет сильную зависимость качества распознавания от геометрических размеров и классов наземных объектов. В условиях невысокой насыщенности моделируемых сцен объектами модель демонстрирует максимальную эффективность при распознавании одиночных деревьев (AP_{50} = 0,990), характеризующихся стабильным объемным рассеянием крон. Высокая точность также достигнута для небольших локализованных объектов с четкими радиолокационными сигнатурами и тенями – легковых автомобилей (AP_{50} = 0,959), грузового транспорта и автобусов (AP_{50} = 0,958), частных малоэтажных домов (AP_{50} = 0,919) и специальной техники (AP_{50} = 0,877), поскольку механизм внимания нейросети уверенно выделяет их компактные центры рассеяния. В то же время качество распознавания резко падает на крупных или высотных объектах – промышленных (AP_{50} = 0,774) и особенно высотных многоквартирных (AP_{50} = 0,477) зданиях. Это в основном обусловлено спецификой наклонной геометрии радиолокационного визирования: поскольку обучение модели проводилось на данных, смоделированных для фиксированной высоты полета 500 м, значительное изменение высоты съемки носителя (от 200 до 2000 м) трансформирует угол падения луча и искажает эталонные пространственные признаки объектов (величину ракурсного наложения вертикальных стен и протяженность радиотеней), что приводит к появлению ложных срабатываний и снижению дифференцирующей способности нейросети.

В современных исследованиях по SAR ATR испытательные протоколы существенно различаются по числу и глубине возмущающих физических факторов, вводимых в независимую тестовую выборку. Как правило, используются режимы с расширенными условиями эксплуатации EOC (Extended Operating Conditions) и кросс-доменного тестирования (Cross-Domain). В режиме EOC проверяется способность нейросети корректно распознавать объекты интереса в условиях существенного отклонения параметров наблюдения, характеристик самого объекта или окружающей среды от тех, что использовались на этапе обучения. В режиме Cross-Domain тестируется качество работы модели при переходе из одного домена данных (генеральной совокупности распределения) в другой, где наблюдается Domain Shift (тестирование синтетических данных на данных натурной съемки, обучение на данных одного РЛС, а тестирование на данных другого). Также применяется совместное тестирование моностатической (однопозиционной) и бистатической (двухпозиционной) конфигураций РСА, направленное на создание универсальных или мультимодальных нейросетевых архитектур, способных эффективно распознавать объекты интереса независимо от геометрии приема.

В работе [Ye25] показано, что при тестировании нейросетевой архитектуры трансформерного типа (CWASA-Net) на моностатических (разрешение 0,1×0,1 м) и бистатических (разрешение 0,16×0,26 м) РЛИ, полученных с БВС в 9 полетных конфигурациях со сменой углов места (от 15° до 45°), бистатических азимутов и коэффициента усиления приемного тракта (в разные даты съемки), выход за пределы обучающей геометрии наблюдения приводит к снижению интегральной точности с внутридоменных $mAP_{50} > 0,9$ (0,901–0,909) до междоменных (Cross-Domain) $mAP_{50} > 0,8$ (0,809–0,904). Для обучения и тестирования использовались натурные наборы моностатических и бистатических РЛИ (11 классов военной техники), сформированные авторами исследования с помощью бортовой РЛС MiniSAR X-диапазона. В то же время в работе [Gen25], где тестирование выполнялось без варьирования базовых условий съемки с БВС, модель YOLOv8, обученная на смешанном наборе из 60% реальных и GAN-синтезированных бистатических РЛИ (сгенерированных на основе кросс-модального преобразования ИК-снимков) для 9 классов военной техники, достигла показателя mAP_{50} = 0,946. Для этой же модели, обученной исключительно на указанной выборке натурных РЛИ без привлечения синтетических данных, зафиксировано снижение показателя качества распознавания до mAP_{50} = 0,927. В работе [Liu25] для двухстадийной архитектуры ST-PA (Swin Transformer + Oriented R-CNN) приведены результаты распознавания на полусинтетическом датасете

MSTAR VDD, содержащем 5 укрупненных классов военной и специальной техники с разрешением 0,3 м. Эксперименты проводились в условиях стандартного внутридоменного тестирования (Standard Operating Condition, SOC) при идентичности параметров съемки, но с внедрением различных уровней многокритериального слияния признаков непосредственно в конвейер обучения и инференса модели. Благодаря этому было достигнуто высокое значение интегрального показателя качества распознавания mAP_{50}: 0,9572 для базовой попиксельной модификации нейросети, а при интеграции текстурных дескрипторов SAR-HOG и SIFT на уровне признаков точность возросла до 0,9783 и 0,9828 соответственно.

Для количественной оценки вклада объединения разноракурсных РЛК в повышение качества распознавания выполнен дополнительный контрольный эксперимент. В качестве исходных данных использована та же независимая тестовая выборка, что и при формировании результатов, приведенных в табл. 7, однако обработке подвергались не пары объединенных разноракурсных РЛК, а одиночные РЛК, полученные отдельно в парах *Tx*1–*Rx* и *Tx*2–*Rx*. Для каждого из двух ракурсов модель YOLOv12S обрабатывала одноканальный РЛК без использования информации о втором ракурсе. Полученные результаты по классам целевых объектов приведены в табл. 8.

Таблица 8

Показатели качества распознавания целевых классов наземных объектов на тестовой выборке при обработке одиночных РЛК 1-го и 2-го ракурсов (конфигурация YOLOv12S)

№ класса	Класс наземного объекта	РЛК 1-го ракурса (*Tx1–Rx*)			РЛК 2-го ракурса (*Tx2–Rx*)		
		Precision	Recall	AP_{50}	Precision	Recall	AP_{50}
1	Легковые автомобили	1,000	0,855	0,931	0,874	0,692	0,823
2	Грузовые автомобили и автобусы	0,739	0,600	0,761	0,777	0,933	0,948
3	Специальная техника	0,706	0,767	0,834	0,876	0,767	0,900
4	Частные здания	0,851	0,875	0,926	0,972	0,861	0,944
5	Многоквартирные и общественные здания	0,427	0,709	0,511	0,589	0,750	0,542
6	Промышленные здания	0,854	0,800	0,894	0,620	0,500	0,501
7	Одиночные деревья	0,901	0,886	0,939	0,926	0,892	0,943
	Macro Average по всем классам	**0,783**	**0,785**	**0,828** (mAP_{50})	**0,805**	**0,771**	**0,800** (mAP_{50})

Анализ таблица 8 показывает, что качество распознавания при обработке одиночных РЛК оказалось в целом ниже, чем при использовании пары объединенных разноракурсных РЛК. Это подтверждает, что предложенный способ формирования трехканального входного изображения на основе РЛК, полученных в парах *Tx*1–*Rx* и *Tx*2–*Rx*, обеспечивает дополнительный прирост качества за счет совместного использования радиолокационных признаков, наблюдаемых с двух ракурсов. Объединение разноракурсных данных позволяет компенсировать ракурсную зависимость радиолокационных сигнатур объектов, уменьшить влияние затенения и частичной потери информативных признаков на одном из ракурсов, а также повысить устойчивость локализации и распознавания объектов при изменении их ориентации относительно оси визирования.

Результаты оценки вычислительной эффективности

Для подтверждения применимости разработанного конвейера обработки РЛК в составе наземного комплекса обработки данных БАС были проведены замеры времени прямого логического вывода (инференса) модели YOLOv12S. Аппаратное тестирование выполнялось на двух типах вычислительных платформ – серверной и мобильной (ноутбук) – с раздельной оценкой производительности на центральных процессорах (CPU) и графических ускорителях (GPU). Оценка среднего времени задержки производилась по серии из 200 независимых запусков для каждого сценария, с использованием РЛК размером 960 × 960 пикселей. Сводные результаты аппаратных измерений представлены в табл. 9.

Таблица 9

Результаты оценки вычислительной эффективности нейросетевой модели YOLOv12S

№ п/п	Тип вычислительной платформы	Конфигурация	Среднее время инференса по 200 запускам, мс	Эквивалентная частота обработки, кадров/с
1	Серверная	1 CPU AMD EPYC 7543	584	1,7
2	Серверная	1 GPU NVIDIA A100 (80 Гб)	10	100,0
3	Мобильная (ноутбук)	1 CPU Intel Core i7-12700H (2,30 GHz)	439	2,3
4	Мобильная (ноутбук)	1 GPU NVIDIA GeForce RTX 3070 Ti Laptop (8 GB)	29	34,5

Полученные результаты свидетельствуют о том, что архитектура YOLOv12S в связке с предложенной схемой параллельной обработки фрагментов на GPU обеспечивает высокий темп обработки РЛК и с запасом перекрывает требования для подсистемы технического зрения БАС (частота 12-15 кадров/с, близкая к видеочастоте). Высокие показатели быстродействия обусловлены одностадийной организацией архитектуры YOLOv12S и линейной масштабируемостью механизма ориентированного внимания по областям, оптимизированного под эффективную тензорную параллелизацию без создания избыточной нагрузки на шину памяти. Однако итоговая сквозная производительность всего конвейера обработки данных для разрабатываемого прототипа многопозиционной системы лимитируется не только временем прямого логического вывода (инференса), но и скоростью формирования и выдачи РЛК в радиолокационном канале. Тем не менее, достигнутый запас быстродействия нейросетевого алгоритма на GPU гарантирует, что блок ИИ-обработки не станет узким местом при интеграции с бортовыми вычислителями.

Заключение

Разработаны и экспериментально апробированы алгоритмические решения для распознавания наземных объектов на РЛК, формируемых в трехпозиционной пространственно-распределенной системе бортовых РСА. Основной научный вклад работы состоит в разработке комплексного подхода, основанного на имитационном моделировании радиолокационных сцен и синтезе РЛК в программном симуляторе, объединении пары разноракурсных РЛК в едином трехканальном представлении и реализации высокопроизводительного конвейера логического вывода с применением облегченной нейросетевой архитектуры YOLOv12S. Автоматизированное формирование синтетического датасета позволило решить проблему дефицита размеченных данных и обеспечить учет изменчивости фоно-целевой обстановки и геометрии радиолокационного наблюдения. Сочетание механизма ориентированного внимания архитектуры YOLO 12-го поколения с пофрагментной обработкой кадров методом скользящего окна, пространственным объединением предсказаний соседних фрагментов и постобработкой границ объектов по ярким отсчетам обеспечило многоклассовое распознавание наземных объектов при сохранении высокой вычислительной эффективности, свойственной одностадийным детекторам. Экспериментальная верификация подтвердила высокую дифференцирующую способность разработанного алгоритма при одновременном распознавании 7 классов транспортных средств, зданий и деревьев на разнородных фонах: интегральное значение показателя *mAP*50 составило 0,942 на валидационной выборке и сохранилось на приемлемом уровне 0,851 на независимой тестовой выборке в условиях варьирования высоты съемки, ориентации объектов и типов подстилающей поверхности. Существенным фактором достижения высокого качества распознавания стало управляемое формирование обучающей выборки с использованием программного симулятора. Перспективы дальнейших исследований связаны с расширением номенклатуры фонов и геометрических условий съемки при имитационном моделировании, проверкой метода на экспериментальных натурных радиолокационных данных, а также его адаптацией к иным частотным диапазонам и конфигурациям многопозиционных систем.

Благодарности и поддержка

Работа выполнена на средства гранта, предоставленного по договору от 28.05.2026 № 70-2026-001129 между АНО ВО «Университет Иннополис» и Фондом поддержки проектов Национальной технологической инициативы.

Список литературы | References

[Aky22] Akyon F. C., Altinuc S. O., Temizel A. Slicing aided hyper inference and fine-tuning for small object detection // 2022 IEEE Int. Conf. Image Processing (ICIP). Bordeaux, France, 2022. pp. 966-970. 10.1109/ICIP46576.2022.9897990.

[Aue16] Auer S., Bamler R., Reinartz P. RaySAR – 3D SAR simulator: Now open source // 2016 IEEE Int. Geoscience and Remote Sensing Symp. (IGARSS). Beijing, China, 2016. pp. 6730-6733. 10.1109/IGARSS.2016.7730757.

[Bal06] Balz T. Real-time SAR simulation on graphics processing units // Proc. 6th European Conf. Synthetic Aperture Radar (EUSAR 2006). Dresden, Germany, 2006. www.vdeverlag.de/proceedings-de/562960184.html.

[Bla20] Blasch E. et al. Review of recent advances in AI/ML using the MSTAR data // Proc. SPIE 11393, Algorithms for Synthetic Aperture Radar Imagery XXVII, 113930C (19 May 2020), 2020. pp. 53-63. 10.1117/12.2559035.

[Che16] Chen S. et al. Target classification using the deep convolutional networks for SAR images // IEEE Trans. Geoscience and Remote Sensing. 2016. Vol. 54, No. 8. pp. 4806-4817. 10.1109/TGRS.2016.2551720.

[Che23] Chen Y., Cong Y., Zhang L. Deformable scattering feature correlation network for aircraft detection in SAR images // IEEE Geoscience and Remote Sensing Letters. 2023. Vol. 20. Art No. 4007205. pp. 1-5. 10.1109/LGRS.2023.3292243.

[El13] El-Darymli K. et al. Target detection in synthetic aperture radar imagery: a state-of-the-art survey // J. Applied Remote Sensing. SPIE, 2013. Vol. 7, No. 1. p. 071598. 10.1117/1.JRS.7.071598.

[Gen25] Geng Z. et al. Vehicle target detection based on cross-modality projective-invariant features extracted from unpaired SAR and infrared images // Electronics Letters. 2025. Vol. 61, No. 1. p. e70336. 10.1049/ell2.70336.

[Gro22] Gromada K. et al. Real-Time object detection and classification by UAV equipped with SAR // Sensors. 2022. Vol. 22, No. 5. p. 2068. 10.3390/s22052068.

[Guo21] Guo Y., Du L., Lyu G. SAR target detection based on domain adaptive faster R-CNN with small training data size // Remote Sensing. 2021. Vol. 13, No. 21. p. 4202. URL: 10.3390/rs13214202.

[Hen13] Henderson F. M., Lewis A. J. (Eds.) Principles & Applications of Imaging Radar 3Ed: Manual of Remote Sensing. New York: Wiley, 2013. 896 p.

[Kan25] Kanz J. et al. Coherent Multistatic UAV-Based SAR // IEEE Trans. Radar Systems. 2025. Vol. 3. pp. 1056-1072. 10.1109/TRS.2025.3593818.

[Kha25] Khanam R., Hussain M. A Review of YOLOv12: Attention-Based Enhancements vs. Previous Versions // arXiv preprint. 2025. 10.48550/arXiv.2504.11995.

[Lan25] Lang P. et al. Recent Advances in Deep-Learning-Based SAR Image Target Detection and Recognition // IEEE J. Selected Topics in Applied Earth Observations and Remote Sensing. 2025. Vol. 18. pp. 6884-6915. 10.1109/JSTARS.2025.3543531.

[Li17] Li J., Qu C., Shao J. Ship detection in SAR images based on an improved faster R-CNN // 2017 SAR in Big Data Era: Models, Methods and Applications (BIGSARDATA). Beijing, China, 2017. pp. 1-6. 10.1109/BIGSARDATA.2017.8124934.

[Li22] Li J. et al. An Accelerated Hybrid Method for Electromagnetic Scattering of a Composite Target-Ground Model and Its Spotlight SAR Image // Remote Sensing. 2022. Vol. 14, No. 24. p. 6332. 10.3390/rs14246332.

[Li23] Li J. et al. A Comprehensive Survey on SAR ATR in Deep-Learning Era // Remote Sensing. 2023. Vol. 15, No. 5. p. 1454. 10.3390/rs15051454.

[Li24] Li Y. et al. SARDet-100K: Towards Open-Source Benchmark and ToolKit for Large-Scale SAR Object Detection // arXiv preprint. 2024. 10.48550/arXiv.2403.06534.

[Lin23] Lin X. et al. SIVED: A SAR Image Dataset for Vehicle Detection Based on Rotatable Bounding Box // Remote Sensing. 2023. Vol. 15, No. 11. p. 2825. 10.3390/rs15112825.

[Liu23] 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. p. 4558. 10.3390/rs15184558.

[Liu25] Liu Q. et al. Intelligent Target Detection in Synthetic Aperture Radar Images Based on Multi-Level Fusion // Remote Sensing. 2025. Vol. 17, No. 1. p. 112. 10.3390/rs17010112.

[Liu25b] Liu Y. et al. ATRNet-STAR: A large dataset and benchmark towards remote sensing object recognition in the wild // IEEE Trans. Pattern Analysis and Machine Intelligence. 2025. Vol. 48, No. 6. pp. 6735-6753. 10.1109/TPAMI.2026.3658649.

[Mor26] Moreira A. et al. Spaceborne Synthetic Aperture Radar: Future Technologies and Mission Concepts // Proc. IEEE. 2026. Vol. 114, No. 2. pp. 183-219. 10.1109/JPROC.2025.3621586.

[Pan24] Panati C., Wagner S., Brüggenwirth S. Multiple Target Recognition Within SAR Scene Achieved Using YOLO and Explanability Investigated Using Gradient-Free Visualisation // 2024 IEEE Radar Conf. (RadarConf24). Denver, CO, USA, 2024. pp. 1-6. 10.1109/RadarConf2458775.2024.10548088.

[Qiu23] Qiu L. et al. YOLO-based multi-class target detection for SAR images // 2023 IEEE 7th Information Technology and Mechatronics Engineering Conf. (ITOEC). 2023. Vol. 7. pp. 1671-1675. 10.1109/ITOEC57671.2023.10292083.

[Ren16] Ren S. et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks // arXiv preprint. 2016. 10.48550/arXiv.1506.01497.

[Roc25] Rocha R. d. L., Figueiredo F. A. P. d. Enhancing YOLO-Based SAR Ship Detection with Attention Mechanisms // Remote Sensing. 2025. Vol. 17, No. 18. p. 3170. 10.3390/rs17183170.

[Son23] Song Y. et al. Vehicle Target Detection Method for Wide-Area SAR Images Based on Coarse-Grained Judgment and Fine-Grained Detection // Remote Sensing. 2023. Vol. 15, No. 13. p. 3242. 10.3390/rs15133242.

[Son25] Song D. et al. Coverage Path Planning for Multi-view SAR-UAV Observation System under Energy Constraint // arXiv preprint. 2505. 10.48550/arXiv.2505.16389.

[Sun24] Sun X. et al. SAR Vehicle Image Generation with Integrated Deep Imaging Geometric Information // Int. J. Applied Earth Observation and Geoinformation. 2024. Vol. 132. p. 104028. 10.1016/j.jag.2024.104028.

[Tia25] Tian Y., Ye Q., Doermann D. YOLOv12: Attention-Centric Real-Time Object Detectors // arXiv preprint. 2025. 10.48550/arXiv.2502.12524.

[Wan26] Wang W. et al. An Improved YOLOv12-Based Object Detection Model for Ship Monitoring in SAR Images // Chinese Journal of Information Fusion. 2026. Vol. 3, No. 2. pp. 125-137. 10.62762/CJIF.2025.869982.

[Xu20] Xu F. et al. UAV Intelligent Optimal Path Planning Method for Distributed Radar Short-Time Aperture Synthesis // IGARSS 2020 – 2020 IEEE Int. Geoscience and Remote Sensing Symp. Waikoloa, HI, USA. 2020. pp. 469-472. 10.1109/IGARSS39084.2020.9324684.

[Ye25] Ye Z., Zhou P. Domain-Unified Adaptive Detection Framework for Small Vehicle Targets in Monostatic/Bistatic SAR Images // Remote Sensing. 2025. Vol. 17, No. 22. p. 3671. 10.3390/rs17223671.

[Zha23] Zhao J., An D., Chen L. Man-Made Target Extraction Based on Multi-Aspect SAR Images // 2023 8th Int. Conf. Image, Vision and Computing (ICIVC). Dalian, China. 2023. pp. 472-477. URL: 10.1109/ICIVC58118.2023.10270403.

[Zha25] Zhao G. et al. Towards SAR Automatic Target Recognition MultiCategory SAR Image Classification Based on Light Weight Vision Transformer // arXiv preprint. 2025. 10.48550/arXiv.2407.06128.

[Zha26] Zhang F. et al. GeoDiff-SAR: A Geometric Prior Guided Diffusion Model for SAR Image Generation // arXiv preprint. 2026. 10.48550/arXiv.2601.03499.

[Вер10] Верба В. С. и др. Радиолокационные системы землеобзора космического базирования. М: Радиотехника, 2010. 680 с. QKJNFV. [Verba V. S. et al. Space-Based Land Survey Radar Systems. M: Radio Engineering, 2010. (In Ross.).]

[Кок25] Кокутин С. Н., Сергиев С. А. Методы интеллектуального анализа радиолокационных изображений, получаемых с бортовых авиационных систем // СИИТ. 2025. Т. 7. № 4 (23). С. 124-154. TTBCME. [Kokutin S. N., Sergiev S. A. Methods for intelligent analysis of radar images obtained from on-board aviation systems // SIIT. 2025. T. 7. No. 4 (23). pp. 124-154. (In Ross.). TTBCME.]

[Кок26] Кокутин С. Н., Сергиев С. А. Автоматизированная генерация синтетических радиолокационных изображений при разработке системы группового технического зрения авиационного базирования // Радиоэлектронные устройства и системы для инфокоммуникационных технологий (РЭУС-ИТ 2026): Всеросс. конф. Москва, 20-21 мая 2026 г. М.: РНТОРЭС им. А. С. Попова, 2026. С. 148-153. [Kokutin S. N., Sergiev S. A. Automated generation of synthetic radar images in the development of an aviation-based group technical vision system // Radio-Electronic Devices and Systems for Infocommunication Technologies (REUS-IT 2026): All-Russian. Conf. Moscow, May 20-21, 2026 M.: RNTORES, 2026. P. 148-153. (In Ross.).]

[Ком26] Кокутин С. Н., Муслимов Т. З. и др. Комплекс группового управления беспилотными воздушными судами многопозиционной системы радиолокационного мониторинга: пат. 2861462 РФ. 2026. Бюл. № 13. LHQCKI. [Kokutin S. N., Muslimov T. Z., et al. Complex of group control of unmanned aerial vehicles of a multi-position radar monitoring system: Pat. 2861462 Russian Federation. 2026. Bull. No. 13. (In Ross.). LHQCKI.]

[Фед22] Федотовских А. В. Особенности разработки и эксплуатации гражданских беспилотных авиационных систем с технологиями искусственного интеллекта в Арктической зоне Российской Федерации. М: Ай Пи Ар Медиа, 2022. 277 с. KFLVAN. [Fedotovskikh A.V. Features of the Development and Operation of Civil Unmanned Aircraft Systems with Artificial Intelligence Technologies in the Arctic Zone of the Russian Federation. M: IP Ar Media, 2022. (In Ross.). KFLVAN.]

ОБ АВТОРАХ | ABOUT THE AUTHORS

КОКУТИН Сергей Николаевич
Университет Иннополис, Россия.
Вед науч. сотр. отдела систем видеоаналитики Ин-та искусств. интеллекта. Дипл. инж. оптик-конструктор (Казанск. нац. иссл. тех. ун-т). Канд. геол.-мин. наук по геоэкологии (Уральск. гос. горный ун-т). Иссл. в обл. дистанц. зондирования Земли.
E-mail: s.kokutin@innopolis.ru.
ORCID: 0009-0007-6474-2838.



KOKUTIN Sergey Nikolaevich
Innopolis University, Russia.
Lead. Researcher, Dept. of Video Analytics Systems, Institute of AI. Dipl. Optical engineer-designer (Kazan National Research Tech. Univ.). Cand. of Tech. Sci. on GeoEcology. (Ural State Mining Univ.). Research in Earth Remote Sensing methods.
E-mail: s.kokutin@innopolis.ru.
ORCID: 0009-0007-6474-2838.

МЕРГАЛИЕВ Ибрагим Расулович
Университет Иннополис, Россия.
Гл. программист-математик Центра разработки и проектирования ПО Ин-та искусственного интеллекта. Дипл. спец. по информатике и выч. технике (Ун-т Иннополис). Иссл. в обл. компьютерного зрения и глубокого обучения.
E-mail: i.mergaliev@innopolis.ru



MERGALIEV Ibragim Rasulovich
Innopolis University, Russia.
Sr. Programmer-mathematician, Software Design and Development Center, Institute of AI. Dipl. Informatics and computer engineering (Innopolis University). Research in the field of computer vision and deep learning.
E-mail: i.mergaliev@innopolis.ru

МЕТАДАННЫЕ | METADATA

Заглавие: Распознавание наземных объектов по разноракурсным данным многопозиционной системы бортовых РСА: разработка и апробация нейросетевого алгоритма на основе имитационного моделирования.

Авторы: Кокутин С. Н., Мергалиев И. Р.

Аннотация: Автоматическое распознавание наземных объектов по радиолокационным изображениям, получаемым с борта беспилотных воздушных судов, требует учета сложной фоно-целевой обстановки и выраженной ракурсной изменчивости радиолокационных сигнатур целей, при этом для подходов на основе глубокого обучения характерен острый дефицит обучающих наборов данных. В работе представлены результаты реализации и экспериментальной апробации нейросетевого алгоритма распознавания неподвижных наземных объектов на радиолокационных кадрах, формируемых в многопозиционной системе малогабаритных бортовых радиолокаторов с синтезированной апертурой X-диапазона, в качестве носителей которых используется группа беспилотных воздушных судов самолетного типа. Предложенный подход базируется на аппарате имитационного моделирования, включающем генерацию синтетических радиолокационных кадров посредством программного симулятора на основе модифицированного алгоритма растеризации и слияние в единое трехканальное изображение двух разноракурсных кадров, получаемых в проекции горизонтальной дальности. Сформирован расширенный синтетический датасет для семи целевых классов, относящихся к категориям транспортных средств, зданий и деревьев. Датасет создан путем управляемого варьирования состава, взаимного расположения и ориентации наземных объектов, а также типа подстилающей поверхности. Реализован конвейер обработки на базе облегченной одностадийной архитектуры YOLOv12S и механизма нарезки входного объединенного радиолокационного кадра на перекрывающиеся фрагменты фиксированного размера с последующим пространственным объединением локальных предсказаний в единый набор результатов распознавания. Результаты экспериментов продемонстрировали эффективность предложенного решения при распознавании наземных объектов в различных условиях наблюдения и подтвердили его потенциальную применимость в подсистемах технического зрения беспилотных авиационных систем.

Ключевые слова: Радиолокатор с синтезированной апертурой; многопозиционная радиолокационная система; беспилотное воздушное судно; имитационное моделирование; синтетический датасет; радиолокационный кадр; нейросетевой алгоритм; распознавание наземных объектов.

Язык: Русский.

Статья поступила в редакцию 22 июля 2026 г.

Title: Ground object recognition from multi-aspect data of a multistatic onboard SAR system: development and experimental evaluation of a neural network algorithm based on simulation modeling.

Authors: Kokutin S. N., Mergaliev I. R.

Abstract: Abstract: Automatic recognition of ground objects from radar images acquired by unmanned aerial vehicles requires accounting for a complex background-target environment and pronounced aspect-dependent variability in target radar signatures, while deep-learning-based approaches suffer from a severe shortage of training datasets. This paper presents the implementation and experimental evaluation of a neural network algorithm for stationary ground object recognition in radar frames generated by a multistatic system of compact onboard X-band synthetic aperture radars, with a group of fixed-wing unmanned aerial vehicles serving as carrier platforms. The proposed approach relies on a simulation-modeling framework that incorporates the generation of synthetic radar frames using a software simulator based on a modified rasterization algorithm, and the fusion of two multi-aspect frames obtained in ground-range projection into a single three-channel image. An extended synthetic dataset was created for seven target classes belonging to the categories of vehicles, buildings, and trees. The dataset was generated by controlled variation of the composition, mutual arrangement, and orientation of ground objects, as well as the type of underlying surface. A processing pipeline was implemented based on the lightweight single-stage YOLOv12S architecture and a mechanism for partitioning the input merged radar frame into fixed-size overlapping tiles, followed by spatial aggregation of local predictions into a unified set of recognition results. Experimental results demonstrated the effectiveness of the proposed solution for ground objects recognition under various observation conditions and confirmed its potential applicability in machine-vision subsystems of unmanned aircraft systems.

Key words: Synthetic aperture radar; multistatic radar system; unmanned aerial vehicle; simulation modeling; synthetic dataset; radar frame; neural network algorithm; ground object recognition.

Language: Russian.

The article was received by the editors on 22 July 2026.